On the Soliton Resolution Conjecture for Wave Maps

Roland Grinis

Merton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2016

On the Soliton Resolution Conjecture for Wave Maps

Roland Grinis

November 25, 2016

Acknowledgements

I dedicate this thesis to my family: my parents Vadim and Gala, my grand-mother Tamara and my sister Inna. Your infinite love, support and care are the most marvelous things I have. You have been unconditionally with me during all those three years of mathematical research at Oxford. Even after hundreds of failures, you were always here to convince me to keep on building new approaches to the puzzles I was trying to break and to seek new insights in the theories I was developing. This thesis would not have existed without you.

My thesis is written in memory of my grand-parents who have passed away: Evgeniya and Inna Grinis, and Artem Hodjaev, whose love will be with me forever.

Abstract

Broadly speaking, the research presented in this thesis is centered around the study of the Soliton Resolution Conjecture (SRC) for the wave maps equation in dimension 2+1, which is rooted in a belief held by the physics community since the 1970's predicting that for a large class of non-linear dispersive/hyperbolic evolution equations in mathematical physics (of which wave maps are an example), the solution should decompose into a decoupled sum of rescaled solitons plus a regular term, up to an error of asymptotically vanishing energy, as one evolves towards its maximal time of existence.

To be more precise, we consider large energy wave maps as in the resolution of the threshold conjecture by Sterbenz and Tataru [29, 30], but more specifically into the unit Euclidean sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ with $n \geq 2$ (although parts of our argument work for general targets). We prove that, on a suitably chosen sequence of time slices approaching maximal existence, there is a decomposition of the map, up to an error with asymptotically vanishing energy, into a decoupled sum of rescaled solitons concentrating in the interior of the light cone and a term having asymptotically vanishing energy dispersion norm. For the latter, we further describe it as a linear gauge co-variant wave, concentrating on the null boundary and converging to a constant locally in the interior of the cone, in the energy space.

Similar and stronger results have been recently obtained in the equivariant setting by several authors [4, 5, 2, 3, 15], where better control on the dispersive term concentrating on the null boundary of the cone is provided and in some cases the asymptotic decomposition is shown to hold for all time. Here however, we do not impose any symmetry condition on the map itself and our strategy follows the one from bubbling analysis of harmonic maps into spheres in the supercritical regime due to Lin and Rivière [20, 21], which we make work here in the hyperbolic context of [30]. A large part of the work presented in this thesis has appeared in author's [12].

Contents

Chapter 1

Introduction: the Wave Maps Equation

1.1 Wave maps

We start by explaining some important facts about smooth wave maps into compact Riemannian manifolds. For a broad introduction to the subject, we shall refer the interested reader to the monograph of Shatah and Struwe [26]. Our exposition will however be self-contained.

1.1.1 The basic definition

In this thesis, we will work with closed Riemannian manifolds $\mathcal{N}$ equipped with an isometric embedding into some Euclidean space. If thinking about applications to mathematical General Relativity, it is important to consider the broader class of non-compact manifolds such as the hyperbolic space. However, as we will elaborate in the next section, for the particular case of the hyperbolic space the questions we are addressing here have been already completely resolved seven years ago. Also, it is interesting to point out that from the analytical perspective, having the manifold $\mathcal{N}$ embedded into $\mathbb{R}^n$ enables us to use techniques form harmonic and Fourier analysis with ease. Moreover, the boundedness of the map will be quite helpful for both the arguments upon which we will rely, and those which we will present.

Let us denote by $\mathcal{S} = (\mathcal{S}^i_{j\ell})$ the second fundamental form corresponding to the embedding $\mathcal{N} \hookrightarrow \mathbb{R}^n$. Wave maps are smooth maps $\phi : I \times \mathbb{R}^2 \to \mathbb{R}^n$, defined on some

time interval $I \subset \mathbb{R}$, with the evolution:

$$\phi[t_0] := (\phi(t_0), \partial_t \phi(t_0)), \quad t_0 \in I,$$

taking values in the tangent bundle $T\mathcal{N}$ of the manifold $\mathcal{N}$, so that

$$\mathcal{S}_{i\ell}^j(\phi) \nabla_{t,x} \phi^j = 0, \tag{1.1.1}$$

with $\phi[t]$ also required to belong to the Sobolev space $C_t^0(I\,;\dot{H}_x^1) \cap C_t^1(I\,;L_x^2)$ and governed by the following semi-linear wave equation:

$$\Box \phi^i = -\mathcal{S}_{j\ell}^i(\phi) \partial_\alpha \phi^\ell \partial^\alpha \phi^j. \tag{1.1.2}$$

Here, the D'Alembertian is given by $\Box := \partial_\alpha \partial^\alpha = -\partial_t^2 + \Delta_x$. Note our convention: we are summing over repeating indices, where α is running from 0 to 2, with $\partial_0 = \partial_t$ and $\partial^0 = -\partial_t$ as we will be always raising the indices with respect to the Minkowski metric $\mu = -dt^{\otimes 2} + dx_1^{\otimes 2} + dx_2^{\otimes 2}$ on $\mathbb{R}^{2+1}$ unless clearly stated otherwise.

Using the identities (1.1.1), we can rewrite the wave maps equation in a more geometric form:

$$\Box \phi + \Omega_\alpha \partial^\alpha \phi = 0, \tag{1.1.3}$$

where Ω_α is an $\mathfrak{so}(n)$ connection matrix defined by:

$$(\Omega_\alpha)_j^i := \left[\mathcal{S}_{j\ell}^i(\phi) - \mathcal{S}_{i\ell}^j(\phi) \right] \partial_\alpha \phi^\ell.$$

One of the main benefits of taking this point of view lies in the fact that given a change of gauge, i.e. change of moving frame on $\mathbb{R}^n$:

$$\nabla_{t,x} \phi \longmapsto U \nabla_{t,x} \phi,$$

for a smooth field U with values in $SO(n)$, the gauge, equation (1.1.3) transforms naturally after application of the latter:

$$\left(\partial_\alpha + \Omega_\alpha^U \right) U \partial^\alpha \phi = 0, \tag{1.1.4}$$

where we have the usual definition $\Omega_\alpha^U = \partial_\alpha U U^\dagger + U \Omega_\alpha U^\dagger$. This extra degree of freedom introduces the need of choosing the right gauge to do analysis in. As we will discuss later, several possibilities have been constructed by various authors. However, as far

as this thesis is concerned, the issue of choosing the right gauge will not be present for most of the arguments we advance.

1.1.2 Important examples

The most interesting example for us would be when the target manifold is chosen to be the Euclidean sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$. We should remark that all of the results we claim in this thesis do extend to the slightly more general case of homogeneous Riemannian manifolds, but for the sake of mathematical clarity we will focus on $\mathbb{S}^{n-1}$ – being the simplest yet fully representative situation. In this case, the evolution taking values in the tangent bundle $T\mathbb{S}^{n-1}$ leads to the following geometric identities:

$$\phi^{\dagger}\phi = 1, \quad \phi^{\dagger}\nabla_{t,x}\phi = 0. \tag{1.1.5}$$

The wave maps equation itself takes the following nice form:

$$\Box\phi = -\phi\partial_{\alpha}\phi^{\dagger}\partial^{\alpha}\phi. \tag{1.1.6}$$

A further special case of the above example, that we would like to single out in our presentation, is the one of equivariant wave maps. Those have the two-sphere as target (certainly more general surfaces of revolution can be considered here, but we leave it to the references in [26]) and take the following form:

$$\phi(t,r,\theta) = \begin{pmatrix} \sin\psi(t,r)\,\cos\ell\theta \\ \sin\psi(t,r)\,\sin\ell\theta \\ \cos\psi(t,r) \end{pmatrix}, \tag{1.1.7}$$

with (r,θ) standing for spatial polar coordinates. The integer $\ell = 1, 2, \ldots$ is the equivariance class of the map, and the function $\psi(t,r)$ solves the following semi-linear wave equation on $\mathbb{R} \times \mathbb{R}_+$:

$$-\partial_t^2\psi + \partial_r^2\psi + \frac{1}{r}\partial_r\psi - \frac{\ell^2\sin\psi\cos\psi}{r^2} = 0.$$

We will say a bit more about this set-up in the next section, as this is the one other authors have been considering.

1.1.3   Conservation laws

Let us now discuss some important properties that wave maps have. First of all, symmetries: we recall that equation (1.1.2) is invariant with respect to the scaling:

$$\phi(t, x) \longmapsto \phi(\lambda t, \lambda x),$$

for any $\lambda > 0$, and also any space-time translation. Moreover, a few important conservation laws are associated to the wave map evolution. Firstly, the energy of a wave map at time $t_0 \in I$, scale invariant in dimension 2+1, is given by:

$$\mathcal{E}[\phi](t_0) := \frac{1}{2} \int_{\mathbb{R}^2} |\partial_t \phi(t_0)|^2 + |\nabla_x \phi(t_0)|^2 \, dx = \frac{1}{2} \|\nabla_{t,x}\phi(t_0)\|^2_{L^2_x},$$

and a conservation of energy law holds:

$$\mathcal{E}[\phi](t_0) = \mathcal{E}[\phi](t_1), \tag{1.1.8}$$

for any $t_0, t_1 \in I$.

Secondly, restricting ourselves to the case $\mathcal{N} = \mathbb{S}^{n-1}$ and using the identities (1.1.5), it is simple to see that equation (1.1.6) is equivalent to the following conservation law:

$$\partial^\alpha(\phi \partial_\alpha \phi^\dagger - \partial_\alpha \phi \phi^\dagger) = 0. \tag{1.1.9}$$

The use of this identity means however, that some of our arguments do not directly generalize to the case when one has an arbitrary closed Riemannian manifold as a target.

A more geometric perspective on (1.1.9) would be to argue through Noether's theorem and the abundance of Killing vector fields on the Euclidean sphere (and similarly for other homogeneous Riemannian manifolds but, as mentioned before, we shall focus on the sphere in the thesis). For, one has to note that wave maps are formally critical points of the Lagrangian:

$$\mathcal{L}(\phi) := \int_{\mathbb{R}^{2+1}} \partial^\alpha \phi^\dagger \partial_\alpha \phi \, dt dx, \tag{1.1.10}$$

of which (1.1.2) is the Euler-Lagrange equation. The reader interested in this formalism is invited to have a look at the beautiful monograph of Hélein [13].

Going back to a general target $\mathcal{N}$, there is another consequence of the variational

point of view and Noether's theorem worth mentioning. In fact, smooth wave maps have the stress energy tensor:

$$T_{\alpha\beta}[\phi] := \partial_\alpha \phi^\dagger \partial_\beta \phi - \frac{1}{2}\mu_{\alpha\beta}\partial^\gamma \phi^\dagger \partial_\gamma \phi, \tag{1.1.11}$$

being divergence free:

$$\partial^\alpha T_{\alpha\beta}[\phi] = 0. \tag{1.1.12}$$

Note that this conservation law can be deduced directly from identities (1.1.1) and equation (1.1.2), without appealing to Noether's theorem.

The energy conservation law (1.1.8) is actually obtained by contracting $T[\phi]$ with ∂_t and using (1.1.12) with Stokes' theorem in $[t_0, t_1] \times \mathbb{R}^2$. As we shall see later, many other monotonicity and Morawetz type estimates, very important in the blow-up analysis of large energy wave maps, are obtained in this way.

Finally, closing our presentation of wave maps, we remark that the Lagrangian $\mathcal{L}$ is invariant under Lorentz transformations of $\mathbb{R}^{2+1}$, which implies that, after composition with Lorentz transformations, the map still solves equation (1.1.2) and in particular, if we target the Euclidean sphere, the conservation law (1.1.9) also stays true.

1.2 From the Threshold Theorem to the Soliton Resolution Conjecture

1.2.1 Some basic notation

Before presenting our main results, let us set up some notation. As usual, for two quantities A and B we will be writing $A \lesssim B$ if $A \leq C \cdot B$ for some implicit constant $C > 0$ whose dependence should be clarified when necessary. We also write $A \sim B$ whenever the additional estimate $B \lesssim A$ holds. Similarly, for the O-notation, we set $A = O(B)$ if $|A| \leq C \cdot B$.

Regarding the asymptotic notation, arising in various statements of the soliton decomposition below, we write $f_\nu = o_X(A)$, as $\nu \to +\infty$ in the background with X some Banach space (typically a Sobolev space), for a sequence of elements $f_\nu \in X$ with $\|f_\nu\|_X \leq c_\nu \cdot A$ where $c_\nu \downarrow 0$. In the same spirit, we will write $A_\nu \ll B_\nu$ whenever $A_\nu/B_\nu \to 0$ holds.

By $B_{r_0}(x_0) \subset \mathbb{R}^2$ we will be always referring to a spatial open ball of radius $r_0 > 0$ and center $x_0 \in \mathbb{R}^2$. In space-time our basic domains should be light cones. We denote

the forward light cone by:

$$C := \{(t,x) \,:\, 0 \le t, \, r \le t\}, \quad r := |x|,$$

and the restriction to some time interval I, as well as time sections, by:

$$C_I := C \cap (I \times \mathbb{R}^2), \quad S_{t_0} := C \cap (\{t_0\} \times \mathbb{R}^2),$$

respectively. We let $\partial C_I := \{(t,x) \,:\, t \in I, \, r = t\}$ standing for the *lateral* boundary, to which we usually refer as the *null boundary*. Given some $\delta > 0$, it will be convenient also to set $C^\delta := (\delta, 0) + C$, with the convention that C^0 stands for $\cup_{\delta>0} C^\delta$, the open interior of C. Accordingly, we have $C_I^\delta := C_I \cap C^\delta$, $S_{t_0}^\delta := S_{t_0} \cap C^\delta$ and if $\delta > 0$, ∂C_I^δ for the lateral boundary of C_I^δ.

1.2.2 Well-posedness versus blow-up

We recall now the set-up from [30], which holds for any closed Riemannian manifold $\mathcal{N}$ as target. By the finite speed of propagation, translation and scaling invariance properties, we shall restrict ourselves to the forward light cone C on which it is convenient to study at the same time both scenarios: the finite time blow-up at the tip of the cone, as well as the problem of scattering as $t \to +\infty$.

We pause here to note that, classically, one says that a wave map *scatters at infinity* if $\phi[t]$ is equal to a linear wave up to an error term which has asymptotically vanishing energy as $t \to +\infty$. In reality, because the wave maps equation transforms naturally under change of gauge, see (1.1.4), one can obtain such a behavior only by constructing a suitable gauge. Therefore, we should always state precisely what we mean by scattering in each concrete situation.

With this understood, we can assume that we are given a wave map ϕ on C, smooth up to but not necessarily including the origin $(0,0)$, and satisfying the energy bound:

$$\mathcal{E}_{S_{t_0}}[\phi] := \frac{1}{2} \|\nabla_{t,x}\phi\|^2_{L^2(S_{t_0})} \le \mathcal{E}, \quad \forall t_0 \in [0, \infty), \tag{1.2.1}$$

where $\mathcal{E}$ is an arbitrarily large but fixed for the rest of the thesis bound on which most of our constants will depend. Let us introduce here the notation for the energy of the

wave map ϕ over some domain $U \subset \mathbb{R}^{2+1}$ at the time slice $\{t = t_0\}$ setting:

$$\mathcal{E}_U[\phi](t_0) := \frac{1}{2} \int_{U \cap \{t=t_0\}} |\nabla_{t,x}\phi(t_0)|^2 \, dx = \frac{1}{2} \|\nabla_{t,x}\phi(t_0)\|^2_{L^2(U \cap \{t=t_0\})},$$

or simply $\mathcal{E}_U[\phi]$ when there is no ambiguity, as for example with $\mathcal{E}_{S_{t_0}}[\phi]$ above. For the latter quantity, we recall the important monotonicity property:

$$\mathcal{E}_{S_{t_0}}[\phi] \leq \mathcal{E}_{S_{t_1}}[\phi] \ \text{ for } \ t_0 \leq t_1.$$

It is obtained, as the conservation of energy law (1.1.8), contracting the stress energy tensor $T[\phi]$ with ∂_t and using (1.1.12) with Stokes' theorem, this time however applied in $C_{[t_0,t_1]}$, giving:

$$\mathcal{E}_{S_{t_1}}[\phi] = \mathcal{F}_{[t_0,t_1]}[\phi] + \mathcal{E}_{S_{t_0}}[\phi], \ \ \mathcal{F}_{[t_0,t_1]}[\phi] := \int_{\partial C_{[t_0,t_1]}} \left(\frac{1}{4} |L\phi|^2 + \frac{1}{2} |r^{-1}\partial_\theta\phi|^2 \right) dA. \ \ (1.2.2)$$

The quantity $\mathcal{F}_{[t_0,t_1]}[\phi]$ is called the *flux* of the wave map from t_1 to t_0, and L is part of the null frame:

$$L := \partial_t + \partial_r, \ \ \underline{L} := \partial_t - \partial_r.$$

The monotonicity property and the global bound (1.2.1) enable us to define the limits:

$$\mathcal{E}_0 := \lim_{t\downarrow 0} \mathcal{E}_{S_t}[\phi], \ \ \mathcal{E}_\infty := \lim_{t\uparrow\infty} \mathcal{E}_{S_t}[\phi],$$

and imply that $\mathcal{F}_{[t_0,t_1]}[\phi] \downarrow 0$ as t_0, t_1 both tend to zero or infinity. The latter can be used, together with the angular part of $\mathcal{F}_{[t_0,t_1]}[\phi]$ from (1.2.2), to construct, given any $\varepsilon > 0$, an extension of ϕ outside the cone C on $(0, t_0]$ for $t_0 = t_0(\varepsilon)$ small enough, and on $[t_\infty, \infty)$ for $t_\infty = t_\infty(\varepsilon)$ large enough, solving the wave maps equation (which is possible by finite speed of propagation, hence we shall slightly abuse notation denoting those extensions by ϕ) such that:

$$\mathcal{E}[\phi](t) - \mathcal{E}_{S_t}[\phi] \leq \varepsilon\mathcal{E}, \ \ \forall t \in (0, t_0] \cup [t_\infty, \infty).$$

The details of this construction can be found in Sections 6.1 and 6.2 of [30], and a similar argument is presented here in Section 3.2.

By the small energy theorem of Tao [32], that we explain carefully in Section 2.2, if $\mathcal{E}[\phi](t_0)$ can be chosen small enough, then $\mathcal{E}_0 = 0$ and ϕ can be extended to a smooth wave map for all time. This guarantees also that the above extensions are smooth

everywhere except possibly $(0,0)$, even if $\mathcal{E}_{S_t}[\phi]$ is large, provided $\varepsilon > 0$ was chosen small enough initially.

Moreover, via a continuity-iteration-renormalization argument, ϕ is proved in [32] to belong to a space $S \subset C_t^0(I\,;\dot{H}_x^1) \cap C_t^1(I\,;L_x^2)$, implying control in all the Strichartz spaces amongst others, in which well-posedness for the Cauchy problem (1.1.6) can be established. We discuss this more precisely with further references later in Section 2.2.

Following the terminology of Sterbenz and Tataru [30], we will say that *scattering* holds if:

$$\phi \in S,$$

noting that, strictly speaking, this means that ϕ behaves like a linear wave as $t \to \pm\infty$ after passing to frequency localization and applying the microlocal gauge (if small energy, see [32]) or the diffusion gauge (necessary if large energy, see [29]). We refer the reader to the structure theorem of Sterbenz and Tataru in [29], and Section 3.2 in this thesis for further information. Therefore, if $\mathcal{E}[\phi](t_\infty)$ could be chosen small enough for some extension we consider the scattering problem for ϕ as $t \to +\infty$ resolved.

Let us take the opportunity here to remark that, if the target manifold is a hyperbolic Riemann surface, then scattering in the classical sense was established by Krieger and Schlag [17] for wave maps in the Coulomb gauge. For the hyperbolic spaces, this was achieved by Tao [33] using the caloric gauge. Those works were considering wave maps with initial data of arbitrarily large energy, without any further assumptions. Therefore, as we explain below, they obtain the proof of the full *Soliton Resolution Conjecture* for those targets, see Section 1.2.4.

Once energy gets large, blow-up can occur and the first examples of finite time singularity for equivariant wave maps into $\mathbb{S}^2$ were constructed by Krieger, Schlag and Tataru [18], as well as Rodnianski and Sterbenz [24] and also Raphaël and Rodnianski [22], where, as for the harmonic map heat flow, the mechanism behind the singular behavior was concentration of a non-trivial harmonic map.

More generally, the wave map ϕ could have concentrated at the origin at least one *soliton*: these are defined to be finite energy smooth maps $\omega : \mathbb{R}^{2+1} \to \mathcal{N}$ solving the wave maps equation (1.1.2) and satisfying:

$$X\omega = 0,$$

for some constant time-like vector field X on $\mathbb{R}^{2+1}$. In particular, precomposing ω with a Lorentz transformation Ψ that takes ∂_t to X, we obtain a finite energy harmonic

map from $\mathbb{R}^2$ steady in the time direction which, upon extending over spatial infinity using the removable singularity theorem of Sacks and Uhlenbeck [25], gives a harmonic two-sphere $\omega \circ \Psi : \mathbb{R} \times \mathbb{S}^2 \to \mathcal{N}$ familiar from the bubbling analysis of harmonic maps and heat flows. Let us note here that this last point of view enables us to set $\omega(\infty) := \lim_{|x| \to \infty} \omega(t, x)$, which is well-defined and independent of time t chosen.

The *Threshold Conjecture*, resolved by Sterbenz and Tataru [29, 30] (for closed Riemannian manifolds), Krieger and Schlag [17] (for hyperbolic surfaces) and Tao [33] (for hyperbolic spaces of any dimension), predicts that concentration of solitons is the essential mechanism behind blow-up. That is if $\mathcal{E}_0, \mathcal{E}_\infty$ are less than the energy threshold below which every harmonic two-sphere is constant, then one has regularity at $t = 0$ and scattering as $t \to +\infty$.

One of the central difficulties in establishing this conjecture, in the general non-symmetric situation, was that relying only on standard Morawetz type estimates obtained from the stress energy tensor, it was not possible to get a non-trivial amount of energy concentrating within the light cone required to produce a non-constant soliton. As far as the program of Sterbenz and Tataru is concerned, the breakthrough was made in [29], where they obtain that, on top of concentrating energy, the map must concentrate a non-trivial amount $\epsilon(\mathcal{E}, \mathcal{N}) > 0$ of the BMO type *energy dispersion norm*. That is if:

$$\sup_k \|P_k \phi\|_{L^\infty_{t,x}((0,t_0] \cup [t_\infty, \infty))} < \epsilon(\mathcal{E}, \mathcal{N}), \tag{1.2.3}$$

where P_k stands for the Littlewood-Paley projection (see Section 2.1), then:

$$\phi \in S((0, t_0] \cup [t_\infty, \infty)),$$

and the map extends smoothly to a neighborhood of $t = 0$ (we shall state a slightly more precise version of this theorem in Section 2.2). This is a large data result and is proved in [29] via an induction on energy argument.

Let us note here, as an aside, that the program of Krieger and Schlag [17], as well as the one of Tao [33], proceeded via a different induction on energy argument and without any smallness assumption as (1.2.3). As there are no non-constant solitons for the targets considered there, one obtains global regularity and scattering for arbitrarily large data in those cases. We point out on the other hand, that the concentration-compactness techniques used in [17] can also lead to a fruitful study of the formation of solitons, as was demonstrated so far for equivariant wave maps in [4, 5, 2, 3, 15]. In the present work however, we shall adopt a more direct approach staying closer to [29, 30],

see Section 1.3 for a detailed summary of our strategy.

In Section 3.2, we will briefly discuss results from [30] that convert concentration of energy dispersion into concentration of a non-trivial amount of time-like energy, as this is how, arguing by contradiction, we get the energy dispersion norm of the term concentrating on the null boundary asymptotically vanishing. On the other hand, the fact that arguments in [30] give that only *some* energy is prevented from escaping into the null boundary at a finite time singularity, is a serious obstacle to controlling null concentration further. In fact, techniques dealing with this phenomenon would have to strengthen [29, 30] considerably in this situation, if not giving a wholly alternative proof to the threshold conjecture (which we shall not attempt in this thesis).

Theorem 1.2.1. (Threshold Theorem of Sterbenz and Tataru [29, 30]). *Suppose that the wave map ϕ is singular at $(0, 0)$, respectively $\phi \notin S[t_\infty, \infty)$ for any extension as discussed above, then there exists a sequence $\lambda_\nu^0 \downarrow 0$, respectively $\lambda_\nu^\infty \uparrow \infty$, the so-called final rescaling, such that setting:*

$$\phi_\nu(\cdot) := \phi(\lambda_\nu^0 \cdot), \quad \text{respectively} \quad \phi(\lambda_\nu^\infty \cdot),$$

we can find a sequence of concentration points $(t_\nu, x_\nu) \in C_{[1, O(1)]}^{\frac{1}{2}}$ and scales $r_\nu \downarrow 0$, for which:

$$\phi_\nu(t_\nu + r_\nu t, x_\nu + r_\nu x) \longrightarrow \omega(t, x) \quad \text{in} \quad (H_{t,x}^1)_{loc}([-\frac{1}{2}, \frac{1}{2}] \times \mathbb{R}^2),$$

for some non-constant soliton ω.

We shall describe in detail the final rescaling ϕ_ν at the beginning of Chapter 3, see Lemma 3.0.1.

1.2.3 Statement of the main results

In our main theorems, we study the above sequence ϕ_ν of Theorem 1.2.1 further, carrying out a blow-up analysis for it and establishing an analogue of the energy identity from the bubbling analysis of harmonic maps and heat flows (and many other geometric variational problems), see for example the works [6, 36, 19] and the references therein for the critical regime, and for a supercritical situation the papers of Lin and Rivière [20, 21], which are of closer flavor to the arguments presented in this thesis.

Below, Theorems 1.2.2 and 1.2.3 hold for a general target $\mathcal{N}$ as considered in Theorem 1.2.1.

Theorem 1.2.2. (Blow-up analysis for asymptotically self-similar sequences of wave maps). *Upon passing to a subsequence for the wave maps $\{\phi_\nu\}_{\nu\in\mathbb{N}}$ obtained in Theorem 1.2.1, or abstractly those satisfying the conclusions of Lemma 3.0.1, the following holds.*

There exists a non-trivial finite collection of time-like geodesics $\varrho_1,\ldots,\varrho_I$, emanating from the origin in Minkowski space $\mathbb{R}^{2+1}$, along which the maps concentrate some threshold $\epsilon_s > 0$ of energy:

$$\liminf_{\nu\to\infty}\mathcal{E}_{B_r(\varrho_i(t))}[\phi_\nu] > \epsilon_s \ \ \forall t\in[1,2], \ \ \forall r>0, \ \ i=1,\ldots,I,$$

where we are writing $\varrho_i(t) := \varrho_i \cap S_t$, and the maps converge locally to a constant away from ϱ_i in the interior of the light cone:

$$\phi_\nu \longrightarrow \text{const.} \ \ on \ \ C^0_{[1,2]}\setminus\cup_i\varrho_i,$$

locally in $C^0_t(H^1_x)\cap C^1_t(L^2_x)$.

In our next theorem, we consider the parts of the maps ϕ_ν that get concentrated on the null boundary ∂C.

Theorem 1.2.3. (Dispersive property for null-concentration). *Given the set up from Theorem 1.2.2, fix a small distance to the null boundary:*

$$\delta_0 := \frac{1}{10}\text{dist}(\cup_i\varrho_i,\partial C_{[1,2]}).$$

The maps ϕ_ν on $C_{[t_0-\delta_0,t_0+\delta_0]}\setminus C^{2\delta_0}_{[t_0-\delta_0,t_0+\delta_0]}$ admit extensions $\varpi_{t_0,\nu}$ to $[t_0-\delta_0,t_0+\delta_0]\times\mathbb{R}^2$, for each $t_0\in[1+\delta_0,2-\delta_0]$, solving the wave maps equation on this short, but independent of ν, time interval and converging to a constant away from ∂C:

$$\nabla_{t,x}\varpi_{t_0,\nu}\longrightarrow 0 \ \ in \ \ C^0_t(L^2_x)_{loc}\left(([t_0-\delta_0,t_0+\delta_0]\times\mathbb{R}^2)\setminus\partial C_{[t_0-\delta_0,t_0+\delta_0]}\right);$$

with a decay for the angular and the null $L = \partial_t + \partial_r$ energy:

$$\sup_{t\in[t_0-\tau,t_0+\tau]}(\left\|L\varpi_{t_0,\nu}(t)\right\|_{L^2_x} + \left\|r^{-1}\partial_\theta\varpi_{t_0,\nu}(t)\right\|_{L^2_x}) \longrightarrow 0.$$

Moreover, the maps $\varpi_{t_0,\nu}$ admit a uniform dispersive bound:

$$\left\|\varpi_{t_0,\nu}\right\|_{S[t_0-\delta_0,t_0+\delta_0]} \lesssim 1,$$

for any $t_0\in[1+\delta_0,2-\delta_0]$, and we can extract the following asymptotic structure for

them:

$$\varpi_{t_0,\nu} = \sum_{k\in\mathbb{Z}} \psi_{\nu,k} + o_{L_t^\infty(\dot{H}_x^1\times L_x^2)[t_0-\delta_0,t_0+\delta_0]}(1), \tag{1.2.4}$$

where each $\psi_{\nu,k}$ solves a gauge co-variant linear wave equation:

$$\begin{cases} \left(\Box + 2A_\alpha^{(k,\nu)}\partial^\alpha\right)\psi_{\nu,k} = 0 \\ \psi_{\nu,k}[t_0] = P_k\varpi_{t_0,\nu}[t_0] \end{cases},$$

with $A_\alpha^{(k,\nu)} \in \mathfrak{so}(n)$ given by:

$$(A_\alpha^{(k,\nu)})_j^i = P_{\leq k-m_\nu}\left[\mathcal{S}_{j\ell}^i(\varpi_{t_0,\nu}) - \mathcal{S}_{i\ell}^j(\varpi_{t_0,\nu})\right]P_{\leq k-m_\nu}\left[\partial_\alpha\varpi_{t_0,\nu}^\ell\right],$$

for a sequence of integers $m_\nu \longrightarrow +\infty$.

Our final theorem is the heart of this thesis and requires the full machinery of compensation theory that we develop in Section 2.3 which relies on the crucial conservation law (1.1.9) not available for general targets.

Theorem 1.2.4. (Quantization of time-like energy for wave maps into spheres). *Consider the set up from Theorems 1.2.2 and 1.2.3 with $\mathcal{N} = \mathbb{S}^{n-1}$. We can find a sequence of time slices:*

$$\{t_\nu\}_{\nu\in\mathbb{N}} \subset [1+\delta_0, 2-\delta_0],$$

on which there exists a non-trivial collection of $J = J(\{t_\nu\}_{\nu\in\mathbb{N}}) \lesssim_\mathcal{E} 1$ sequences of points $a_\nu^j \in \mathbb{R}^2$, $|a_\nu^j| < t_\nu - 5\delta_0$, with associated scales $\lambda_\nu^j \downarrow 0$ for $j = 1, \ldots, J$, satisfying the orthogonality conditions:

$$\frac{\lambda_\nu^i}{\lambda_\nu^j} + \frac{\lambda_\nu^j}{\lambda_\nu^i} + \frac{|a_\nu^i - a_\nu^j|^2}{\lambda_\nu^i\lambda_\nu^j} \longrightarrow \infty$$

as $\nu \to +\infty$ for distinct $i \neq j$, such that:

$$\phi_\nu(t,x) = \sum_{j=1}^J \left(\omega_j\left(\frac{t-t_\nu}{\lambda_\nu^j}, \frac{x-a_\nu^j}{\lambda_\nu^j}\right) - \omega_j(\infty)\right) + \varpi_{t_\nu,\nu}(t,x) + o_{\dot{H}_x^1\times L_x^2}(1) \text{ on } S_{t_\nu},$$

where $\omega_j : \mathbb{R}^{2+1} \to \mathbb{S}^{n-1}$ are solitons for which:

$$\phi_\nu(t_\nu + \lambda_\nu^j t, a_\nu^j + \lambda_\nu^j x) \longrightarrow \omega_j(t,x) \text{ on } \mathbb{R}^{2+1} \setminus \cup_q \varrho_q^j, \tag{1.2.5}$$

locally in $C_t^0(H_x^1) \cap C_t^1(L_x^2)$, for a finite collection, $q = 1, \ldots, q(\omega_j, \mathcal{E})$, of parallel time-

like geodesics ϱ_q^j.

In other words, we have *Energy Quantization* in the interior of the light cone for wave maps into spheres.

1.2.4 Towards the Soliton Resolution Conjecture

Our Theorems 1.2.2, 1.2.3 and 1.2.4 represent a little first step towards understanding the *Soliton Resolution Conjecture* for the (2+1)-dimensional wave maps equation with target $\mathbb{S}^{n-1}$. It states that in addition, such a decomposition should be unique holding for all time and that $\varpi_{t_0,\nu}$ should have asymptotically vanishing energy in the case of finite time blow-up, or correspond to the scattering part of the wave map in the case of global existence.

We mention here that the soliton resolution conjecture has recently been shown to hold for the 1-equivariant wave maps into $\mathbb{S}^2 \subset \mathbb{R}^3$ (see equation (1.1.7) from Section 1.1.2) with initial data having topological degree one and energy strictly less than 3 times 4π (note that 4π is the energy threshold) by Côte, Kenig, Lawrie and Schlag at finite time singularity in [4], and in [5] for the case of global existence (more general surfaces of revolution are also considered). Note that in this situation, one knows a priori the uniqueness of the possible configurations of solitons that can be concentrated (in fact there is only one of them and it is the unique equivariant degree one harmonic map). The conjecture is also established for the examples constructed by Krieger, Schlag and Tataru [18], as well as Raphaël and Rodnianski [22].

Without this restriction on the initial data, the soliton resolution along a sequence of times was obtained in the 1-equivariant setting by Côte [2, 3] building upon [4, 5], and more generally for the ℓ-equivariant case for any integer $\ell \geq 1$ by Jia and Kenig [15] relying on a method different from [4, 5, 2] (in both works, the finite time singularity and the global existence case have been considered). We refer the reader to [15] for more references and an overview with some history of the various beautiful techniques used to tackle the soliton resolution conjecture in the radial/equivariant cases for a variety of non-linear wave equations initiated by Duyckaerts, Kenig and Merle, see for example [8]. We also note that those techniques have been very recently applied to prove the sequential soliton resolution conjecture without any symmetry assumptions for some focusing semi-linear wave equations by Duyckaerts, Jia, Kenig and Merle [14, 7, 9]. The strategy of the present thesis will have a very different flavor though. An outline can be found in Section 1.3.

It is worth mentioning here that our theorems give an alternative proof for some of the results obtained in [4, 2, 15]. For, we have to recall the following self-similar energy decay estimate for equivariant wave maps due to Christodoulou and Tahvildar-Zadeh [1], as well as Shatah and Tahvildar-Zadeh [27] from the early 90's.

Theorem 1.2.5. (Self-similar energy decay [1, 27]). *For an equivariant wave map* $\phi : C \to \mathbb{S}^2$, *given any constant* $0 < \lambda < 1$, *we have:*

$$\mathcal{E}_C[\phi](t) - \mathcal{E}_{\lambda C}[\phi](t) \longrightarrow 0 \ \ as \ \ t \downarrow 0.$$

Combined with the above result, Theorems 1.2.2, 1.2.3 and 1.2.4 immediately imply the sequential version of SRC in the case of finite time blow-up for equivariant wave maps. For the case of global existence, one has to rely on [5] to extract the scattering profile of the map before applying our quantization result. We should, however, abstain from reporting on this further here.

To close this section, let us say that the techniques we use to establish the above theorems leave completely open the question of uniqueness of the set of solitons. In fact, as suggested by an example of Topping [36] for the harmonic map heat flow, this, and therefore the soliton resolution conjecture, could fail for certain targets (in view of the work of Simon [28] however, such pathologies are believed to be excluded when working with real analytic targets like $\mathbb{S}^{n-1}$). Therefore, there is a notoriously difficult and long way from Theorem 1.2.4 to the full soliton resolution conjecture as one should expect the quantization result to hold for any closed Riemannian manifold as a target and the only place where we use the fact that our target is a sphere is when relying on the conservation law (1.1.9) in the proof of the compensation estimates in Section 2.3. Establishing the analogue of those estimates for general targets is an important open question even in the elliptic theory, see the work of Rivière [23] for a further discussion.

1.3 Discussion of the strategy and overview of the argument

We should close the introduction by outlining the proofs of Theorems 1.2.2, 1.2.3 and 1.2.4, which are contained in Section 3.

Firstly, Theorem 1.2.2 is obtained in Section 3.1. For the sequence of wave maps $\{\phi_\nu\}_{\nu \in \mathbb{N}}$ at the final rescaling, Sterbenz and Tataru [30] obtain a decay estimate along

the scaling vector field $\partial_\rho = \frac{1}{(t^2-r^2)^{1/2}}(t\partial_t + r\partial_r)$:

$$\int \int_{C^{\epsilon_\nu^{\frac{1}{2}}}_{[\varsigma_\nu,\varsigma_\nu^{-1}]}} \frac{1}{(t^2-r^2)^{\frac{1}{2}}} \, |\partial_\rho\phi_\nu|^2 \, dxdt \longrightarrow 0,$$

for some sequences $\varsigma_\nu \downarrow 0$, $\epsilon_\nu^{\frac{1}{2}} \ll \varsigma_\nu$, see Lemma 3.0.1. If one uses a local version of the latter, by contracting the stress energy tensor (1.1.11) with $\varphi\partial_\rho$, for some compactly supported cut-off φ on the unit hyperbolic plane $\mathbb{H}^2$, it is possible to spread a given energy control on some ball $B_{r_0}(x_0) \Subset S_1^0$, at the time slice $t = 1$ say, along the flow of the vector field ∂_ρ for any finite amount of time; in other words the wave maps ϕ_ν would have small energy, uniformly in ν, on the whole of:

$$\{\lambda z \; : \; \lambda \in [1,2], \, z \in \{t = 1\} \times B_{r_0}(x_0)\}\,,$$

provided they did so initially at $t = 1$. This is a simple analogue of the fact, from the blow-up analysis of supercritical harmonic maps, that one must have the tangent Radon measures monotone under scaling (see the work of Lin [20], and Lemma 3.1.1 here).

This way, relying as well on concentration-compactness at $t = 1$ and the small energy compactness result under control of a time-like direction due to Sterbenz and Tataru [30], see Lemma 2.2.2 here, we are able to obtain a subsequence for $\{\phi_\nu\}_{\nu\in\mathbb{N}}$ which converges on $C^0_{[1,2]}$, away from a finite set of time-like rays passing through the origin, to a regular self-similar wave map ϕ. By homogeneity and the singularity removable theorem of Sacks and Uhlenbeck [25], the map ϕ extends to a smooth wave map on the whole of the open forward light cone C^0 (the details of this argument are contained in Lemma 3.1.2). We note that similar arguments give also the convergence to solitons statement (1.2.5) claimed in Theorem 1.2.2 (see Lemma 3.3.1 for this point). We recall, however, that self-similar wave maps of finite energy must be constant. This is a well-known result, the proof of which can be found in [30] (see also Proposition 3.1.3 here for a precise statement).

On the other hand, another crucial property of the wave maps at the final scaling of Sterbenz and Tataru [30], is that a non-trivial amount of energy is uniformly held at a fixed distance away from the null boundary. Hence, our configuration of time-like rays, along which the wave maps concentrate, must be non-trivial. At this stage, this yields the proof of Theorem 1.2.2.

Because only some time-like energy is obtained in [30] (as some of the total energy

can be consumed by the linear part of the map if one considers the non-scattering problem for example), our second result, Theorem 1.2.3 treated in Section 3.2, tries to address the issue of null concentration. By cutting the parts of the map concentrating at the time-like geodesics, we are able to solve the wave maps equation for a uniform amount of time, even though the energy of the initial data is a priori large (thanks to the finite speed of propagation property and the fact the configuration of time-like rays was fixed initially). Running the arguments of Sterbenz and Tataru [30] backwards, yields then the claimed control for the energy dispersion norm (see Lemma 3.2.1), as well as decay for the angular and L null energy. Finally, the decomposition (1.2.4) claimed in Theorem 1.2.3 is a consequence of the theory developed in [32, 29].

Up until now, all of our arguments were valid for any closed Riemannian manifold $\mathcal{N}$ as target. For our last result, we have to restrict ourselves to the case when $\mathcal{N} = \mathbb{S}^{n-1}$, the Euclidean sphere.

The construction of the asymptotic decomposition and the proof of the energy quantization, Theorem 1.2.4, is contained in Section 3.3. Upon choosing a suitable sequence of time slices $\{t_\nu^{(1)}\}_{\nu \in \mathbb{N}} \subset (1,2)$ and scales $\delta_\nu \downarrow 0$, we study the wave maps:

$$\phi_{i,\nu}(\cdot) := \phi_\nu(t_\nu^{(1)} + \delta_\nu \cdot, \varrho_i(t_\nu^{(1)}) + \delta_\nu \cdot) \ \text{on} \ [-1,1] \times B_1,$$

for each geodesic ϱ_i, from Theorem 1.2.2. The maps $\phi_{i,\nu}$ converge to the constant c_ϕ corresponding to the self-similar wave map ϕ mentioned previously, locally in $L_t^\infty(H_x^1 \times L_x^2)$ away from ϱ_i, and in fact strongly in $L_t^\infty(L_x^2)$. The time slices $\{t_\nu^{(1)}\}_{\nu \in \mathbb{N}}$ have been chosen such that:

$$X_i \phi_{i,\nu} \longrightarrow 0 \ \text{in} \ L_{t,x}^2,$$

for the constant time-like vector field X_i pointing in the direction of the ray ϱ_i. The concentration scales $\{\delta_\nu\}_{\nu \in \mathbb{N}}$ have been chosen decaying slowly enough, to avoid losing energy in the process:

$$\lim_{\nu \to \infty} \sup_{t \in [1,2]} \mathcal{E}_{S_t^{\delta_\nu} \setminus \cup_i B_{\delta_\nu}(\varrho_i(t))}[\phi_\nu] = 0.$$

From there, we appeal to the compensation type estimates from Section 2.3 (the only place where we use the fact that our target is the sphere $\mathbb{S}^{n-1}$), decomposing the gradient as:

$$\nabla_{t,x} \phi_{i,\nu} = \Theta_{i,\nu} + \Xi_{i,\nu},$$

$$\text{with} \ \Theta_{i,\nu} \longrightarrow 0 \ \text{in} \ L_{t,x}^2 \ \text{and} \ \sum_{k \in \mathbb{Z}} \|P_k \Xi_{i,\nu}\|_{L_t^1(L_x^2)} \lesssim 1,$$

which is obtained in Proposition 2.3.1. To construct $\Theta_{i,\nu}$, we rely essentially on the time-like decay above, and for $\Xi_{i,\nu}$ the div-curl type structure of the non-linearity:

$$\Omega_\alpha^{i,\nu}\partial^\alpha\phi_{i,\nu}, \text{ where } \Omega_\alpha^{i,\nu} := \phi_{i,\nu}\partial_\alpha\phi_{i,\nu}^\dagger - \partial_\alpha\phi_{i,\nu}\phi_{i,\nu}^\dagger,$$

coming from the conservation law (1.1.9). Furthermore, we obtain a decomposition for the higher order time-like derivatives of $\phi_{i,\nu}$:

$$X_i^2\phi_{i,\nu} = \Gamma_{i,\nu} + \Pi_{i,\nu},$$

where the first term is a linear combination of:

$$\sum_{k\in\mathbb{Z}} P_k\nabla_x[\Omega_x^{i,\nu}(P_{>k+10}\phi_{i,\nu})], \quad \sum_{k\in\mathbb{Z}} P_k[\Omega_x^{i,\nu}(P_{\leq k+10}\nabla_x\phi_{i,\nu})], \text{ and } \Omega_{t,x}^{i,\nu}\nabla_{t,x}\phi_{i,\nu}, \quad (1.3.1)$$

that we note being local in time and quadratic in the gradient, and the second one satisfies a favorable decay estimate:

$$\sum_{k\in\mathbb{Z}} 2^{-2k}\|P_k\Pi_{i,\nu}\|_{L_{t,x}^2[-1,1]}^2 \longrightarrow 0.$$

This is obtained in Lemma 2.3.2 of Section 2.3, relying crucially on the conservation law (1.1.9) again, and plays an important role in the proof of the Besov decay estimate for wave maps on neck domains of Lemma 3.3.3 in Section 3.3, to which we come in few moments here.

We proceed then by constructing the soliton decomposition for the wave maps $\phi_{i,\nu}$, up to terms called *necks* in the literature on harmonic maps, which are given by $\phi_{i,\nu}$ restricted to a finite collection of conformally degenerating annuli:

$$[-\frac{r_{i,\nu}^k}{2}, \frac{r_{i,\nu}^k}{2}] \times \left(B_{R_{i,\nu}^k}(x_{i,\nu}^k)\setminus B_{r_{i,\nu}^k}(x_{i,\nu}^k)\right) \subset [-1,1]\times B_1 \text{ with } r_{i,\nu}^k \ll R_{i,\nu}^k$$

and $k = 1,\ldots, K_i(\mathcal{E})$, satisfying the local energy decay estimate:

$$\sup_{2^{-\ell}r_{i,\nu}^k\leq r\leq 2^\ell R_{i,\nu}^k} \sup_{t\in[-\frac{r}{2},\frac{r}{2}]} \mathcal{E}_{B_{2r}(x_{i,\nu}^k)\setminus B_r(x_{i,\nu}^k)}[\phi_{i,\nu}](t) \longrightarrow 0, \quad (1.3.2)$$

for any positive integer $\ell \in \mathbb{N}$. This is the content of Lemma 3.3.1, and represents essentially a standard argument of concentration-compactness. The whole of Theorem 1.2.4 is then reduced to showing that those necks have asymptotically vanishing energy.

In doing so, upon picking up suitable time slices $\{t_\nu^{(2)}\}_{\nu\in\mathbb{N}} \subset (-\frac{1}{2}, \frac{1}{2})$ before applying Lemma 3.3.1, and taking the fastest concentrating scale $\lambda_{\min,\nu} := \min_i\{\lambda_\nu^i\}$, we consider the maps:

$$\phi_{\nu,x_{i,\nu}^k}(t,x) := \phi_{i,\nu}(t_\nu^{(2)} + \lambda_{\min,\nu}t, x_{i,\nu}^k + \lambda_{\min,\nu}x) \text{ on } [-1,1] \times \mathbb{R}^2,$$

together with:

$$\Theta_{\nu,x_{i,\nu}^k}(t,x) := \lambda_{\min,\nu}\Theta_{i,\nu}(t_\nu^{(2)} + \lambda_{\min,\nu}t, x_{i,\nu}^k + \lambda_{\min,\nu}x),$$
$$\Xi_{\nu,x_{i,\nu}^k}(t,x) := \lambda_{\min,\nu}\Xi_{i,\nu}(t_\nu^{(2)} + \lambda_{\min,\nu}t, x_{i,\nu}^k + \lambda_{\min,\nu}x),$$
$$\Pi_{\nu,x_{i,\nu}^k}(t,x) := \lambda_{\min,\nu}^2\Pi_{i,\nu}(t_\nu^{(2)} + \lambda_{\min,\nu}t, x_{i,\nu}^k + \lambda_{\min,\nu}x),$$

and $\{t_\nu^{(2)}\}_{\nu\in\mathbb{N}}$ was chosen in such a way that:

$$\left\|\Theta_{\nu,x_{i,\nu}^k}(0)\right\|_{L_x^2} + \left\|X_i\phi_{\nu,x_{i,\nu}^k}(0)\right\|_{L_x^2} + \sum_{k\in\mathbb{Z}} 2^{-2k}\left\|P_k\Pi_{\nu,x_{i,\nu}^k}(0)\right\|_{L_x^2}^2 \longrightarrow 0.$$

We use then the second and third items of the decay statement above, to write for the gradient of $\phi_{\nu,x_{i,\nu}^k}$ on the neck domain:

$$\nabla_{t,x}\phi_{\nu,x_{i,\nu}^k} = \Upsilon_{\nu,x_{i,\nu}^k} \text{ on } [-1,1] \times (B_{\lambda_{\min,\nu}^{-1}R_{i,\nu}^k} \setminus B_{\lambda_{\min,\nu}^{-1}r_{i,\nu}^k}),$$

with the RHS supported on $[-1,1] \times (B_{2\lambda_{\min,\nu}^{-1}R_{i,\nu}^k} \setminus B_{2^{-1}\lambda_{\min,\nu}^{-1}r_{i,\nu}^k})$ and satisfying:

$$\left\|\Upsilon_{\nu,x_{i,\nu}^k}\right\|_{L_t^\infty(L_x^2)[-1,1]} \lesssim 1, \quad \sup_{k\in\mathbb{Z}}\left\|P_k\Upsilon_{\nu,x_{i,\nu}^k}(0)\right\|_{L_x^2} \longrightarrow 0.$$

This is proved in Lemma 3.3.3 using the decay for $X_i\phi_{\nu,x_{i,\nu}^k}$, localizing to the neck region the already obtained favorable estimate for $\Pi_{\nu,x_{i,\nu}^k}$, and relying on the local energy control (1.3.2) to get a weak $\dot{B}_\infty^{-1,2}$ decay estimate for the non-linear terms at high frequency, which are quadratic in the gradient of the map $\phi_{\nu,x_{i,\nu}^k}$ such as (1.3.1) left over from Lemma 2.3.2.

Finally, we are brought to the following control for the energy of $\phi_{\nu,x_{i,\nu}^k}$ on the neck

domain at time $t = 0$:

$$\left\| \nabla_{t,x} \phi_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2(B_{\lambda_{\min,\nu}^{-1} R_{i,\nu}^k} \setminus B_{\lambda_{\min,\nu}^{-1} r_{i,\nu}^k})}^2$$

$$\lesssim (\sup_{k \in \mathbb{Z}} \left\| P_k \Upsilon_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2}) \sum_{k \in \mathbb{Z}} \left\| P_k \Xi_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2}$$

$$+ \left\| \Upsilon_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2} \left\| \Theta_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2} + o(1),$$

and this gives the desired energy collapsing result.

Chapter 2

Technical results and Compensation Theory

In this section we gather some of the technical results, mainly restricted to the regularity theory of wave maps, that we will be using in Section 3 to establish our Main Theorems. The crucial compensation estimate is proved in Section 2.3.

2.1 Some harmonic analysis

2.1.1 Fourier Analysis and Littlewood–Paley theory

We will be mainly relying on the spatial Fourier transform. For $\phi(t,x) \in \mathcal{S}(\mathbb{R}^2)$, a Schwartz function on $\mathbb{R}^2$ at some fixed time t, we define:

$$\hat{\phi}(t,\xi) := \int_{\mathbb{R}^2} e^{-2\pi i x \cdot \xi} \phi(t,x) dx,$$

together with the inverse transform given by:

$$\check{\varphi}(t,x) = \int_{\mathbb{R}^2} e^{2\pi i x \cdot \xi} \varphi(t,\xi) d\xi,$$

for a Schwartz function $\varphi(t,\xi)$ on the frequency space. The space-time Fourier transform:

$$\mathcal{F}\psi(\tau,\xi) = \int_{\mathbb{R}^2} \int_{\mathbb{R}} e^{-2\pi i(t\tau + x \cdot \xi)} \psi(t,x) dt dx, \quad \psi \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^2),$$

with inverse denoted by $\mathcal{F}^{-1}$, will however appear in Section 2.3 while treating high modulations.

The use of Littlewood–Paley theory will be quite beneficial to our analysis and general references for it are the monographs of Taylor [35] and Grafakos [11]. We shall rely on the discrete version here only: the Littlewood--Paley projection $P_{\leq k}$, with $k \in \mathbb{Z}$, is defined to be a Fourier multiplier with symbol $m_{\leq k}(\xi) := m_{\leq 0}(2^{-k} |\xi|)$, i.e. via the convolution:

$$P_{\leq k}\phi(t, x) := 2^{2k} \int_{\mathbb{R}^2} \check{m}_{\leq 0}\left(2^k(x - y)\right)\phi(t, y)dy, \tag{2.1.1}$$

for some radial non-negative function $m_{\leq 0}(|\xi|)$ in frequency space, identically 1 on $|\xi| \leq 1$ and 0 for $|\xi| \geq 2$.

We also set P_k to be a multiplier with symbol $m_k(\xi) := m_0(2^{-k} |\xi|)$, where $m_0(|\xi|) := m_{\leq 0}(|\xi|) - m_{\leq 0}(2 |\xi|)$, and the operators $P_{<k}$, $P_{k_1 \leq \cdot \leq k_2}$, $P_{\geq k}$, etc. are then defined in the usual way. Note that LP-projections make sense for functions defined only at some given time t, or restricted to any time interval, and more generally commute with time cut-offs. Furthermore they are *disposable multipliers*, i.e. have the distributional convolution kernels of bounded mass, even when considered on the whole of space-time which in practice means that they are bounded on any translation-invariant Banach space of functions on $\mathbb{R} \times \mathbb{R}^2$ and therefore can be discarded from the estimates as one wishes.

Two elementary but important facts about LP-projections that we would like to mention here are the *finite band property* that states:

$$\|\nabla_x P_{\leq k}\phi\|_{L_x^p} \lesssim 2^k \|P_{\leq k}\phi\|_{L_x^p}, \tag{2.1.2}$$

and further:

$$\|\nabla_x P_k\phi\|_{L_x^p} \sim 2^k \|P_k\phi\|_{L_x^p}, \tag{2.1.3}$$

for any $1 \leq p \leq \infty$, as well as *Bernstein's inequality*:

$$\|P_k\phi\|_{L_x^p} \lesssim 2^{2k\left(\frac{1}{q} - \frac{1}{p}\right)} \|P_k\phi\|_{L_x^q}, \tag{2.1.4}$$

for any $1 \leq q \leq p \leq \infty$. The latter is especially useful converting integrability into regularity at low frequencies.

We can decompose any Schwartz function using LP-projections, and as we typically consider maps taking values in the sphere, we will be considering affinely (i.e. upon adding a constant) Schwartz functions, obtaining:

$$\phi = P_{\leq 0}\phi + \sum_{k>0} P_k\phi = \text{const.} + \sum_{k \in \mathbb{Z}} P_k\phi \quad \text{in } \mathcal{S}(\mathbb{R}^2). \tag{2.1.5}$$

While working with the gradient $\nabla_{t,x}\phi$, this will make no difference of course. By duality, the above decompositions hold also for tempered distributions and are used to define various Besov and Triebel-Lizorkin spaces, see [11]. Let us present here some examples important for our argument.

In this thesis, we will be mainly working with the Besov spaces $B_q^{s,p}(\mathbb{R}^2)$, for $s \in \mathbb{R}$ and $1 \leq p, q \leq \infty$, together with the homogeneous versions $\dot{B}_q^{s,p}(\mathbb{R}^2)$, defined as completions with respect to the norms:

$$\|\phi\|_{B_q^{s,p}}^q := \|P_{\leq 0}\phi\|_{L_x^p}^q + \sum_{k>0} 2^{qsk} \|P_k\phi\|_{L_x^p}^q, \quad \|\phi\|_{\dot{B}_q^{s,p}}^q := \sum_{k\in\mathbb{Z}} 2^{qsk} \|P_k\phi\|_{L_x^p}^q,$$

and taking the ℓ^∞ norm if $q = \infty$ instead, of subspaces of $\mathcal{S}(\mathbb{R}^2)$ for which those norms are finite. We remark that the case $p, q = 2$ corresponds to the familiar Sobolev spaces H_x^s, and their homogeneous versions $\dot{H}_x^s$ respectively.

We introduce also the local Hardy space $\mathcal{H}_{loc}^1(\mathbb{R}^2)$ with its homogeneous counterpart $\mathcal{H}^1(\mathbb{R}^2)$, as Triebel-Lizorkin spaces $F_2^{0,1}(\mathbb{R}^2) = \mathcal{H}_{loc}^1(\mathbb{R}^2)$ and $\dot{F}_2^{0,1}(\mathbb{R}^2) = \mathcal{H}^1(\mathbb{R}^2)$ (this characterization is obtained in [11]), both subspaces of L_x^1, defined as the completion of Schwartz functions with respect to the norms:

$$\|\phi\|_{F_2^{0,1}} := \|P_{\leq 0}\phi\|_{L_x^1} + \| \big(\sum_{k\geq 1} |P_k\phi|^2\big)^{1/2} \|_{L_x^1}, \quad \|\phi\|_{\dot{F}_2^{0,1}} := \| \big(\sum_{k\in\mathbb{Z}} |P_k\phi|^2\big)^{1/2} \|_{L_x^1},$$

and which admit the local and homogeneous BMO spaces as a duals, $(\mathcal{H}_{loc}^1)' = bmo$ and $(\mathcal{H}^1)' = \mathrm{BMO}$ respectively. Although the latter does not admit a Littlewood-Paley type characterization, the former does via the Triebel-Lizorkin space $F_2^{0,\infty} = bmo$, which is defined to be the Banach space of all tempered distributions $\varphi \in \mathcal{S}'(\mathbb{R}^2)$ having the following norm finite:

$$\|\varphi\|_{F_2^{0,\infty}} := \inf_{\{\varphi_k\}\subset L^\infty} \{\|P_{\leq 0}\varphi_0\|_{L^\infty} + \| \big(\sum_{k\geq 1} |P_k\varphi_k|^2\big)^{1/2} \|_{L^\infty} : \varphi = P_{\leq 0}\varphi_0 + \sum_{k\geq 1} P_k\varphi_k\},$$

the series above required to hold in $\mathcal{S}'$, see the monograph of Taylor [35] for further information. Hardy spaces are especially useful in estimating paraproducts (see below), and let us mention here, with this in mind, that $\mathcal{H}^1$ embeds into a Besov space with lower regularity but better summability:

$$\dot{F}_2^{0,1}(\mathbb{R}^2) \subset \dot{B}_1^{-1,2}(\mathbb{R}^2). \tag{2.1.6}$$

This fact, that we will exploit in the proof of Proposition 2.3.1 later, is taken from Lemma 7.19 of Krieger and Schlag [17] (page 250). For a related result in the Lorentz space setting see the monograph of Hélein [13] (Theorem 3.3.10 and also the references mentioned there).

2.1.2   The Littlewood–Paley trichotomy

Littlewood-Paley decompositions are also very useful in studying non-linear expressions, and one central example is the product $\theta\vartheta$ of two Schwartz functions θ and $\vartheta \in \mathcal{S}$. Applying the decomposition (2.1.5), we can write:

$$P_k\left(\theta\vartheta\right) = P_k \sum_{k_1, k_2} (P_{k_1}\theta)(P_{k_2}\vartheta),$$

but recalling that the Fourier transform of a product is a convolution leads to the so-called *Littlewood-Paley trichotomy* decomposition (also called paraproduct decomposition), which simplifies the above double sum into:

$$
\begin{aligned}
P_k\left(\theta\vartheta\right) = P_k[&\sum_{k_1, k_2 \geq k-6 : |k_1 - k_2| \leq O(1)} (P_{k_1}\theta)(P_{k_2}\vartheta) \\
&+ (P_{\leq k-7}\theta)(P_{k-3\leq\cdot\leq k+3}\vartheta) \\
&+ (P_{k-3\leq\cdot\leq k+3}\theta)(P_{\leq k-7}\vartheta)],
\end{aligned}
$$

•*the high-high interactions*: both θ and ϑ have Fourier support well above the scale $|\xi| \sim 2^k$, but the only way the sum of two annuli at larger scales $|\xi| \sim 2^{k_1}, 2^{k_2}$ with $k_1, k_2 \geq k + 6$ can intersect the small annulus at $|\xi| \sim 2^k$, is if they are approximately at the same scale, we should have $|k_1 - k_2| \leq 3$.

•*the low-high interactions*: if θ has Fourier support in the ball of radius 2^{k-6}, it will contribute to the frequency scale $|\xi| \sim 2^k$ if it is multiplied by ϑ frequency localized to the annuli $|\xi| \sim 2^{k_2}$ with $k - 3 \leq k_2 \leq k + 3$. The rougher components of ϑ bring up the low frequency parts of θ. The sum in k of the low-high interactions is sometimes called a *paraproduct* in the literature. By symmetry, we have the same picture with the roles of θ and ϑ interchanged: these are *the high-low interactions*.

We are then left only with the contribution of $\theta_{k_1}\vartheta_{k_2}$ where both terms are frequency localized at $2^{k_1}, 2^{k_2} \sim 2^k$, these are *the low-low interactions* and in our case it will be often convenient to incorporate them in the high-high interactions.

2.1.3 Tools for analysis in space-time

Finally, let us set up here the notation for some space-time function spaces and related tools that we use. We define the Sobolev spaces $H_{t,x}^s = H_{t,x}^s(\mathbb{R} \times \mathbb{R}^2)$, for $s \in \mathbb{R}$, by using the space-time Fourier transform and taking the completion of $\mathcal{S}(\mathbb{R} \times \mathbb{R}^2)$ with respect to the norm:

$$\|\psi\|_{H_{t,x}^s} := \left\| (1 + \tau^2 + |\xi|^2)^{\frac{s}{2}} \mathcal{F}\psi(\tau, \xi) \right\|_{L_{t,x}^2}.$$

We define the modulation projections $Q_{\leq j}$ and Q_j for $j \in \mathbb{Z}$ to be the Fourier multipliers with symbols:

$$m_0(|\frac{|\tau| - |\xi|}{2^j}|) \text{ and } m(|\frac{|\tau| - |\xi|}{2^j}|),$$

respectively (and similarly for $Q_{<j}$, $Q_{j_1 \leq \cdot \leq j_2}$ and $Q_{\geq j}$). We note that those are not disposable so that one needs to be careful when discarding them off from the estimates in general, but as their symbols are bounded and smooth, they are directly seen to be bounded on $L_{t,x}^2$ by Plancherel. Otherwise, we have the following lemma due to Tao (Lemmata 3 and 4 in [32]).

Lemma 2.1.1. *The operators $P_k Q_j$, $P_k Q_{\leq j}$, $P_{\leq k} Q_{\leq j}$ and $P_{\leq k} Q_j$ are disposable for any pair of integers j and k with $j \geq k + O(1)$. Moreover, for any $1 \leq p \leq \infty$ and $j, j_1, j_2 \in \mathbb{Z}$, the operators $Q_{\leq j}$, $Q_{j_1 \leq \cdot \leq j_2}$ and Q_j are bounded on the spaces $L_t^p(L_x^2)$.*

Using the modulation projections Q_j, we define following Tao [32] the homogeneous $\dot{X}_k^{s,b,q}$ spaces associated to the cone $\{|\tau| = |\xi|\}$ at the spatial frequency scale k, for any fixed integer $k \in \mathbb{Z}$ and some given real $b \in \mathbb{R}$, to be the completion of the space of Schwartz functions ψ on $\mathbb{R} \times \mathbb{R}^2$ with respect to the norm:

$$\|\psi\|_{\dot{X}_k^{s,b,q}} := 2^{sk} \left[\sum_j 2^{qbj} \|Q_j P_k \psi\|_{L_{t,x}^2}^q \right]^{\frac{1}{q}},$$

provided the latter is finite for ψ, and adopting the usual convention if q is infinite. For $q = 1$ we obtain an atomic space. As our methods here have more of an elliptic rather than dispersive character in the end, we shall not use those spaces directly (other than stating the estimates from regularity theory). However, the distinction between the *high modulations* regime $P_k Q_{>k+10}$, and the one of *frequency space-like* $P_k Q_{\leq k+10}$, is absolutely crucial for our analysis.

To close this section, let us recall here the convention that function spaces over domains are defined via minimal extensions. For example, we shall write $X(I)$, where

X is a function space over $\mathbb{R} \times \mathbb{R}^2$ and I some time interval, for the Banach space of functions f in $I \times \mathbb{R}^2$ admitting an extension f' to the whole of $\mathbb{R} \times \mathbb{R}^2$ and set:

$$\|f\|_{X(I)} := \inf \left\{ \|f'\|_X \ : \ f' \in X, \ \ f' = f \ \text{ on } \ I \times \mathbb{R}^2 \right\}.$$

2.2 Regularity theory for wave maps

2.2.1 Small data well-posedness theory

We shall not give here the full definition of the space DS, and its undifferentiated version S, used in the iteration arguments of the proofs of well-posedness for the wave maps equation, referring to [32] Section 10 or [29] Section 5.2, but we will briefly summarize here some characteristic properties.

At a given frequency scale $k \in \mathbb{Z}$, the space DS is defined as an intersection of several different spaces and for us it will be enough to note that we have the control:

$$\|P_k\psi\|_{L_t^\infty(L_x^2)} + \|P_k\psi\|_{\dot{X}_k^{0,\frac{1}{2},\infty}} + \sup_{(q,r):\frac{1}{q}+\frac{1}{2r}\leq\frac{1}{4}} 2^{(\frac{1}{q}+\frac{2}{r}-1)k} \|P_k\psi\|_{L_t^q(L_x^r)} \leq \|P_k\psi\|_{DS}, \quad (2.2.1)$$

for any Schwartz function ψ on $\mathbb{R} \times \mathbb{R}^2$ (under frequency localization, for the space S we have $P_k\phi \in S$ if $\nabla_{t,x}P_k\phi \in DS$ for a Schwartz ϕ). The first component is the natural energy component on which we should mainly rely in this work. The second one is the dispersive component to be used only indirectly here but being important in gaining extra regularity for the part of the wave map that has Fourier support away from the light cone. The latter observation is exploited by Sterbenz and Tataru [30] in their compactness result that we discuss below. The third component represents the standard Strichartz spaces.

We note that, for the regularity theory, the Q_0-null structure in the non-linearity of equation (1.1.2) is crucial and the components mentioned above are not enough by themselves to exploit it so that one needs to introduce further suitable null frame Strichartz spaces. However, as this structure will not play any direct role in our arguments we should not elaborate more on this point here. Let us simply remark in the end that DS contains the atomic Fourier restriction space:

$$\|P_k\psi\|_{DS} \lesssim \|P_k\psi\|_{\dot{X}_k^{0,\frac{1}{2},1}}, \quad (2.2.2)$$

referring to Lemma 8 in Tao's paper [32] for the proof of this fact, ideas from which we

will actually use later in the proof of Lemma 2.3.2.

By default in [29], the authors define the spaces DS and S as completions of Schwartz functions in $\mathbb{R} \times \mathbb{R}^2$ with respect to the norms obtained by ℓ^2-summing the control on the LP-projections and adding the L^∞ norm for S:

$$\|\psi\|_{DS}^2 := \sum_{k \in \mathbb{Z}} \|P_k \psi\|_{DS}^2, \quad \|\phi\|_S^2 := \|\phi\|_{L_{t,x}^\infty}^2 + \sum_{k \in \mathbb{Z}} \|\nabla_{t,x} P_k \phi\|_{DS}^2. \tag{2.2.3}$$

In practice however, it is sometimes convenient to replace the ℓ^2 summation in (2.2.3) with a control with respect to a *frequency envelope*. Following Sterbenz and Tataru [29], we call a sequence $c := \{c_k\}_{k \in \mathbb{Z}} \in \ell^2$ of positive numbers $c_k > 0$ a (σ_0, σ_1)-*admissible frequency envelope* if $0 < \sigma_0 < \sigma_1$ and for any $k_0 < k_1$ we have:

$$2^{-\sigma_0(k_1-k_0)} c_{k_1} \leq c_{k_0} \leq 2^{\sigma_1(k_1-k_0)} c_{k_1}.$$

Given some smooth initial data $\phi[0] = (\phi(0), \partial_t \phi(0))$ we can naturally attach to it an admissible frequency envelope by setting:

$$c_k^2 = \sum_{k_0 < k} 2^{-2\sigma_1(k-k_0)} \|P_{k_0} \nabla_{t,x} \phi(0)\|_{L_x^2}^2 + \sum_{k_1 \geq k} 2^{-2\sigma_0(k_1-k)} \|P_{k_1} \nabla_{t,x} \phi(0)\|_{L_x^2}^2, \tag{2.2.4}$$

for which we note that:

$$\left(\sum_{k \in \mathbb{Z}} 2^{2\sigma} c_k^2 \right)^{\frac{1}{2}} \sim \|\nabla_{t,x} \phi(0)\|_{\dot{H}_x^\sigma}, \quad -\sigma_0 < \sigma < \sigma_1, \tag{2.2.5}$$

so that given any function ψ on $\mathbb{R}^2$, $\|P_k \psi\|_{L_x^2} \lesssim c_k$ implies:

$$\|\psi\|_{\dot{H}_x^\sigma} \lesssim \|\nabla_{t,x} \phi(0)\|_{\dot{H}_x^\sigma}, \quad -\sigma_0 < \sigma < \sigma_1,$$

which is very useful in controlling the regularity of an evolution like the wave map.

Well-posedness theory for the wave maps equation with small energy initial data is due to Tao [32] for the Euclidean sphere as target, and Tataru [34] for general closed Riemannian manifolds, and also Krieger [16] who considered the hyperbolic plane as target. We will be using here a local version that we state below appearing as Theorem 1.3 in [34].

Theorem 2.2.1. (Tao [32], Tataru [34]). *There exists a constant $\epsilon_0 := \epsilon_0(\mathcal{N}) > 0$ such that:*

- Regularity: *given some smooth initial data $\phi[0] \in T\mathcal{N}$ at time $t = 0$ constant outside a compact domain with energy:*

$$\mathcal{E}[\phi](0) < \epsilon_0,$$

there exists a unique smooth wave map ϕ defined on the whole of Minkowski space $\mathbb{R}^{2+1}$ such that:

$$\|P_k \phi\|_S \lesssim c_k, \tag{2.2.6}$$

taking the frequency envelope c from (2.2.4) for $\phi[0]$ and where $\sigma_0 = \sigma_0(\mathcal{N})$ is some fixed small positive constant but σ_1 can be chosen arbitrarily large;

- Continuous dependence on initial data and rough solutions: *given a sequence of smooth tuples $\phi_\nu[0] \in T\mathcal{N}$ of initial data equal to a fixed constant outside some fixed compact domain, with energy:*

$$\mathcal{E}[\phi_\nu](0) < \epsilon_0,$$

and converging strongly in $H_x^1 \times L_x^2$ to some $\phi[0]$, there exist smooth wave maps ϕ_ν with the properties as stated in the first point above and a map:

$$\phi \in S,$$

solving weakly the wave maps equation (1.1.6), to which ϕ_ν converge in $C_t^0(H_x^1) \cap C_t^1(L_x^2)$ on bounded time intervals, and further for $0 < s < \sigma_0$:

$$\nabla_{t,x} \phi_\nu \to \nabla_{t,x} \phi \quad \text{in} \;\; DS \cap L_t^\infty(\dot{H}_x^{-s})(\mathbb{R}^{2+1}).$$

2.2.2 A compactness result under time-like control

We state now a compactness result due to Sterbenz and Tataru [30] for a sequence of small energy wave maps which become constant in the direction of some smooth time-like vector field. The absence of such a result in the general small energy case is precisely what makes the study of wave maps near the null boundary of the light cone a very challenging affair, requiring global non-linear techniques going beyond the present article. We mention that the arguments in [30] rely on the elliptic flavor given to the situation by the assumption that the sequence is asymptotically constant along a time-like vector field, the use of the Fourier restriction component of DS to gain compactness and regularity for the limiting map, as well as the small energy weak stability theory developed by Tataru [34] (which we have presented in the second point of Theorem

2.2.1 here).

Lemma 2.2.2. (Sterbenz and Tataru [30]). *Consider a sequence of smooth wave maps ϕ_ν in $[-3,3] \times B_3$ with small energy:*

$$\sup_{t \in [-3,3]} \mathcal{E}_{B_3}[\phi_\nu](t) \leq \epsilon_s, \tag{2.2.7}$$

where $\epsilon_s > 0$ depends only on ϵ_0 from Theorem 2.2.1, and such that:

$$\|X\phi_\nu\|_{L^2_{t,x}([-3,3] \times B_3)} \longrightarrow 0, \tag{2.2.8}$$

for some smooth time-like vector field X. Then there exists a wave map:

$$\phi \in H^{\frac{3}{2}-\epsilon}_{t,x}([-1,1] \times B_1), \tag{2.2.9}$$

for any $0 < \epsilon < \frac{1}{2}$, satisfying:

$$X\phi = 0 \ \ on \ \ [-1,1] \times B_1,$$

to which the maps ϕ_ν converge in $C^0_t(H^1_x) \cap C^1_t(L^2_x)$ after passing to a subsequence, and further:

$$\nabla_{t,x}\phi_\nu \longrightarrow \nabla_{t,x}\phi \ \ in \ \ DS(\{t \in [-1,1], \ r \leq 2 - |t|\}). \tag{2.2.10}$$

Remark 2.2.3. The proof of this lemma can be found in Proposition 5.1 of [30] and we remark that convergence in $H^1_{t,x}(U)$ for any domain $U \Subset (-3,3) \times B_3$ only is claimed there. But the stronger statement (2.2.10), to be understood in terms of minimal extensions, can be obtained as follows. Let us fix $U = [-\frac{5}{2}, \frac{5}{2}] \times B_{5/2}$, then upon passing to a further subsequence we would have:

$$\|\phi_\nu(t) - \phi(t)\|^2_{L^2_x(B_{\frac{5}{2}})} + \|\nabla_{t,x}\phi_\nu(t) - \nabla_{t,x}\phi(t)\|^2_{L^2_x(B_{\frac{5}{2}})} \longrightarrow 0 \ \text{ for a.e. } t, \tag{2.2.11}$$

therefore ϕ_ν converge strongly to ϕ in $(H^1_x \times L^2_x)(B_{5/2})$ for almost every t that we can fix as close to 0 as we wish. Hence, assuming that ϵ_s was chosen small enough initially, by the pigeonhole principle we have for $\sigma \in (2, \frac{5}{2})$:

$$\fint_{\partial B_\sigma} |\nabla_{t,x}\phi(t)|^2 \, d\theta \lesssim \epsilon_s,$$

away from a set of measure $\frac{1}{10}$ say. Fixing such a σ, we would have $\phi(t, \partial B_\sigma)$ contained

in a single chart of $\mathcal{N}$ of diameter $O(\sqrt{\epsilon_s})$ around a point $c \in \mathcal{N}$. Moreover, upon passing to a further subsequence, by the strong convergence (2.2.11) we can choose $\sigma \in (2, \frac{5}{2})$ such that $\phi_\nu(t)|_{\partial B_\sigma} \to \phi(t)|_{\partial B_\sigma}$ in the Hölder space $C^\alpha(\partial B_\sigma)$ with $\alpha \in (0, \frac{1}{2})$, using Morrey's inequality. Hence, we would have $\phi_\nu(t, \partial B_\sigma)$ contained in the chart around $c \in \mathcal{N}$ of diameter $O(\sqrt{\epsilon_s})$ as well, for all $\nu \in \mathbb{N}$ large enough. Therefore, we can construct extensions $\phi'_\nu[t] \in T\mathcal{N}$ of $\phi_\nu[t]|_{B_\sigma}$, smooth as the latter are, with the energy bound:

$$\mathcal{E}[\phi'_\nu](t) \lesssim \epsilon_s,$$

by smoothly interpolating between $\phi_\nu[t]|_{\partial B_\sigma}$ and $(c, 0) \in T\mathcal{N}$ on $B_3 \setminus B_\sigma$. By construction, we obtain $\phi'_\nu[t]$ strongly convergent in $H^1_x \times L^2_x$ to some map $\phi'[t]$ agreeing with $\phi[t]$ on B_2. In the end, setting the constant $\epsilon_s > 0$ small enough and the time t close enough to 0, the convergence statements are justified by the continuous dependence on the initial data part of Theorem 2.2.1 and the finite speed of propagation property.

In particular, the assumption (2.2.8) gets upgraded to:

$$X\phi_\nu \longrightarrow 0 \ \text{in} \ C^0_t(L^2_x)\left([-1, 1] \times B_1\right),$$

and going further, the regularity theory of Theorem 2.2.1 tells us that in fact we have:

$$\phi \in C^0_t([-1, 1] \,; H^{\frac{3}{2}-\epsilon}_x(B_1)) \cap C^1_t([-1, 1] \,; H^{\frac{1}{2}-\epsilon}_x(B_1)),$$

for any $0 < \epsilon < \frac{1}{2}$ improving upon (2.2.9), although it is unfortunately impossible to obtain convergence in such a stronger space without further assumptions, especially regarding the decay (2.2.8).

2.2.3 Energy dispersed wave maps

Let us close this section by mentioning the result of Sterbenz and Tataru [29], see both Theorem 1.3 and Proposition 3.9 there, which relaxes the assumption of small energy in the work of Tao [32] and Tataru [34] to small *energy dispersion*. This represents a crucial technical ingredient in the proof by Sterbenz and Tataru [30] of the threshold conjecture. Let us consider an open interval $I = (t_0, t_1)$, which can be unbounded.

Theorem 2.2.4. (Sterbenz and Tataru [29]). *Given an energy bound $\mathcal{E} > 0$, there exist constants $0 < \epsilon(\mathcal{N}, \mathcal{E}) \ll 1$ and $1 \ll F(\mathcal{N}, \mathcal{E})$ such that for any smooth wave map ϕ on*

(t_0, t_1) *with energy bounded by* $\mathcal{E}$ *and* $\nabla_{t,x}\phi$ *spatially Schwartz, if we have:*

$$\sup_k \|P_k\phi\|_{L^\infty_{t,x}(t_0,t_1)} \le \epsilon(\mathcal{N}, \mathcal{E}),$$

then

$$\|\phi\|_{S(t_0,t_1)} \le F(\mathcal{N}, \mathcal{E}).$$

Moreover, considering an admissible frequency envelope c *attached to some* $\phi[t]$ *for* $t_0 < t < t_1$, *as in (2.2.4) and* σ_0 *as in Theorem 2.2.1, we obtain:*

$$\|P_k\phi\|_{S(t_0,t_1)} \lesssim c_k,$$

and the map ϕ *extends to a smooth wave map on a neighborhood of the time interval* (t_0, t_1).

2.3 Compensation type estimates

We prove here two compensation estimates for wave maps into spheres with a good bound in the direction of some constant time-like vector field, relying on the conservation law (1.1.9) to treat high-high frequency interactions (this phenomena goes back essentially to Henry Wente, see the references in [13]). These estimates will play a key role in the proof of no loss of energy in formation of solitons, and as in the case of higher dimensional harmonic maps considered by Lin and Rivière [21], this is the only place where we use the fact that our target manifold is $\mathbb{S}^{n-1}$.

2.3.1 The hyperbolic version of the div-curl lemma

Proposition 2.3.1. *Let* $\phi : [-1, 1] \times \mathbb{R}^2 \to \mathbb{S}^{n-1}$ *be a smooth wave map equal to a constant c outside a compact domain in space, with energy bounded by some positive* $\mathcal{E} > 0$:

$$\|\nabla_{t,x}\phi\|^2_{L^\infty_t(L^2_x)[-1,1]} \le \mathcal{E}, \tag{2.3.1}$$

and X a constant time-like vector field, that we may take to be:

$$X = \cosh(\zeta)\partial_t + \sinh(\zeta)\partial_{x_1}, \tag{2.3.2}$$

for some rapidity constant $\zeta \geq 0$. Denote by $\chi = \chi(t) \in C_0^\infty(-1,1)$ a smooth time cut-off function, then there exists a decomposition holding in $\mathcal{S}(\mathbb{R} \times \mathbb{R}^2)$:

$$\chi \nabla_{t,x} \phi = \Theta_X + \Xi_X, \tag{2.3.3}$$

satisfying:

$$\|\Theta_X\|_{L_{t,x}^2} \lesssim \|X\phi\|_{L_{t,x}^2[-1,1]} + \|\phi - c\|_{L_t^\infty(L_x^2)[-1,1]} \tag{2.3.4}$$

and

$$\sum_{k \in \mathbb{Z}} \|P_k \Xi_X\|_{L_t^1(L_x^2)} \lesssim 1, \tag{2.3.5}$$

with the implicit constants depending only on n the dimension of $\mathbb{R}^n$, the energy bound $\mathcal{E}$, the rapidity constant ζ and the cut-off χ (most notably on $\|\partial_t \chi\|_{L_t^\infty}$).

Proof. We start by noting that, expressing ∂_t as a linear combination of X and ∂_{x_1} via (2.3.2), it suffices to consider the spatial gradient $\chi \nabla_x \phi$.

For low frequencies, we proceed claiming immediately:

$$\|\chi P_{\leq 0} \nabla_x \phi\|_{L_{t,x}^2} \lesssim \|\phi - c\|_{L_t^\infty(L_x^2)[-1,1]}, \tag{2.3.6}$$

which simply follows from the finite band property (2.1.2), passing to $L_t^\infty(L_x^2)$ as necessary. This is an acceptable contribution.

For high modulations, we claim:

$$\left\| \sum_{k \in \mathbb{Z}} Q_{\geq k+10} P_k[\chi \nabla_x \phi] \right\|_{L_{t,x}^2} \lesssim \|X\phi\|_{L_{t,x}^2[-1,1]} + \|\phi - c\|_{L_t^\infty(L_x^2)[-1,1]}, \tag{2.3.7}$$

and the idea here, as in [30], is to note that the vector field X being time-like, the Fourier multiplier $X^{-1} \nabla_x Q_{\geq k+10} \widetilde{P}_k$, where $\widetilde{P}_k = P_{k-1 \leq \cdot \leq k+1}$, has symbol smooth and bounded uniformly in $k \in \mathbb{Z}$. By Plancherel in $L_{t,x}^2$, this gives rise to the favorable elliptic estimate:

$$\|Q_{\geq k+10} P_k[\chi \nabla_x \phi]\|_{L_{t,x}^2} \lesssim \|\chi P_k X\phi\|_{L_{t,x}^2} + \|(\partial_t \chi) P_k \phi\|_{L_{t,x}^2}, \tag{2.3.8}$$

and so (2.3.7) follows square-summing in k the above and dropping the cut-off. This is again acceptable.

The main term to consider is $Q_{<k+10} P_k(\chi \nabla_x \phi)$ with $k > 0$, and for this we rely on the wave maps equation (1.1.6), that we rewrite to make the vector X to appear,

introducing the operator:

$$\Delta_{x,\beta} := (1 - \beta^2)\partial_{x_1}^2 + \partial_{x_2}^2, \ \ \beta := \tanh(\zeta) \in [0,1), \tag{2.3.9}$$

which is elliptic in the frequency region considered. So, using (2.3.2), together with (1.1.5), we rewrite the wave maps equation (1.1.6) as:

$$\Delta_{x,\beta}(\chi\phi) = - \chi(\phi\partial_\alpha\phi^\dagger - \partial_\alpha\phi\phi^\dagger)\partial^\alpha\phi \tag{2.3.10}$$
$$+ \operatorname{sech}^2(\zeta)(X - 2\sinh(\zeta)\partial_{x_1})(\chi X\phi) - \operatorname{sech}(\zeta)(\partial_t\chi)X\phi,$$

and inverting $\Delta_{x,\beta}$ we have:

$$P_{>0}\chi\nabla_x\phi = \frac{\nabla_x}{\Delta_{x,\beta}}P_{>0}(\Delta_{x,\beta}(\chi\phi)),$$

holding in $\mathcal{S}(\mathbb{R} \times \mathbb{R}^2)$, hence let us treat each term in (2.3.10) one by one.

Considering the second line in (2.3.10), we control the first two terms by claiming, for any $k \in \mathbb{Z}$:

$$\left\| \nabla_x \frac{\nabla_{t,x}}{\Delta_{x,\beta}} Q_{<k+10} P_k(\chi X\phi) \right\|_{L^2_{t,x}} \lesssim \|P_k X\phi\|_{L^2_{t,x}[-1,1]}, \tag{2.3.11}$$

which follows immediately discarding, via Plancherel in $L^2_{t,x}$, the Fourier multiplier $\nabla_x\nabla_{t,x}\Delta_{x,\beta}^{-1}Q_{<k+10}\widetilde{P}_k$ of symbol bounded uniformly in $k \in \mathbb{Z}$, and dropping the time cut-off χ. For the third term, we have, for any $k \in \mathbb{Z}$:

$$\left\| \frac{\nabla_x}{\Delta_{x,\beta}} Q_{<k+10} P_k[(\partial_t\chi)X\phi] \right\|_{L^2_{t,x}} \lesssim 2^{-k} \|\partial_t\chi\|_{L^\infty_{t,x}} \|P_k X\phi\|_{L^2_{t,x}[-1,1]}, \tag{2.3.12}$$

where we discarded by Plancherel in $L^2_{t,x}$ the Fourier multiplier $2^k\nabla_x\Delta_{x,\beta}^{-1}Q_{<k+10}\widetilde{P}_k$, having here again the symbol bounded uniformly in $k \in \mathbb{Z}$. Therefore, square-summing over $k > 0$, both (2.3.11) and (2.3.12) lead to acceptable contributions.

We consider now the non-linear term on the first line of (2.3.10). Let us introduce some notation for the connection matrices:

$$\Omega_\alpha := \phi\partial_\alpha\phi^\dagger - \partial_\alpha\phi\phi^\dagger, \ \text{with } \partial^\alpha\Omega_\alpha = 0 \ \text{and} \ \|\Omega_\alpha\|_{L^\infty_t(L^2_x)[-1,1]} \lesssim 1, \tag{2.3.13}$$

by (1.1.9), respectively the global energy bound (2.3.1) and the boundedness of the

wave map. We claim then the following compensation estimate:

$$\sum_{k>0} \left\| \frac{\nabla_x}{\Delta_{x,\beta}} Q_{<k+10} P_k(\chi\Omega_\alpha\partial^\alpha\phi) \right\|_{L^1_t(L^2_x)} \lesssim 1. \tag{2.3.14}$$

Thanks to the conservation law, the term $\Omega_\alpha\partial^\alpha\phi$ exhibits a div-curl type structure, and we should treat this using the Littlewood-Paley trichotomy in very much the same standard way as the actual div-curl structure, see Taylor's monograph [35]. We start by writing:

$$P_k\left(\chi\Omega_\alpha\partial^\alpha\phi\right) = P_k[(\partial_t\chi) \sum_{k_1,k_2\geq k-6:|k_1-k_2|\leq O(1)} \Omega_{\alpha,k_1}\phi_{k_2} \tag{2.3.15}$$

$$+ \partial^\alpha \sum_{k_1,k_2\geq k-6:|k_1-k_2|\leq O(1)} \chi\Omega_{\alpha,k_1}\phi_{k_2}$$

$$+ \chi\Omega_{\alpha,\leq k-7}\partial^\alpha\phi_{k-3\leq\cdot\leq k+3}$$

$$+ \chi\Omega_{\alpha,k-3\leq\cdot\leq k+3}\partial^\alpha\phi_{\leq k-7}],$$

where $\Omega_{\alpha,k_1} := P_{k_1}\Omega_\alpha$ and similarly for ϕ_{k_2}, $\Omega_{\alpha,\leq k_1}$, etc. We are going to prove claim (2.3.14) for each of the terms in (2.3.15) separately. Note that the Fourier multipliers:

$$\frac{\nabla_x\nabla_{t,x}}{\Delta_{x,\beta}} Q_{<k+10}\widetilde{P}_k \ \ \text{and} \ \ \frac{2^k\nabla_x}{\Delta_{x,\beta}} Q_{<k+10}\widetilde{P}_k, \tag{2.3.16}$$

are disposable, which is essentially contained in Lemma 2.1.1 (precomposing, for example, with the space-time LP-projections to $|\tau| + |\xi| \sim 2^k$ that we do not use here otherwise). This justifies the fact that we can work with the space $L^1_t(L^2_x)$ instead of $L^2_{t,x}$ (on which, of course, (2.3.16) are bounded by Plancherel).

Let us start with the high-high interactions on the first and second lines of (2.3.15), for which we control (2.3.14), discarding the multipliers (2.3.16) and dropping $2^{-k}\partial_t\chi$ for the first term, by:

$$\sum_{k>0} \left\| P_k \sum_{k_1,k_2\geq k-6:|k_1-k_2|\leq O(1)} \Omega_{\alpha,k_1}\phi_{k_2} \right\|_{L^1_t(L^2_x)[-1,1]} \tag{2.3.17}$$

$$\lesssim \sup_{t\in[-1,1]} \sum_{k>0} 2^k \sum_{k_1,k_2\geq k-6:|k_1-k_2|\leq O(1)} \left\| \Omega_{\alpha,k_1}(t)\phi_{k_2}(t) \right\|_{L^1_x},$$

where we applied Bernstein's inequality (2.1.4), commuted the sum $\sum_{k>0}$ with L^1_t and

discarded P_k. Using Cauchy-Schwarz in L_x^1 and recalling the finite band property (2.1.3) for ϕ_{k_2}, we can bound the contribution of (2.3.17) via:

$$\sup_{t\in[-1,1]} \sum_{k>0} \sum_{k_1,k_2\geq k-6:|k_1-k_2|\leq O(1)} 2^{-(k_2-k)} \left\|\Omega_{\alpha,k_1}(t)\right\|_{L_x^2} \left\|\nabla_x\phi_{k_2}(t)\right\|_{L_x^2},$$

and summing this over $k > 0$, letting $i := k_1 - k_2$ and $j := k_2 - k$, we obtain:

$$\sup_{t\in[-1,1]} \sum_{i=O(1)} \sum_{j\geq O(1)} 2^{-j} \sum_{k>0} \left\|\Omega_{\alpha,k+j+i}(t)\right\|_{L_x^2} \left\|\nabla_x\phi_{k+j}(t)\right\|_{L_x^2}$$

$$\lesssim \sup_{t\in[-1,1]} \left(\sum_{k_1\geq O(1)} \left\|\Omega_{\alpha,k_1}(t)\right\|_{L_x^2}^2\right)^{\frac{1}{2}} \left(\sum_{k_2\geq O(1)} \left\|\nabla_x\phi_{k_2}(t)\right\|_{L_x^2}^2\right)^{\frac{1}{2}},$$

where we have used Cauchy-Schwarz in k. By the global energy bound, we get that high-high interactions make an acceptable contribution to (2.3.14).

Finally, let us consider the contribution of the paraproducts from lines three and four in (2.3.15), and we focus on the latter as the former is treated in the same way by symmetry (or in fact, could have already been absorbed in the argument for high-high interactions). Here, the div-curl structure is not playing any role, and is actually counter-productive. Hence, discarding the second multiplier from (2.3.16) and commuting the discrete sum $\sum_{k>0}$ with L_t^1 as previously, it suffices control:

$$\sup_{t\in[-1,1]} \sum_{k>0} 2^{-k} \left\|P_k[\Omega_{\alpha,k-3\leq\cdot\leq k+3}(t)\partial^\alpha\phi_{\leq k-7}(t)]\right\|_{L_x^2}.$$

Recalling the embedding (2.1.6) we are reduced to showing:

$$\sup_{t\in[-1,1]} \left\|\sum_{k>0} P_k[\Omega_{\alpha,k-3\leq\cdot\leq k+3}(t)\partial^\alpha\phi_{\leq k-7}(t)]\right\|_{F_2^{0,1}(\mathbb{R}^2)} \lesssim 1.$$

Using the duality $(F_2^{0,1})' = F_2^{0,\infty}$, as discussed in section 2.1, we take an arbitrary $\varphi \in F_2^{0,\infty}$ together with a representation $\varphi = \sum_{k\geq 0} \varphi_k$ in $\mathcal{S}_x'$ such that each φ_k has Fourier support in $|\xi| \sim 2^k$ ($|\xi| \lesssim 1$ for φ_0) and:

$$\left\|\left(\sum_{k\geq 0}|\varphi_k|^2\right)^{1/2}\right\|_{L_x^\infty} \leq 2\left\|\varphi\right\|_{F_2^{0,\infty}}.$$

Then, recalling the fact that LP-projections are self-adjoint, we must show that:

$$\sum_{j=O(1)} \sum_{k\geq 0} \int |\Omega_{\alpha,k-3\leq\cdot\leq k+3}(t)\partial^\alpha \phi_{\leq k-7}(t)\varphi_{k+j}|\, dx \lesssim \|\varphi\|_{F_2^{0,\infty}},$$

with the convention that φ_k with k negative simply stands for φ_0. Using Cauchy-Schwartz we bound this via:

$$\left\| \Big(\sum_{k\geq 0} |\Omega_{\alpha,k-3\leq\cdot\leq k+3}(t)|^2\Big)^{1/2}\right\|_{L_x^2} \left\| \sup_{k\in\mathbb{Z}} |P_{\leq k}\nabla_{t,x}\phi(t)| \right\|_{L_x^2} \sum_{j=O(1)} \left\| \Big(\sum_{k\geq 0} |\varphi_{k+j}|^2\Big)^{1/2}\right\|_{L_x^\infty}.$$

It is a well-known fact from harmonic analysis, to which we shall refer as the *Littlewood–Paley square-function estimate*, see e.g. [35] (Section 2 of Chapter 1), that:

$$\| \Big(\sum_{k\in\mathbb{Z}} |\Omega_{\alpha,k}(t)|^2\Big)^{1/2} \|_{L_x^2} \lesssim \|\Omega_\alpha(t)\|_{L_x^2},$$

$$\| \sup_{k\in\mathbb{Z}} |P_{\leq k}\nabla_{t,x}\phi(t)| \|_{L_x^2} \lesssim \|\nabla_{t,x}\phi(t)\|_{L_x^2}.$$

Hence, by the global energy bound, the contribution of the paraproducts is acceptable. Therefore we have shown the compensation estimate (2.3.14).

Proposition 2.3.1 is proved. □

2.3.2 Controlling the second order time-like derivatives

We present now a compensation estimate for higher order time-like derivatives of wave maps as considered in the previous proposition. It holds up to a non-linear bulk, essentially quadratic in the gradient and local in time, that we shall consider on neck regions later in the proof of the weak Besov $\dot{B}_\infty^{1,2}$ decay estimate in Lemma 3.3.3. Parts of this estimate are non-linear, and will be established via a duality argument in the spirit of the energy collapsing result itself.

As for Proposition 2.3.1, the conservation law (1.1.9) is absolutely crucial, and so our arguments do not generalize directly to the case of a general target beyond the Euclidean sphere $\mathbb{S}^{n-1}$.

Lemma 2.3.2. *Consider a wave map* $\phi : [-1,1] \times \mathbb{R}^2 \to \mathbb{S}^{n-1}$ *with the same set-up as in Proposition 2.3.1, then we have the following decomposition holding in* $\mathcal{S}(\mathbb{R} \times \mathbb{R}^2)$,

using notation from (2.3.13) and recalling that $\beta = \tanh(\zeta)$:

$$\text{sech}^2(\zeta)\chi X^2\phi = -\sum_{k\in\mathbb{Z}} P_k\left[\chi((1-\beta^2)\partial_{x_1}\Omega_{x_1} + \partial_{x_2}\Omega_{x_2})(P_{>k+10}\phi)\right] \tag{2.3.18}$$
$$+ \text{sech}^2(\zeta)\chi\left(-\Omega_X X\phi + \sinh(\zeta)(\Omega_X\partial_{x_1}\phi + \Omega_{x_1}X\phi)\right)$$
$$+ \Pi_X,$$

the error term satisfying:

$$\sum_{k\in\mathbb{Z}} 2^{-2k}\left\|P_k\Pi_X\right\|_{L^2_{t,x}[-1,1]}^2 \tag{2.3.19}$$
$$\lesssim (1 + \|X\phi\|_{L^2_{t,x}[-1,1]} + \|\phi - c\|_{L^\infty_t(L^2_x)[-1,1]})(\|X\phi\|_{L^2_{t,x}[-1,1]} + \|\phi - c\|_{L^\infty_t(L^2_x)[-1,1]}),$$

with the same dependence as for the implicit constant as in Proposition 2.3.1.

Let us note here, for later use in Lemma 3.3.3, that we can rewrite the decomposition (2.3.18) as follows. Introducing the notation:

$$\Omega_{x,\beta} := (1-\beta^2)\Omega_{x_1}dx_1 + \Omega_{x_2}dx_2,$$

we can write:

$$\text{sech}^2(\zeta)\chi X^2\phi - \chi\Omega_\alpha\partial^\alpha\phi = -\sum_{k\in\mathbb{Z}}\chi P_k\left[\nabla_x\cdot(\Omega_{x,\beta}P_{>k+10}\phi) + \Omega_{x,\beta}\cdot P_{\leq k+10}\nabla_x\phi\right]$$
$$+ \Pi_X.$$

To obtain this equation, one should rely on the conservation law (1.1.9), making the vector X appear through (2.3.2) and adding up on both sides some low-high interactions - see the proof below for more details.

Proof. Let us start with the frequency space-like region, that we can treat directly and for which we claim the stronger estimate:

$$2^{-k}\left\|P_kQ_{<k+10}(\chi X^2\phi)\right\|_{L^2_{t,x}[-1,1]} \lesssim \|P_kX\phi\|_{L^2_{t,x}[-1,1]}, \tag{2.3.20}$$

for any $k \in \mathbb{Z}$. To see this, we simply commute X with the time cut-off χ, getting:

$$2^{-k}\left\|P_kQ_{<k+10}(\chi X^2\phi)\right\|_{L^2_{t,x}[-1,1]} \lesssim \|P_k(\chi X\phi)\|_{L^2_{t,x}} + 2^{-k}\left\|P_kQ_{<k+10}(\partial_t\chi X\phi)\right\|_{L^2_{t,x}[-1,1]},$$

where for the first term we discarded the multiplier $2^{-k}X\widetilde{P}_k Q_{<k+10}$ using Plancherel in $L^2_{t,x}$. Regarding the second one, passing to $L^\infty_t(L^2_x)$, which is possible as we are working over a bounded time interval in (2.3.20), we can apply the inversion formula for the space-time Fourier transform $\mathcal{F}$, to get:

$$2^{-k}\left\|P_k Q_{<k+10}(\partial_t \chi X\phi)\right\|_{L^\infty_t(L^2_x)} \lesssim 2^{-k}\left\|\mathcal{F}P_k Q_{<k+10}(\partial_t \chi X\phi)\right\|_{L^1_\tau(L^2_\xi)},$$

combining Minkowski's inequality and then Plancherel in L^2_x. But the integrand on the RHS has τ-support of length $O(2^k)$, hence we can bound this simply via:

$$\left\|\mathcal{F}P_k(\partial_t \chi X\phi)\right\|_{L^\infty_\tau(L^2_\xi)} \lesssim \left\|P_k(\partial_t \chi X\phi)\right\|_{L^1_t(L^2_x)} \lesssim \left\|\partial_t \chi\right\|_{L^2_t(L^\infty_x)}\left\|P_k X\phi\right\|_{L^2_{t,x}[-1,1]},$$

where we applied the inversion formula for $\mathcal{F}^{-1}$ this time (note that this argument is essentially a manifestation of Bernstein's one dimensional inequality). This gives claim (2.3.20) as desired.

For high modulations, we use the wave maps equation as in (2.3.10). Following the Littlewood-Paley trichotomy (passing to the convention $\phi_k := P_k\phi$, etc. as before), we write:

$$
\begin{aligned}
P_k Q_{\geq k+10}(\operatorname{sech}^2(\zeta)\chi X^2\phi) = P_k Q_{\geq k+10}\big[&\Delta_{x,\beta}(\chi\phi) + 2\operatorname{sech}^2(\zeta)\sinh(\zeta)\chi\partial_{x_1}X\phi \\
&+ \operatorname{sech}^2(\zeta)\chi\left(-\Omega_X X\phi + \sinh(\zeta)(\Omega_X \partial_{x_1}\phi + \Omega_{x_1}X\phi)\right) \\
&+ \chi\Omega_{x,\beta}\cdot\nabla_x\phi_{\leq k+10} + \chi\nabla_x\cdot(\Omega_{x,\beta}\phi_{>k+10}) \\
&- \chi(\nabla_x\cdot\Omega_{x,\beta})(\phi_{>k+10})\big],
\end{aligned}
$$

where we set:
$$\Omega_X := \cosh(\zeta)\Omega_t + \sinh(\zeta)\Omega_{x_1}.$$

From there, we add and subtract the frequency space-like part of the terms on second and last lines above, and use the conservation law (1.1.9), that we rewrite as:

$$\nabla_x\cdot\Omega_{x,\beta} = \operatorname{sech}^2(\zeta)(X\Omega_X - \sinh(\zeta)(\partial_{x_1}\Omega_X + X\Omega_{x_1})).$$

This yields the following decomposition:

$$P_k Q_{\geq k+10}(\operatorname{sech}^2(\zeta)\chi X^2 \phi) = P_k Q_{\geq k+10}\left[\Delta_{x,\beta}(\chi\phi)\right. \tag{2.3.21}$$
$$+ 2\operatorname{sech}^2(\zeta)\sinh(\zeta)\chi\partial_{x_1}X\phi]$$
$$+ \operatorname{sech}^2(\zeta)\chi P_k\left[-\Omega_X X\phi + \sinh(\zeta)(\Omega_X\partial_{x_1}\phi + \Omega_{x_1}X\phi)\right]$$
$$+ \operatorname{sech}^2(\zeta)Q_{<k+10}[\psi_k^{(1)} + \psi_k^{(2)} + \psi_k^{(3)} + \psi_k^{(4)}]$$
$$+ Q_{\geq k+10}[\varphi_k^{(1)} + \varphi_k^{(2)}] - P_k\left[\chi(\nabla_x \cdot \Omega_{x,\beta})(\phi_{>k+10})\right],$$

where we define:

$$\psi_k^{(1)} := \chi P_k\left[\Omega_X X\phi_{\leq k+10} - \sinh(\zeta)(\Omega_X\partial_{x_1}\phi_{\leq k+10} + \Omega_{x_1}X\phi_{\leq k+10})\right],$$
$$\psi_k^{(2)} := P_k\left[(X\chi)(-\Omega_X + \sinh(\zeta)\Omega_{x_1})\phi_{>k+10}\right]$$
$$\psi_k^{(3)} := P_k\left[[X - \sinh(\zeta)\partial_{x_1}](\chi\Omega_X\phi_{>k+10})\right],$$
$$\psi_k^{(4)} := P_k\left[-\sinh(\zeta)X(\chi\Omega_{x_1}\phi_{>k+10})\right],$$

as well as:

$$\varphi_k^{(1)} := P_k\left[\chi\Omega_{x,\beta} \cdot \nabla_x\phi_{\leq k+10}\right],$$
$$\varphi_k^{(2)} := P_k\left[\chi\nabla_x \cdot [\Omega_{x,\beta}\phi_{>k+10}]\right].$$

We proceed proving the estimate (2.3.19) for the first line of (2.3.21) and each of the $\psi_k^{(i)}$ and $\varphi_k^{(i)}$ separately.

For the Laplacian, inverting X, we have the stronger estimate:

$$2^{-k}\left\|P_k Q_{\geq k+10}\Delta_{x,\beta}(\chi\phi)\right\|_{L^2_{t,x}[-1,1]} \lesssim \|P_k X\phi\|_{L^2_{t,x}[-1,1]} + \|P_k\phi\|_{L^2_{t,x}[-1,1]},$$

that follows immediately by discarding, via Plancherel in $L^2_{t,x}$, the Fourier multiplier $2^{-k}X^{-1}\Delta_{x,\beta}\widetilde{P}_k Q_{\geq k+10}$ having symbol bounded uniformly in $k \in \mathbb{Z}$, which leads to an acceptable contribution.

For the second term on the RHS of (2.3.21) we immediately have:

$$2^{-k}\left\|P_k Q_{\geq k+10}[\chi\partial_{x_1}X\phi]\right\|_{L^2_{t,x}} \lesssim \|P_k X\phi\|_{L^2_{t,x}[-1,1]},$$

by the finite band property (2.1.3), which is acceptable.

Regarding $\psi_k^{(1)}$, we remark that it has a paraproduct structure and so at least one of the factors will be frequency localized to $|\xi| \sim 2^k$, which is favorable for square-

summing. More precisely, discarding $Q_{<k+10}$ before dropping the cut-off χ, and using Bernstein's inequality (2.1.4) to pass to $L_t^2(L_x^1)$, it is enough to note that for any $1 \leq p, q, r \leq n$ and any time slice $t \in [-1, 1]$:

$$\sum_{k \in \mathbb{Z}} \sum_{k'=k+O(1)} \Big(\|(\phi^p X \phi^q \nabla_{t,x} \phi_{k'}^r)(t)\|_{L_x^1}^2 + \|(\phi^p \nabla_{t,x} \phi^q X \phi_{k'}^r)(t)\|_{L_x^1}^2$$
$$+ \|(P_{k'}[\phi^p \nabla_{t,x} \phi^q] X \phi^r)(t)\|_{L_x^1}^2 + \|(P_{k'}[\phi^p X \phi^q] \nabla_{t,x} \phi^r)(t)\|_{L_x^1}^2 \Big)$$
$$\lesssim \|\nabla_{t,x} \phi(t)\|_{L_x^2}^2 \|X \phi(t)\|_{L_x^2}^2 ,$$

by Cauchy-Schwarz. Upon integrating in time, this is an acceptable contribution by the energy bound (2.3.1).

For the expression $\psi_k^{(2)}$, it is already convenient to proceed via a duality argument:

$$\sum_{k \in \mathbb{Z}} 2^{-2k} \left\| Q_{<k+10} \psi_k^{(2)} \right\|_{L_{t,x}^2}^2 \lesssim \sum_{k \in \mathbb{Z}} \left\| \psi_k^{(2)} \right\|_{L_{t,x}^1} 2^{-k} \left\| Q_{<k+10} \psi_k^{(2)} \right\|_{L_t^\infty(L_x^2)}$$
$$\lesssim \left(\sum_{k \in \mathbb{Z}} 2^k \left\| \psi_k^{(2)} \right\|_{L_{t,x}^1} \right) \left(\sup_{k \in \mathbb{Z}} \left\| \psi_k^{(2)} \right\|_{L_{t,x}^1} \right) ,$$

where we used Bernstein (2.1.4) for the first factor, and for the second one we proceeded as for the frequency space-like term (2.3.20), using time frequency localization to estimate it via the Fourier inversion formula:

$$2^{-k} \left\| Q_{<k+10} \psi_k^{(2)} \right\|_{L_t^\infty(L_x^2)} \lesssim \left\| \psi_k^{(2)} \right\|_{L_t^1(L_x^2)} .$$

The first factor is universally bounded for us, as for any $1 \leq p, q, r \leq n$:

$$\sum_{k \in \mathbb{Z}} 2^k \left\| P_k[(\partial_t \chi)(\phi^p \nabla_{t,x} \phi^q) \phi_{>k+10}^r] \right\|_{L_{t,x}^1} \lesssim \|\partial_t \chi\|_{L_t^1(L_x^\infty)} \|\nabla_{t,x} \phi\|_{L_t^\infty(L_x^2)[-1,1]}^2 ,$$

which follows directly from the analogous treatment of high-high interactions in the proof of Proposition 2.3.1. On the other hand, the second factor is controlled via:

$$\left\| P_k[(\partial_t \chi)(\phi^p \nabla_{t,x} \phi^q) \phi_{>k+10}^r] \right\|_{L_{t,x}^1} \lesssim \|\partial_t \chi\|_{L_t^1(L_x^\infty)} \|\nabla_{t,x} \phi\|_{L_t^\infty(L_x^2)[-1,1]} \|\phi - c\|_{L_t^\infty(L_x^2)[-1,1]} ,$$

which yields an acceptable contribution to the non-linear part of (2.3.19).

Regarding $\psi_k^{(3)}$, it is a linear combination of:

$$\sum_{k\in\mathbb{Z}} 2^{-2k} \left\| Q_{<k+10} P_k \nabla_{t,x}(\chi\Omega_X \phi_{>k+10}) \right\|_{L^2_{t,x}}^2$$

$$\lesssim \left\| \sum_{k\in\mathbb{Z}} \sum_{k_1,k_2 \geq k+O(1): |k_1-k_2|\leq O(1)} 2^{-(k_2-k)} \left\| \chi\Omega_{X,k_1}(t) \right\|_{L^2_x} \left\| \nabla_x \phi_{k_2}(t) \right\|_{L^2_x} \right\|_{L^2_t}^2,$$

where we discarded via Plancherel in $L^2_{t,x}$ the Fourier multiplier $2^{-k}\nabla_{t,x}Q_{<k+10}\widetilde{P}_k$ having bounded symbol, passed from ℓ^2 to ℓ^1 summation in k after commuting time integration with the discrete sum $\sum_k$, and applied Bernstein (2.1.4) with Cauchy-Schwarz. This contribution is directly seen to be bounded by $O(\|X\phi\|_{L^2_{t,x}[-1,1]}^2 \|\nabla_{t,x}\phi\|_{L^\infty_t(L^2_x)[-1,1]}^2)$ as required.

The terms $\psi_k^{(4)}$, $\varphi_k^{(1)}$ and $\varphi_k^{(2)}$ are similar and require a duality argument relying heavily on their compensated structure to obtain estimate (2.3.19) at ℓ^2 modulation.

First for $\psi_k^{(4)}$, using the self-adjointness of $Q_{<k+10}$ and then commuting $\sum_k$ with time integration, we have:

$$\sum_{k\in\mathbb{Z}} 2^{-2k} \left\| Q_{<k+10} \psi_k^{(4)} \right\|_{L^2_{t,x}}^2$$

$$\lesssim \left\| \sum_{k\in\mathbb{Z}} 2^{-k} \left\| (Q_{<k+10}^2 \psi_k^{(4)})(t) \right\|_{L^2_x} \right\|_{L^2_t} \cdot \left\| \sup_{k\in\mathbb{Z}} 2^{-k} \left\| \psi_k^{(4)}(t) \right\|_{L^2_x} \right\|_{L^2_t[-1,1]}.$$

For the first factor, we claim that it is universally bounded due to its compensated structure. Indeed, passing to the Hardy space on each time slice via the embedding (2.1.6), we estimate it by:

$$\left\| (\sum_{k\in\mathbb{Z}} |Q_{<k+10}^2 \psi_k^{(4)}|^2)^{\frac{1}{2}} \right\|_{L^2_t(L^1_x)} \lesssim \left\| (\sum_{k\in\mathbb{Z}} |2^k P_k[\chi\Omega_{x_1}\phi_{>k+10}]|^2)^{\frac{1}{2}} \right\|_{L^2_t(L^1_x)},$$

where we relied on the Calderón-Zygmund theory for the Littlewood-Paley square function and the vector valued operator $(2^{-k}XQ_{<k+10}^2\widetilde{P}_k)_{k\in\mathbb{Z}}$, precomposing with the space-time LP-projections to $|\tau|+|\xi| \sim 2^k$ as necessary. From there, proceeding as previously, we immediately bound the latter by $O(\|\nabla_x\phi\|_{L^\infty_t(L^2_x)[-1,1]}^2)$ as required.

The set-up is similar for $\varphi_k^{(1)}$ and $\varphi_k^{(2)}$. Here however, being at high modulations, we start by inverting the time-like vector X for one of the factors. Then, using the skew-adjointness of $2^k X^{-1} Q_{\geq k+10}$, but proceeding identically to the above otherwise,

we obtain:

$$\sum_{k\in\mathbb{Z}} 2^{-2k} \int\int (Q_{\geq k+10}\varphi_k^{(i)})(\frac{X}{X}Q_{\geq k+10}\varphi_k^{(i)})dxdt$$

$$\lesssim \left\|\sum_{k\in\mathbb{Z}} 2^{-k}\left\|(\frac{2^k}{X}Q_{\geq k+10}^2\varphi_k^{(i)})(t)\right\|_{L_x^2}\right\|_{L_t^2} \cdot \left\|\sup_{k\in\mathbb{Z}} 2^{-2k}\left\|X\varphi_k^{(i)}(t)\right\|_{L_x^2}\right\|_{L_t^2[-1,1]}$$

$$\lesssim \sup_{j\geq 10}\left\|(\sum_{k\in\mathbb{Z}}|\frac{2^{k+j}}{X}Q_{k+j}\widetilde{Q}_{k+j}\varphi_k^{(i)}|^2)^{\frac{1}{2}}\right\|_{L_t^2(L_x^1)} \cdot \left\|\sup_{k\in\mathbb{Z}} 2^{-2k}\left\|X\varphi_k^{(i)}(t)\right\|_{L_x^2}\right\|_{L_t^2[-1,1]}$$

$$\lesssim \left\|(\sum_{k\in\mathbb{Z}}|\varphi_k^{(i)}|^2)^{\frac{1}{2}}\right\|_{L_t^2(L_x^1)} \cdot \left\|\sup_{k\in\mathbb{Z}} 2^{-2k}\left\|X\varphi_k^{(i)}(t)\right\|_{L_x^2}\right\|_{L_t^2[-1,1]},$$

where $\widetilde{Q}_{k+j} = Q_{k+j-1\leq\cdot\leq k+j+1}$ is the slightly enlarged modulation projection, and we relied as previously on Calderón-Zygmund theory to discard the vector valued operator $(2^{k+j}X^{-1}Q_{k+j}\widetilde{Q}_{k+j}\widetilde{P}_k)_{k\in\mathbb{Z}}$, precomposing with the space-time LP-projections to $|\tau| + |\xi| \sim 2^{k+j}$ as necessary, for any integer $j \geq 10$.

From there, we note that the first factor is bounded by $O(\|\nabla_x\phi\|_{L_t^\infty(L_x^2)[-1,1]}^2)$ as required. This follows essentially from the arguments used to treat the high-high interactions and the paraproducts, for $\varphi_k^{(1)}$ and $\varphi_k^{(2)}$ respectively, in the proof of Proposition 2.3.1 that we shall not reproduce here.

Given this, to prove estimate (2.3.19) for the terms $\psi_k^{(4)}$, $\varphi_k^{(1)}$ and $\varphi_k^{(2)}$, it is enough by (2.1.3) and (2.1.4) to establish the following couple of weak estimates:

$$2^{-k}\left\|XP_k[(\chi\phi^p\nabla_x\phi^q)(P_{\leq k+10}\nabla_x\phi^r)](t)\right\|_{L_x^1} \lesssim \|X\phi(t)\|_{L_x^2} + \|(\phi-c)(t)\|_{L_x^2}, \quad (2.3.22)$$

$$2^{-k}\left\|XP_k[\chi\phi^p(\nabla_x\phi^q)(\phi_{>k+10}^r)](t)\right\|_{L_x^2} \lesssim \|X\phi(t)\|_{L_x^2} + \|(\phi-c)(t)\|_{L_x^2}, \quad (2.3.23)$$

for any $1 \leq p,q,r \leq n$ and any time slice $t \in [-1,1]$.

Consider (2.3.22). For convenience, let us suppress the time t from the notation. Moving X inside the bracket, we first differentiate the time cut-off getting by Cauchy-Schwarz:

$$2^{-k}\left\|(\partial_t\chi)(\phi^p\nabla_x\phi^q)(P_{\leq k+10}\nabla_x\phi^r)]\right\|_{L_x^1} \lesssim \|\phi\|_{L_x^\infty}\|\nabla_x\phi\|_{L_x^2}\|\phi-c\|_{L_x^2},$$

where we relied on the finite band property (2.1.2) for ϕ^r, which is a permissible bound for (2.3.22).

Next, if X falls on ϕ^p, then we have:

$$2^{-k} \left\| \chi X \phi^p \nabla_x \phi^q (P_{\leq k+10} \nabla_x \phi^r)] \right\|_{L^1_x} \lesssim \left\| X \phi^p \right\|_{L^2_x} \left\| \chi \nabla_x \phi^q \right\|_{L^2_x} \left\| \phi^r \right\|_{L^\infty_x} ,$$

with again the finite band property (2.1.2) applied to ϕ^r, but this time in L^∞_x, and this is an acceptable bound.

When X falls on $\nabla_x \phi^q$, we shall first insert the projection $P_{\leq k+O(1)}$ in front of $\phi^p X \nabla_x \phi^q$, which is possible by the localization of $\nabla_x \phi^r_{\leq k+10}$, and untangle the high-high interactions:

$$\begin{aligned}
P_{\leq k+O(1)}(\phi^p X \nabla_x \phi^q) = P_{\leq k+O(1)} [&\phi^p_{\leq k+O(1)} X \nabla_x \phi^q_{\leq k+O(1)} \\
&+ \sum_{k_1, k_2 \geq k+O(1): |k_1 - k_2| \leq O(1)} \phi^p_{k_1} X \nabla_x \phi^q_{k_2}].
\end{aligned}$$

Given this decomposition, we have for the low frequency interactions:

$$\begin{aligned}
2^{-k} &\left\| \chi \phi^p_{\leq k+O(1)} (X \nabla_x \phi^q_{\leq k+O(1)}) (\nabla_x \phi^r_{\leq k+10}) \right\|_{L^1_x} \\
&\lesssim \left\| \phi^p \right\|_{L^\infty_x} \left\| X \phi^q \right\|_{L^2_x} \left\| \nabla_x \phi^r \right\|_{L^2_x} ,
\end{aligned}$$

where we used the finite band property (2.1.2) for ϕ^q, and this is acceptable. For the high-high frequency interactions:

$$\begin{aligned}
\sum_{k_1, k_2 \geq k+O(1): |k_1 - k_2| \leq O(1)} &2^{-k} \left\| \chi \phi^p_{k_1} (X \nabla_x \phi^q_{k_2}) (\nabla_x \phi^r_{\leq k+10}) \right\|_{L^1_x} \\
&\lesssim \left\| \phi^r \right\|_{L^\infty_x} \sum_{k_1, k_2 \geq k+O(1): |k_1 - k_2| \leq O(1)} \left\| \nabla_x \phi^p_{k_1} \right\|_{L^2_x} \left\| X \phi^q_{k_2} \right\|_{L^2_x} ,
\end{aligned}$$

where we have used the finite band property (2.1.2) for ϕ^r in L^∞_x, and transferred the spatial gradient from ϕ^q to ϕ^p by relying on (2.1.3) this time and the fact that $|k_1 - k_2| \leq O(1)$. This control is acceptable applying the discrete Cauchy-Schwarz inequality in $k_1 = k_2 + O(1)$.

The last case we need to consider, in order to finish with (2.3.22), is when X falls on ϕ^r. This follows however at once, applying (2.1.2) to the latter:

$$2^{-k} \left\| \chi \phi^p (\nabla_x \phi^q)(P_{\leq k+10} \nabla_x X \phi^r)] \right\|_{L^2_t(L^1_x)} \lesssim \left\| \phi^p \right\|_{L^\infty_{t,x}} \left\| \nabla_x \phi^q \right\|_{L^\infty_t (L^2_x)[-1,1]} \left\| X \phi^r \right\|_{L^2_{t,x}[-1,1]} ,$$

which is certainly acceptable and gives (2.3.22).

The estimate (2.3.23) is very much similar to (2.3.22). As previously, we move X into the bracket, first estimating the term when the derivative falls on the time cut-off, passing initially to L_x^1 via Bernstein's inequality (2.1.4):

$$\left\|(\partial_t\chi)(\phi^p\nabla_x\phi^q)(\phi^r_{>k+10})\right\|_{L_x^1} \lesssim \|\phi\|_{L_x^\infty}\|\nabla_x\phi\|_{L_x^2}\|\phi-c\|_{L_x^2},$$

simply noting that $P_{>k+10}\phi = P_{>k+10}(\phi-c)$ and then discarding the LP-projection. When X differentiates ϕ^p, we pass again to L_x^1, and then immediately get:

$$\left\|\chi(X\phi^p)(\nabla_x\phi^q)(\phi^r_{>k+10})\right\|_{L_x^1} \lesssim \|X\phi^p\|_{L_x^2}\|\nabla_x\phi^q\|_{L_x^2}\|\phi^r\|_{L_x^\infty}.$$

Both estimates are acceptable for (2.3.23).

We consider now the term with X falling on ϕ^q, and untangling the high-high interactions in the product we should regroup together ϕ^p and ϕ^r, obtaining:

$$P_k[(\phi^p X\nabla_x\phi^q)(\phi^r_{>k+10})] = P_k[P_{\leq k+O(1)}(\phi^p\phi^r_{>k+10})X\nabla_x\phi^q_{\leq k+O(1)}$$
$$\sum_{k_1,k_2\geq k+O(1):|k_1-k_2|\leq O(1)} P_{k_1}(\phi^p\phi^r_{>k+10})X\nabla_x\phi^q_{k_2}].$$

Now, given this decomposition, we control the first term directly by applying the finite band property (2.1.2) to ϕ^q :

$$2^{-k}\left\|\chi P_{\leq k+O(1)}(\phi^p\phi^r_{>k+10})X\nabla_x\phi^q_{\leq k+O(1)}\right\|_{L_x^2} \lesssim \left\|\phi^p\phi^r_{>k+10}\right\|_{L_x^\infty}\|X\phi^q\|_{L_x^2},$$

which is acceptable by the boundedness of wave maps. For the high-high interactions we proceed as for (2.3.22) above, passing initially to L_x^1 via Bernstein's inequality (2.1.4) and transferring the spatial gradient ∇_x from ϕ^q to $\phi^p\phi^r_{>k+10}$ via the finite band property (2.1.3), which gives:

$$\sum_{k_1,k_2\geq k+O(1):|k_1-k_2|\leq O(1)}\left\|\chi P_{k_1}(\phi^p\phi^r_{>k+10})X\nabla_x\phi^q_{k_2}\right\|_{L_x^1}$$
$$\lesssim \sum_{k_1,k_2\geq k+O(1):|k_1-k_2|\leq O(1)}\left\|P_{k_1}\nabla_x(\phi^p\phi^r_{>k+10})\right\|_{L_x^2}\left\|X\phi^q_{k_2}\right\|_{L_x^2},$$

and using the discrete Cauchy-Schwarz, we can bound this by:

$$(\|\nabla_x\phi^p\|_{L_x^2}\|\phi^r\|_{L_x^\infty} + \|\phi^p\|_{L_x^\infty}\|\nabla_x\phi^r\|_{L_x^2})\|X\phi^q\|_{L_x^2},$$

which is certainly acceptable.

Lastly, if X differentiates ϕ^r, we pass to L_x^1 and this immediately yields the desired control:

$$\|\chi\phi^p(\nabla_x\phi^q)(P_{>k+10}X\phi^r)\|_{L_x^1} \lesssim \|\phi^p\|_{L_x^\infty} \|\nabla_x\phi^q\|_{L_x^2} \|X\phi^r\|_{L_x^2},$$

hence we have (2.3.23).

Lemma 2.3.2 is proved. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Chapter 3

Proof of the Main Theorems: Bubbling Analysis

We prove our main Theorem 1.2.2 in Section 3.1, Theorem 1.2.3 in Section 3.2 and Theorem 1.2.4 in Section 3.3. We start by recording, in the lemma just below, some of the important properties of the wave map ϕ, we were considering in the statement of the Threshold Theorem 1.2.1, at the final rescaling obtained by Sterbenz and Tataru in Section 6.6 of [30].

Lemma 3.0.1. (Sterbenz and Tataru [30]). *The maps $\{\phi_\nu\}_{\nu \in \mathbb{N}}$ from Theorem 1.2.1 represent a sequence of smooth wave maps of bounded energy on increasingly large domains of the forward light cone C:*

$$\phi_\nu : C_{[\varsigma\nu, \varsigma_\nu^{-1}]} \longrightarrow \mathcal{N}, \;\; \mathcal{E}_{S_t}[\phi_\nu] \leq \mathcal{E} \;\; \forall t \in [\varsigma_\nu, \varsigma_\nu^{-1}], \tag{3.0.1}$$

where $\varsigma_\nu \downarrow 0$ as $\nu \to \infty$, with the following properties:
*   • There exists a sequence $\epsilon_\nu \downarrow 0$, with $\epsilon_\nu^{\frac{1}{2}} \ll \varsigma_\nu$, such that:*

$$\mathcal{F}_{[\varsigma\nu, \varsigma_\nu^{-1}]}[\phi_\nu] < \epsilon_\nu^{\frac{1}{2}} \mathcal{E}; \tag{3.0.2}$$

*   • A decay to the self-similar mode holds:*

$$\int\int_{C_{[\varsigma\nu, \varsigma_\nu^{-1}]}^{\epsilon_\nu^{\frac{1}{2}}}} \frac{1}{\rho} |\partial_\rho \phi_\nu|^2 \, dxdt \lesssim |\log \epsilon_\nu|^{-\frac{1}{2}} \mathcal{E}, \tag{3.0.3}$$

where $\rho = (t^2 - r^2)^{\frac{1}{2}}$ and $\partial_\rho = \frac{1}{\rho}(t\partial_t + r\partial_r)$ is the scaling vector field which we recall is uniformly time-like $\mu(\partial_\rho, \partial_\rho) = -1$;

49

- *There is a uniform amount of energy $\mathcal{E}_c > 0$ getting concentrated by the maps ϕ_ν in the interior of the light cone:*

$$\frac{1}{2} \int_{|x| < \gamma_c t_0} |\nabla_{t,x} \phi_\nu(t_0)|^2 \, dx \geq \mathcal{E}_c \quad \forall t_0 \in [\varsigma_\nu, \varsigma_\nu^{-1}], \tag{3.0.4}$$

for some $0 < \gamma_c < 1$.

Let us write here a few lines of comments regarding the above lemma, referring the reader to [30] for more details. Given a sequence of concentration points (t_ν, x_ν) for the energy dispersion norm:

$$2^{-k_\nu} |P_{k_\nu} \nabla_{t,x} \phi(t_\nu, x_\nu)| > \epsilon(\mathcal{N}, \mathcal{E}),$$

with $t_\nu \to 0$ in the case of a finite time blow-up, or $t_\nu \to +\infty$ in a non-scattering scenario, the sequence $\epsilon_\nu \downarrow 0$ is chosen such that:

$$\mathcal{F}_{[\epsilon_\nu t_\nu, t_\nu]}[\phi] < \epsilon_\nu^{\frac{1}{2}} \mathcal{E}.$$

In [30], Sections 6.3 and 6.4, the authors use the above lower bound to prove that there is a non-trivial amount of time-like energy concentrating on the time slice S_{t_ν}. As we shall later rely on those results in Section 3.2, we gathered them in Lemma 3.2.1 here. From there, a weighted energy estimate (see Lemma 3.4 in [30]) propagates this energy backwards in time, leading to (3.0.4) for any $t \in [\epsilon_\nu^{1/2} t_\nu, \epsilon_\nu^{1/4} t_\nu]$.

In parallel to this, a Morawetz type estimate (see Lemma 3.3 in [30]) and the pigeonhole principle enable Sterbenz and Tataru to find a sequence of time intervals $[\tau_\nu, N_\nu \tau_\nu] \subset [\epsilon_\nu^{1/2}, \epsilon_\nu^{1/4}]$, with $N_\nu = \exp(\sqrt{|\log \epsilon_\nu|})$, such that the following decay estimate holds:

$$\int \int_{C_{[\tau_\nu, N_\nu \tau_\nu]}^{\epsilon_\nu}} \frac{1}{\rho} |\partial_\rho[\phi(t_\nu t, t_\nu x)]|^2 \, dxdt \lesssim |\log \epsilon_\nu|^{-\frac{1}{2}} \mathcal{E},$$

see section 6.6 in [30]. Then for the final rescaling, the authors in [30] choose $t_\nu \tau_\nu$ for the scales λ_ν^0 (or λ_ν^∞), obtaining a sequence of wave maps $\phi(\lambda_\nu^0 \cdot)$ with the desired properties on the growing cones $C_{[1, N_\nu]}$. In our case, it will be more convenient (for notational purposes mainly, as to respect the CMC foliation in Section 3.1 below), to asymptotically cover all of the forward light cone C^0, so we should simply fix any:

$$t_\nu \tau_\nu \ll \lambda_\nu^0, \lambda_\nu^\infty \ll N_\nu t_\nu \tau_\nu,$$

and choose then $\varsigma_\nu \downarrow 0$ decaying slowly enough, for Lemma 3.0.1 to hold.

Finally, we bring the reader's attention here to our convention that, in any of the results stated in this last section, we assume (3.0.1)-(3.0.4) holding without mentioning it. In fact, one might directly consider those as the assumptions under which claims of Theorem 1.2.2 are made.

3.1 Blow-up analysis for asymptotically self-similar sequences of wave maps

We start the proof of Theorem 1.2.2 with a study of the energy concentration sets. Our approach here will be close in spirit to the work of Freire, Müller and Struwe [10]. We will rely on a monotonicity lemma for asymptotically self-similar wave maps, see Lemma 3.1.1 below, which is a rough analogue of part (ii) from Lemma 1.7 in Lin's work [20], but mainly parallels the computations in the proof of Morawetz type estimates from section 3 of [30]. Note that we do not use here the fact that our target manifold is a sphere.

3.1.1 An asymptotic monotonicity formula

It will be convenient to use hyperbolic coordinates, also known as CMC foliation of the (forward) light cone C^0, where we recall that C^0 denotes the open interior of the forward light cone, $C^0 = C \setminus (\partial C \cup \{(0,0)\})$. Those are defined by:

$$t = \rho \cosh(y), \ \ r = \rho \sinh(y) \ \text{ and } \ \theta.$$

Associated to those coordinates, we recall the expression for the volume element:

$$dV := rdtdrd\theta = \rho^2 \sinh(y) d\rho dy d\theta,$$

and for the hyperbolic planes $\mathbb{H}^2_{\rho_0} = \{\rho = \rho_0\}$ the area element:

$$dA_{\rho_0} := \rho_0^2 \sinh(y) dy d\theta,$$

with respect to the Minkowski metric μ on $\mathbb{R}^{2+1}$. These formulae will be useful below in applying Stokes' theorem in the hyperbolic annulus $\{\rho_1 \leq \rho \leq \rho_2\}$. Let us also record

here that, using the identities:

$$\partial_t = \frac{t}{\rho}\partial_\rho - \frac{r}{\rho^2}\partial_y, \ \ \partial_r = \frac{t}{\rho^2}\partial_y - \frac{r}{\rho}\partial_\rho,$$

one computes, for a smooth map ϕ into $\mathbb{S}^{n-1}$:

$$\partial^\gamma \phi^\dagger \partial_\gamma \phi = -|\partial_t\phi|^2 + |\partial_r\phi|^2 + \frac{1}{r^2}|\partial_\theta\phi|^2$$

$$= -|\partial_\rho\phi|^2 + \frac{1}{\rho^2}|\nabla_{\mathbb{H}^2}\phi|^2, \tag{3.1.1}$$

where $\nabla_{\mathbb{H}^2}$ denotes the gradient on the unit hyperboloid $\mathbb{H}^2 := \mathbb{H}_1^2$:

$$|\nabla_{\mathbb{H}^2}\phi|^2 = |\partial_y\phi|^2 + \frac{1}{\sinh^2(y)}|\partial_\theta\phi|^2.$$

For every given $\rho_0 > 0$, let us define the Radon measures:

$$\sigma_{\nu,\rho_0} := \left(|\partial_\rho\phi_\nu|^2 + \frac{1}{\rho^2}|\nabla_{\mathbb{H}^2}\phi_\nu|^2\right)dA_{\rho_0} \in \mathcal{R}(\mathbb{H}_{\rho_0}^2).$$

We can naturally view them as measures on the unit hyperbolic plane $\mathbb{H}^2$ since for any given test function φ on $\mathbb{H}^2$, that we should view as a function $\varphi(y,\theta)$ independent of ρ on the whole of the light cone C^0, we have:

$$\int \varphi d\sigma_{\nu,\rho_0} = \int_{\mathbb{H}^2}\left(|\partial_\rho\phi_\nu(\rho_0)|^2 + \frac{1}{\rho_0^2}|\nabla_{\mathbb{H}^2}\phi_\nu(\rho_0)|^2\right)\varphi(y,\theta)\rho_0^2\sinh(y)dyd\theta.$$

Using the decay (3.0.3) to a self-similar mode, we can establish the following asymptotic monotonicity property for the family $\{\sigma_{\nu,\rho}\}_{\rho>0} \subset \mathcal{R}(\mathbb{H}^2)$.

Lemma 3.1.1. *For every pair $\rho_2 > \rho_1 > 0$ and every $\lambda > 0$, we have the decay:*

$$\int_{\rho_1}^{\rho_2}\left(\int \varphi d\sigma_{\nu,\rho_0}\right)d\rho_0 - \int_{\rho_1+\lambda}^{\rho_2+\lambda}\left(\int \varphi d\sigma_{\nu,\rho_0}\right)d\rho_0 \longrightarrow 0, \tag{3.1.2}$$

holding as $\nu \to +\infty$ for any test function $\varphi \in C_0^\infty(\mathbb{H}^2)$.

Proof. Given a continuously differentiable vector field $\psi = \psi^\beta\partial_\beta$ compactly supported in (y,θ), contracting the stress-energy tensor $T[\phi_\nu]$ with ψ, we obtain the associated Noether current:

$$^{(\psi)}P_\alpha = T_{\alpha\beta}[\phi_\nu]\psi^\beta.$$

Hence, if we set:

$$D_{\{\rho' \leq \rho \leq \rho''\}}(\psi) := \int_{\{\rho' \leq \rho \leq \rho''\}} \partial^\alpha \left({}^{(\psi)}P_\alpha \right) dV = \int_{\{\rho' \leq \rho \leq \rho''\}} T_{\alpha\beta}[\phi_\nu] \partial^\alpha \psi^\beta dV,$$

where we relied on the conservation law (1.1.12), and:

$$B_{\tilde{\rho}}(\psi) := \int_{\{\rho = \tilde{\rho}\}} {}^{(\psi)}P(\partial_\rho) dA_{\tilde{\rho}} = \int_{\{\rho = \tilde{\rho}\}} T_{\alpha\beta}[\phi_\nu] \frac{x^\alpha}{\tilde{\rho}} \psi^\beta dA_{\tilde{\rho}},$$

where our convention follows $x^0 := t$ and $x_0 = -t$, so that $x^\alpha = \mu^{\alpha\gamma} x_\gamma$, applying Stokes' theorem over the region $\{\rho_0 \leq \rho \leq \rho_0 + \lambda\}$ leads to the identity:

$$D_{\{\rho_0 \leq \rho \leq \rho_0 + \lambda\}}(\psi) = B_{\rho_0}(\psi) - B_{\rho_0 + \lambda}(\psi). \tag{3.1.3}$$

Taking $\psi = \varphi(y, \theta)\partial_\rho$, we compute using the expression (1.1.11) for $T_{\alpha\beta}[\phi_\nu]$:

$$D_{\{\rho_0 \leq \rho \leq \rho_0 + \lambda\}}(\psi) = \int_{\{\rho_0 \leq \rho \leq \rho_0 + \lambda\}} \left(\frac{1}{\rho} |\partial_\rho \phi_\nu|^2 \varphi + \partial_\rho \phi_\nu^\dagger \partial_\alpha \phi_\nu \partial^\alpha \varphi \right) dV,$$

and for the boundary terms:

$$\begin{aligned} B_{\tilde{\rho}}(\psi) &= \int_{\{\rho = \tilde{\rho}\}} \left(|\tilde{\rho} \partial_\rho \phi_\nu|^2 + \tilde{\rho}^2 \frac{1}{2} \partial^\gamma \phi_\nu^\dagger \partial_\gamma \phi_\nu \right) \frac{\varphi}{\tilde{\rho}^2} dA_{\tilde{\rho}} \\ &= \frac{1}{2} \int \varphi d\sigma_{\nu, \tilde{\rho}}, \end{aligned}$$

where to pass to the second line we have used the identity (3.1.1). Therefore, plugging the above back into (3.1.3) we obtain:

$$\int \varphi d\sigma_{\nu, \rho_0} - \int \varphi d\sigma_{\nu, \rho_0 + \lambda} = 2 \int_{\{\rho_0 \leq \rho \leq \rho_0 + \lambda\}} \left(\frac{1}{\rho} |\partial_\rho \phi_\nu|^2 \varphi + \partial_\rho \phi_\nu^\dagger \partial_\alpha \phi_\nu \partial^\alpha \varphi \right) dV.$$

Integrating over $\rho_0 \in [\rho_1, \rho_2]$ and using Cauchy-Schwarz for the second term on RHS above, appealing to the decay (3.0.3) and the global energy bound (3.0.1), we obtain (3.1.2). Hence Lemma 3.1.1 is proved. $\qquad \square$

3.1.2 The concentration-compactness argument

From now on we restrict ourselves to the time interval $1 \leq t \leq 2$. We will study there the sets in space-time where our wave maps concentrate a non-trivial amount

of energy as in the work of Freire, Müller and Struwe [10], where some general state-ments about the structure of energy concentration loci can be found (for instance, it is shown in Proposition 4.1 and Theorem B.1 of [10] that, upon passing to a suitable subsequence, the concentration set of an energy threshold will be contained in a finite union of Lipschitz curves). Our assumptions however enable us to go beyond [10] via more elementary arguments and prove that picking a suitable subsequence will lead to an energy concentration set which is in fact given by a finite collection of time-like geodesics, relying on Lemmata 2.2.2 and 3.1.1.

To use the latter, we remark that for a fixed open domain U with closure $\overline{U} \subset C^0_{[\frac{1}{2},3]}$, we have:

$$\frac{1}{C} |\nabla_{t,x}\phi_\nu|^2 \leq |\partial_\rho \phi_\nu|^2 + \frac{1}{\rho^2} |\nabla_{\mathbb{H}^2}\phi_\nu|^2 \leq C |\nabla_{t,x}\phi_\nu|^2 \quad \text{on} \ \ U, \tag{3.1.4}$$

with $C := C(\text{dist}(U, \partial C_{[\frac{1}{2},3]}))$, and this will enable us to transfer control back and forward between the Radon measures $\sigma_{\nu,\rho}$ and the energy densities $|\nabla_{t,x}\phi_\nu|^2 \, dx dt$ of which we want to study the concentration sets (with the small energy compactness Lemma 2.2.2 enabling us to obtain some uniformity in time).

Lemma 3.1.2. *There exists a subsequence of $\{\phi_\nu\}_{\nu \in \mathbb{N}}$ restricting to which, without changing notation, we can find a finite collection of time-like geodesics $\varrho_1, \ldots, \varrho_I$ passing through the origin in Minkowski space such that defining the energy concentration set by:*

$$\Sigma := \left\{ (t,x) \in C^0_{[1,2]} \ : \ \liminf_{\nu \to \infty} \mathcal{E}_{B_r(x)}[\phi_\nu](t) > \epsilon_s \ \forall r > 0 \right\},$$

we have:

$$\Sigma = C^0_{[1,2]} \cap \bigcup_{i=1}^{I} \varrho_i,$$

and away from Σ, there exist a wave map ϕ satisfying:

$$\partial_\rho \phi = 0 \ \ on \ \ C^0_{[1,2]} \setminus \Sigma \ \ with \ \ \phi \in (H_{t,x}^{\frac{3}{2}-\epsilon})_{loc}\left(C^0_{[1,2]} \setminus \Sigma\right),$$

for any $0 < \epsilon < \frac{1}{2}$, of finite energy on $C^0_{[1,2]}$, $\mathcal{E}_{S^0_t}[\phi] \leq \mathcal{E} \ \forall t \in [1,2]$, such that:

$$\phi_\nu \longrightarrow \phi \ \ on \ \ \left(C^0_t(H_x^1) \cap C^1_t(L_x^2)\right)_{loc}\left(C^0_{[1,2]} \setminus \Sigma\right), \tag{3.1.5}$$

as dictated by Lemma 2.2.2.

Proof. In view of the asymptotic monotonicity provided by Lemma 3.1.1, let us denote

for a set $U \subset S_{t=1}$ the cone over U by:

$$C(U) := \{\lambda(t,x) \,:\, \lambda > 0, \, x \in U \text{ at } t = 1\},$$

and by $C_I(U) := C(U) \cap C_I$ the corresponding truncation to a time interval I.

Considering the time slice S_1^0, given the global energy bound (3.0.1) we can pass to a subsequence for $\{\phi_\nu\}_{\nu \in \mathbb{N}}$, without changing notation, such that for some Radon measure $\iota \in \mathcal{R}(S_1^0)$ we have:

$$|\nabla_{t,x}\phi_\nu(1)|^2 \, dx \rightharpoonup \iota \text{ in } \mathcal{R}(S_1^0), \tag{3.1.6}$$

from where we also see that there exist only finitely many points $\{x_i\}_{i=1}^I \subset S_1^0$ such that:

$$\{x_i\}_{i=1}^I = \left\{x \in B_1 \,:\, \lim_{\nu \to \infty} \mathcal{E}_{B_r(x)}[\phi_\nu](1) > \epsilon_s \; \forall r > 0\right\}, \tag{3.1.7}$$

and we set $\varrho_i := C(\{x_i\})$.

Let us start by showing that:

$$\Sigma \subset C_{[1,2]}^0 \cap \bigcup_{i=1}^I \varrho_i, \tag{3.1.8}$$

obtaining on the way claim (3.1.5). Fix any point $x_0 \in S_1^0 \setminus \Sigma$, then there exists a radius $r_1 = r_1(x_0) > 0$ such that for all $\nu \in \mathbb{N}$:

$$\mathcal{E}_{B_{r_1}(x_0)}[\phi_\nu] \le \epsilon_s,$$

hence by the energy-flux identity (1.2.2), shrinking r_1 to $r_2 > 0$ as necessary, we obtain that:

$$\sup_{t \in [1-3r_2, 1+3r_2]} \mathcal{E}_{B_{3r_2}(x_0)}[\phi_\nu](t) \le \epsilon_s.$$

By the decay assumption (3.0.3), we can apply the compactness Lemma 2.2.2 obtaining that on a subsequence $\{\phi_{\nu'}\}_{\nu' \in \mathbb{N}}$ we have convergence in $C_t^0(H_x^1) \cap C_t^1(L_x^2)$ to a wave map ϕ in $[1 - r_2, 1 + r_2] \times B_{r_2}(x_0)$, satisfying $\partial_\rho \phi = 0$ and having regularity as dictated by (2.2.9) there.

Hence, given any positive constant $\eta > 0$ there exist a radius $r_\eta > 0$ such that:

$$\sup_{\nu' \in \mathbb{N}} \sup_{t \in [1-r_\eta, 1+r_\eta]} \mathcal{E}_{C(B_{r_\eta}(x_0))}[\phi_{\nu'}](t) \le \eta.$$

Therefore, using (3.1.4) we get for any test function $\varphi(y, \theta)$ on the hyperboloid $\mathbb{H}^2$, having support in $C(B_{r_\eta}(x_0)) \cap \mathbb{H}^2$ and satisfying $0 \leq \varphi \leq 1$, the bound:

$$\sup_{\nu' \in \mathbb{N}} \frac{1}{\rho_2 - \rho_1} \int_{\rho_1}^{\rho_2} \left(\int \varphi d\sigma_{\nu', \rho} \right) d\rho \lesssim \eta,$$

for some suitably chosen $0 < \rho_1 < \rho_2$. The implicit constant here does not depend on the parameter η, and in fact depends only on the distance of the point $(1, x_0)$ to the null boundary.

Recalling Lemma 3.1.1, we obtain by (3.1.2) for every fixed $\lambda > 0$ the estimate:

$$\limsup_{\nu' \to \infty} \frac{1}{\rho_2 - \rho_1} \int_{\rho_1 + \lambda}^{\rho_2 + \lambda} \left(\int \varphi d\sigma_{\nu', \rho} \right) d\rho \lesssim \eta.$$

Given this, shrinking r_2 to $r_3 = r_3(x_0, \eta) > 0$ and picking a suitable cut-off function φ on $\mathbb{H}^2$ as necessary, we can rely on the other inequality in (3.1.4) this time and the energy-flux identity (1.2.2) to find, arguing via the pigeonhole principle, a finite cover of:

$$C_{[1,2]}(B_{r_3}(x_0)) \subset \bigcup_{j=1}^{N} [t_j - s_j, t_j + s_j] \times B_{s_j}(y_j)$$

with $N = N(x_0, \eta) \in \mathbb{N}$ satisfying:

$$\bigcup_{j=1}^{N} [t_j - 3s_j, t_j + 3s_j] \times B_{3s_j}(y_j) \subset C_{[\frac{1}{2}, 3]}(B_{r_3}(x_0)),$$

and such that:

$$\limsup_{\nu' \to \infty} \sup_{t \in [t_j - 3s_j, t_j + 3s_j]} \mathcal{E}_{B_{3s_j}(y_j)}[\phi_{\nu'}](t) \lesssim \eta, \quad j = 1, \ldots, N,$$

where the implicit constant is independent of η. Hence, choosing $\eta > 0$ small enough we can claim:

$$\limsup_{\nu' \to \infty} \sup_{t \in [t_j - 3s_j, t_j + 3s_j]} \mathcal{E}_{B_{3s_j}(y_j)}[\phi_{\nu'}](t) \leq \frac{1}{2} \epsilon_s, \quad j = 1, \ldots, N,$$

with $N = N(x_0)$ and $r_3 = r_3(x_0)$ now.

Proceeding this way for a countable dense set of points $x_0 \in S_1^0 \setminus \Sigma$, we obtain ultimately a countable cover of $C_{[1,2]}^0 \setminus \cup_i \varrho_i$ that we can use together with the compactness

Lemma 2.2.2 to construct a subsequence for $\{\phi_\nu\}_{\nu\in\mathbb{N}}$ via the diagonal process, to which we restrict ourselves without changing notation this time, such that (3.1.5) hold for a wave map $\phi \in (H^{3/2-\epsilon}_{t,x})_{loc}(C^0_{[1,2]} \setminus \cup_i \varrho_i)$ with $\partial_\rho\phi = 0$. By construction, it can be seen immediately that the obtained map ϕ has energy bounded by $\mathcal{E}$ and we note the argument also yields (3.1.8) as desired.

To finish the proof of the lemma, we need to get the reverse inclusion to (3.1.8). This follows however from a simple argument by contradiction: suppose that there exists a point $(s_i, y_i) \in \varrho_i$ which is not contained in Σ. We can then run the above proof with (s_i, y_i) instead of $(1, x_0)$ and obtain that the full ray ϱ_i is not contained in Σ, but that contradicts the definition of x_i from (3.1.7). Lemma 3.1.2 is therefore proved. $\qquad\square$

3.1.3 Ruling out self-similar behavior

To close the proof of the first part of Theorem 1.2.2 it is enough now to prove that the wave map ϕ obtained above must in fact be constant. For this point, we will rely on a folklore fact that finite energy self-similar wave maps do not exist in dimension $2 + 1$ which we state in Proposition 3.1.3 below. A self-contained proof of this proposition can be found in the work of Sterbenz and Tataru [30] (see section 4 there).

Proposition 3.1.3. *Let ϕ be a smooth wave map in the interior of the forward light cone C^0, having finite energy, $\mathcal{E}_{S^0_t}[\phi] \lesssim 1 \ \forall t > 0$, and satisfying the self-similarity condition $\partial_\rho\phi = 0$. Then ϕ must be constant.*

Consider the wave map ϕ from Lemma 3.1.2. By homogeneity, we can extend it to:

$$\phi : C^0 \setminus \bigcup_{i=1}^{I} \varrho_i \longrightarrow \mathcal{N},$$

with finite energy $\mathcal{E}_{S^0_t}[\phi] \leq \mathcal{E} \ \forall t > 0$, locally in $H^{\frac{3}{2}-\epsilon}_{t,x}$ and satisfying $\partial_\rho\phi = 0$. Let us note here that we were considering the unit time interval $[1, 2]$ in (3.1.5) just in order to simplify the task of keeping track of the dependence of implicit constants. It is easy to see that the arguments above lead to local convergence of the sequence ϕ_ν to the map ϕ on all of $C^0 \setminus \cup_i \varrho_i$. This is however a purely qualitative statement.

Restricting ϕ to the unit hyperbolic plane $\mathbb{H}^2$ gives rise to a harmonic map of locally finite energy, by (3.1.4), defined away from a finite set of points given by $\mathbb{H}^2 \cap \bigcup_{i=1}^{I} \varrho_i$. By the regularity theory due to Hélein [13], we obtain in fact a smooth harmonic map away from the above collection of points. But then, by the removable singularity theorem of

Sacks and Uhlenbeck [25] we can extend ϕ to a smooth harmonic map on the whole of the hyperbolic plane $\mathbb{H}^2$, which in turn means that, by homogeneity again, we could have extended ϕ across the rays ϱ_i to a smooth finite energy self-similar wave map on C^0. By Proposition 3.1.3, ϕ has to be a constant.

Theorem 1.2.2 is therefore established, given that Σ must be non-trivial by the concentration of time-like energy assumption (3.0.4).

3.2 Dispersive property for null-concentration

This section is devoted to the description of the parts of the sequence that escape into the null boundary. We will prove Theorem 1.2.3.

3.2.1 Cutting the bubbles off the body of the map

We proceed first, borrowing arguments from Section 6.1 of [30], by constructing extensions for the maps ϕ_ν outside the light cone with asymptotically vanishing energy there (we note that, if considering the non-scattering problem, those have been already constructed in Section 6.2 of [30]).

Relying on the flux decay estimate (3.0.2) and using the angular part of $\mathcal{F}_{[\varsigma_\nu,\varsigma_\nu^{-1}]}[\phi_\nu]$, see the expression in (1.2.2), we can find by the pigeonhole principle a sequence $\tau_\nu \in [2,3]$ such that:

$$\int_{\partial S_{\tau_\nu}} \left| r^{-1} \partial_\theta \phi_\nu(\tau_\nu) \right|^2 d\theta \lesssim \epsilon_\nu^{\frac{1}{2}}.$$

Hence, as in Remark 2.2.3, we get that $\phi_\nu(\partial S_{\tau_\nu})$ is contained in a chart of radius $O(\epsilon_\nu^{1/4})$ and so we can build smooth spatial extensions $\phi_\nu'[\tau_\nu] \in T\mathcal{N}$ of $\phi_\nu[\tau_\nu]$, satisfying the energy control:

$$\mathcal{E}[\phi_\nu'](\tau_\nu) - \mathcal{E}_{S_{\tau_\nu}}[\phi_\nu] \lesssim \epsilon_\nu^{\frac{1}{2}}.$$

We solve then the wave maps equation with initial data $\phi_\nu'[\tau_\nu]$ backwards in time for $t \in [\varsigma_\nu, \tau_\nu]$. By the finite speed of propagation property, the solution agrees with ϕ_ν on $C_{[\varsigma_\nu,\tau_\nu]}$, hence let us denote it by ϕ_ν (abusing notation slightly). Moreover, relying again on the assumption (3.0.2) and using the conservation of energy law (1.1.8) together with the energy-flux identity (1.2.2), we propagate to all of the time interval $[\varsigma_\nu, \tau_\nu]$ the smallness of the energy exterior to the light cone:

$$\sup_{t \in [\varsigma_\nu,\tau_\nu]} \left(\mathcal{E}[\phi_\nu](t) - \mathcal{E}_{S_t}[\phi_\nu] \right) \lesssim \epsilon_\nu^{\frac{1}{2}},$$

which in particular guarantees smoothness of the extension on all of $[\varsigma_\nu, \tau_\nu] \times \mathbb{R}^2$.

Another consequence of the flux decay estimate (3.0.2) that we record here, is the following weighted control:

$$\sup_{t \in [1,2]} \int_{S_t} \frac{1}{(t - |x| + \epsilon_\nu)^{\frac{1}{2}}} \left(|L\phi_\nu(t)|^2 + \left| r^{-1} \partial_\theta \phi_\nu(t) \right|^2 \right) dx \lesssim 1, \qquad (3.2.1)$$

a direct consequence of Lemma 3.2 in [30], and constitutes an important ingredient in the elimination of sharp pockets of null energy (see section 6.3 of [30]).

Regarding the interior of the cone, by the previous section we can pick a monotonically decreasing sequence of scales $\delta_\nu \downarrow 0$, starting with $\delta_0 := \frac{1}{10} \text{dist}(\cup_i \varrho_i, \partial C_{[1,2]})$, such that:

$$\lim_{\nu \to \infty} \sup_{t_0 \in [1,2]} \mathcal{E}_{S_{t_0}^{\delta_\nu} \setminus \cup_i B_{\delta_\nu}(\varrho_i(t_0))}[\phi_\nu] = 0, \qquad (3.2.2)$$

which are in some sense the slowest concentration scales, i.e. have the property that:

$$\phi_\nu(t_0 + \delta_\nu t, \varrho_i(t_0) + \delta_\nu x) \longrightarrow c_\phi \in \mathcal{N} \quad \text{on} \quad ([-4, 4] \times B_4) \setminus \varrho_i, \qquad (3.2.3)$$

locally in $C_t^0(H_x^1) \cap C_t^1(L_x^2)$, where the constant c_ϕ corresponds to the wave map ϕ from (3.1.5), for any given $t_0 \in (1, 2)$ and $i = 1, \ldots, I$. This can be obtained upon taking δ_ν tending slower to 0, which will not break condition (3.2.2). Hence, by pigeonholing, we can choose a sequence of radii $\sigma_\nu = \sigma_\nu(t_0, i) \in (3, 4)$ such that:

$$\int_{\partial B_{\sigma_\nu}} |\nabla_{t,x} [\phi_\nu(t_0 + \delta_\nu t, \varrho_i(t_0) + \delta_\nu x)]|^2 d\theta \longrightarrow 0,$$

which enables us, as before, to construct extensions into B_{σ_ν} that have asymptotically vanishing energy. That is we cut off the bubbles from the body of the map. More precisely, we choose a sequence of maps $(\varpi_{i,t_0,\nu}, \partial_t \varpi_{i,t_0,\nu}) \in T\mathcal{N}$ defined on B_{σ_ν} such that:

$$\nabla_{t,x} \left[\phi_\nu(t_0 + \delta_\nu \cdot, \varrho_i(t_0) + \delta_\nu \cdot)|_{\{t=0\} \times B_4 \setminus B_{\sigma_\nu}} + \varpi_{i,t_0,\nu}(\cdot) \right] \longrightarrow 0 \quad \text{in} \quad L_x^2(B_4),$$

and performing this surgery for each $i = 1, \ldots, I$, we obtain smooth maps:

$$\varpi_{t_0,\nu}[t_0] := \phi_\nu[t_0]|_{\mathbb{R}_x^2 \setminus \cup_i B_{\delta_\nu \sigma_\nu}(\varrho_i(t_0))} + \sum_{i=1}^{I} (\varpi_{i,t_0,\nu}, \frac{1}{\delta_\nu} \partial_t \varpi_{i,t_0,\nu}) \left(\frac{x - \varrho_i(t_0)}{\delta_\nu} \right),$$

satisfying by construction:

$$\nabla_{t,x}\varpi_{t_0,\nu}(t_0) \longrightarrow 0 \ \text{in} \ (L_x^2)_{loc}\left(\mathbb{R}^2 \setminus \{r = t_0\}\right). \tag{3.2.4}$$

3.2.2 Decay for the energy dispersion norm

Fixing $t_0 \in [1 + \delta_0, 2 - \delta_0]$, we can naturally view $\varpi_{t_0,\nu}[t_0]$ as defined on the time slice S_{t_0}, and solve the wave maps equation with initial data $\varpi_{t_0,\nu}[t_0]$ obtaining a smooth solution on $[t_0 - \delta_0, t_0 + \delta_0]$ provided we work with ν large enough, relying on the finite speed of propagation property (which tells us that $\varpi_{t_0,\nu}$ agrees with ϕ_ν near and beyond the null boundary, at least away from $C_{[t_0-\delta_0,t_0+\delta_0]}^{2\delta_0}$), and the small energy regularity via (3.2.4). The choice of δ_0 is not the most optimal one, but here we are rather concerned with its independence from ν. It is immediate then that,

$$\nabla_{t,x}\varpi_{t_0,\nu} \longrightarrow 0 \ \text{in} \ C_t^0(L_x^2)_{loc}\left(([t_0 - \delta_0, t_0 + \delta_0] \times \mathbb{R}^2) \setminus \partial C_{[t_0-\delta_0,t_0+\delta_0]}\right),$$

as desired in Theorem 1.2.3. Furthermore, the weighted estimate (3.2.1) is inherited by the maps $\varpi_{t_0,\nu}$:

$$\sup_{t\in[t_0-\tau,t_0+\tau]} \int_{S_t} \frac{1}{(t - |x| + \epsilon_\nu)^{\frac{1}{2}}} \left(|L\varpi_{t_0,\nu}(t)|^2 + \left|r^{-1}\partial_\theta\varpi_{t_0,\nu}(t)\right|^2\right) dx \lesssim 1. \tag{3.2.5}$$

This immediately implies the desired decay for the angular and the L null energy, as required in Theorem 1.2.3.

This whole construction opens us the possibility to apply the following lemma of Sterbenz and Tataru from [30] (see Sections 6.3 and 6.4 there), which forces the energy dispersion norm of $\varpi_{t_0,\nu}$ to asymptotically vanish:

$$\limsup_{\nu\to\infty} \sup_k \left(2^{-k} \|P_k \nabla_{t,x}\varpi_{t_0,\nu}\|_{L_{t,x}^\infty[t_0-\delta_0,t_0+\delta_0]}\right) = 0.$$

Lemma 3.2.1. (Sterbenz and Tataru [30]). *Consider tuples $\{(\varphi_\nu, \partial_t\varphi_\nu)\}_{\nu\in\mathbb{N}}$ of Schwartz functions on $\mathbb{R}^2$ satisfying, for some sequence $\epsilon_\nu \downarrow 0$ and a bound $\mathcal{E} > 0$:*

$$\|\nabla_{t,x}\varphi_\nu\|_{L_x^2}^2 \lesssim \mathcal{E}, \ \ \|\nabla_{t,x}\varphi_\nu\|_{L_x^2(\mathbb{R}^2\setminus B_1)}^2 \lesssim \epsilon_\nu^{\frac{1}{2}}\mathcal{E},$$

$$\int_{B_1} \frac{1}{(1 - |x| + \epsilon_\nu)^{\frac{1}{2}}} \left(|L\varphi_\nu|^2 + \left|r^{-1}\partial_\theta\varphi_\nu\right|^2\right) dx \lesssim \mathcal{E},$$

such that for some given $\epsilon > 0$:

$$\sup_k \left(2^{-k} \left\| P_k \nabla_{t,x} \varphi_\nu \right\|_{L_x^\infty} \right) > \epsilon.$$

Then, there exist constants $0 < \gamma(\epsilon, \mathcal{E}) < 1$ and $\varepsilon(\epsilon, \mathcal{E}) > 0$ for which:

$$\int_{B_{\gamma(\epsilon, \mathcal{E})}} |\nabla_{t,x} \varphi_\nu|^2 \, dx \geq \varepsilon(\epsilon, \mathcal{E}), \quad \forall \nu \in \mathbb{N}.$$

As immediate consequence of the decay of the energy dispersion norm and Theorem 2.2.4 of Sterbenz-Tataru from Section 2.2, we obtain the uniform bound:

$$\left\| \varpi_{t_0, \nu} \right\|_{S[t_0 - \delta_0, t_0 + \delta_0]} \lesssim 1,$$

claimed in Theorem 1.2.3.

3.2.3 Extracting the gauge co-variant linear wave

Because, by construction, the maps $\varpi_{t_0, \nu}$ converge to a constant strongly in $L_t^\infty(L_x^2)$, their low frequencies $P_{\leq 0} \varpi_{t_0, \nu}$ are not interesting. The purpose of this section is to extract some structure, as $\nu \to +\infty$, for the high frequencies of $\varpi_{t_0, \nu}$. For, without losing generality, we will consider abstractly a sequence of wave maps:

$$\varpi_\nu : [-1, 1] \times \mathbb{R}^2 \longrightarrow \mathcal{N} \quad \text{with} \quad \varepsilon_\nu := \sup_k \| P_k \varpi_\nu \|_{L_{t,x}^\infty} \longrightarrow 0. \tag{3.2.6}$$

We will present here the main points from the theory developed by Terence Tao in [32], and later refined by Sterbenz and Tataru in [29], that leads to the decomposition (1.2.4) claimed in Theorem 1.2.3.

Before we start, let us adopt the convention, for this section only, that all function spaces we work with are to be understood as restricted to the time interval $[-1, 1]$, even if the notation does not reflect this directly. We shall not mention this further.

Let us introduce the space N, companion to the space S in iteration arguments establishing various well-posedness results for wave maps, like those in [32, 16, 34, 29, 17]. As with S, we are not going to give the full definition referring the reader to Section 5 of [29]. But we mention that N is built upon atomic spaces N_k, one for each frequency level $k \in \mathbb{Z}$, by ℓ^2 summing in k or also conveniently relying on frequency envelopes.

Spaces N_k contain several kind of atoms. For example, we have the typical energy/Strichartz type atoms satisfying

$$\|F\|_{L_t^1(L_x^2)} \leq 1,$$

with F frequency localized to $|\xi| \sim 2^k$, or the ones based on Fourier restriction spaces having

$$\|F\|_{\dot{X}_k^{0,-\frac{1}{2},1}} \leq 1,$$

again $|\xi| \sim 2^k$ for F. There are also the null frame atoms, which are absolutely crucial for dealing with the Q_0-null form present in the non-linearity of the wave maps equation (1.1.2). However, those are complicated enough for us to avoid reporting on them, as we actually did with the angular Strichartz spaces for S. But the reader is reassured that those spaces are not going to arise directly in any of the arguments in this thesis.

It might be helpful to think of N as a partial dual to S via the estimate:

$$\left| \int \int \phi \, P_k F \, dx dt \right| \lesssim 2^{-k} \|\phi\|_S \|P_k F\|_{N_k},$$

for Schwartz functions ϕ and N, see Section 10 in [32]. However, their main relationship is dictated by the following energy estimate, proved by Terence Tao in [32] (Section 11 there):

$$\|P_k \phi\|_S \lesssim \|P_k \phi[0]\|_{\dot{H}_x^2 \times L_x^2} + \|\Box P_k \phi\|_{N_k}.$$

Now, one certainly imagines how the iteration argument proceeds: its main aim is to suitably estimate the non-linearity of equation (1.1.2) in N. In fact, with this in mind, Tao developed a whole range of bi-linear and tri-linear estimates in [32] for the spaces N and S, which later got refined by Tataru [34] and Sterbenz-Tataru [29].

As one can imagine, looking at the papers cited above, this story is pretty long and complicated and reporting on it simply does not find its place in this thesis. What we shall do is to state the relevant outcome, that the interested reader can immediately obtain by combining Propositions 3.6 and 3.9 from [29] as black boxes.

For wave maps as in (3.2.6), there exists a positive constant $\delta_1 = \delta_1(\mathcal{N}) > 0$ such that setting:

$$T_{\nu,k}^i := (A_\alpha^{(k,\nu)})_j^i \, P_k \left[\partial^\alpha \varpi_\nu^j \right] - P_k \left[\mathcal{S}_{j\ell}^i(\varpi_\nu) \partial_\alpha \varpi_\nu^\ell \partial^\alpha \varpi_\nu^j \right],$$

where

$$(A_\alpha^{(k,\nu)})_j^i := P_{\leq k-m_\nu} \left[\mathcal{S}_{j\ell}^i(\varpi_\nu) - \mathcal{S}_{i\ell}^j(\varpi_\nu) \right] P_{\leq k-m_\nu} \left[\partial_\alpha \varpi_\nu^\ell \right],$$

for $m_\nu = \delta_1^{1/2} |\log \varepsilon_\nu|$ which degenerates as $\nu \to +\infty$, we have:

$$\sum_{k \in \mathbb{Z}} \|T_{\nu,k}\|_{N_k}^2 \lesssim \varepsilon_\nu^{2\delta_1}. \tag{3.2.7}$$

Perhaps, at this level of exposition, an important take-away point from the above estimate, is the fact that such a result cannot be true when considering, in the definition of $T_{\nu,k}$, the first term alone. Those are the infamous low-low-high frequency interactions which cause breakdown of well-posedness for critical semi-linear wave equations of this class.

That not everything was lost, and well-posedness under suitable smallness assumptions was still achievable for the wave maps equation (1.1.2), which we rewrite with the above notation under frequency localization as:

$$\left(\Box + 2A_\alpha^{(k,\nu)} \partial^\alpha \right) P_k \varpi_\nu = T_{\nu,k},$$

is consequence of the fact that $A_\alpha^{(k,\nu)} \in \mathfrak{so}(n)$. This anti-symmetric property is geometric in nature, and is obtained thanks to the identities (1.1.2) for the second fundamental form.

Tao's breakthrough in [31, 32] was to invent a gauge transformation for the above gauge co-variant wave equation which turned the difficult low-low-high interactions into high-low-high ones, amenable to the multi-linear estimates such as the ones used to establish (3.2.7). This gauge, called by Tao the microlocal gauge, was constructed under the assumption of smallness of energy for the wave map. However, the construction turned out to be so robust that it was with little pain that Sterbenz and Tataru generalized it to arbitrary energy levels in [29], coining the name diffusion gauge.

We will not attempt to report on the details of this whole construction. We will simply content ourselves with extracting the result that leads to the decomposition (1.2.4) in Theorem 1.2.3. The reader should arrive at the same conclusions combining Propositions 3.1, 3.2 and 3.9 from [29]. An analogous result was proved in [17].

Under assumptions (3.2.6) and choices of integers m_ν and matrices $A_\alpha^{(k,\nu)}$ as above, the Cauchy problem:

$$\begin{cases} \left(\Box + 2A_\alpha^{(k,\nu)} \partial^\alpha \right) \psi = \widetilde{P}_k F \\ \psi[0] = \widetilde{P}_k \varphi[0] \end{cases}, \tag{3.2.8}$$

given a Schwartz space-time function F and a tuple of initial data $\varphi[0] = (\varphi, \partial_t \varphi)$, admits a unique solution ψ satisfying:

$$\|P_{k'}\psi\|_S \lesssim 2^{-\delta_0|k'-k|} \left(\left\| \widetilde{P}_k \varphi[0] \right\|_{\dot{H}^2_x \times L^2_x} + \left\| \widetilde{P}_k F \right\|_{N_k} \right), \qquad (3.2.9)$$

for some universal $\delta_0 > 0$. The projections $\widetilde{P}_k$ stand for enlarged LP-projections to $|\xi| \sim 2^k$.

In the end, to obtain Theorem 1.2.3, it suffices to consider (3.2.8) with $F = 0$ and $\varphi[0] = \varpi_\nu[0]$, which gives us the linear gauge co-variant wave $\psi_{\nu,k}$. The difference $P_k \varpi_\nu - \psi_{\nu,k}$ solves then (3.2.8) with $F = T_{\nu,k}$ and $\varphi[0] = 0$. Finally, square-summing the estimates (3.2.9) for $P_k \varpi_\nu - \psi_{\nu,k}$ in $k \in \mathbb{Z}$ and applying Minkowski's inequality, we have:

$$\left\| \sum_{k\in\mathbb{Z}} \nabla_{t,x}(P_k\varpi_\nu - \psi_{\nu,k}) \right\|_{L^\infty_t(L^2_x)} \lesssim \left(\sum_{k'\in\mathbb{Z}} \left\| \sum_{k\in\mathbb{Z}} \nabla_{t,x} P_{k'}(P_k\varpi_\nu - \psi_{\nu,k}) \right\|^2_{L^\infty_t(L^2_x)} \right)^{1/2}$$

$$\lesssim \left(\sum_{k',k\in\mathbb{Z}} \|\nabla_{t,x} P_{k'}(P_k\varpi_\nu - \psi_{\nu,k})\|^2_{DS} \right)^{1/2}$$

$$\lesssim \left(\sum_{k',k\in\mathbb{Z}} 2^{-2\delta_0|k'-k|} \|T_{\nu,k}\|^2_{N_k} \right)^{1/2},$$

recalling the definitions of spaces S and DS from Section 2.2. The tri-linear estimate (3.2.7) leads then to the desired result, and Theorem 1.2.3 is proved.

3.3   Quantization of time-like energy for wave maps into spheres

In this whole section, we restrict ourselves to the case $\mathcal{N} = \mathbb{S}^{n-1}$.

To prove our last result, Theorem 1.2.4, we have to carry out the bubbling analysis for our sequence of wave maps $\{\phi_\nu\}_{\nu \in \mathbb{N}}$ near the set of time-like energy concentration:

$$(\cup_i B_{\delta_\nu}(\varrho_i)) \cap C^0_{[1,2]} \subset C^{\delta_0}, \tag{3.3.1}$$

recalling the set-up from Section 3.2, where $\delta_0 > 0$ controls the distance to the null boundary ∂C of the light cone, on which dependence of our constants will be considered universal.

3.3.1   Choosing suitable time slices

The dynamics of the maps ϕ_ν near distinct rays ϱ_i are completely disjoint and to get the claimed asymptotic decomposition from Theorem 1.2.4 we will have to select the time slices t_ν rather carefully.

To start, in order to obtain from the decay assumption (3.0.3) the asymptotic stationarity at all scales for some suitably chosen time slices, we consider a sequence of positive functions on the time interval $[1, 2]$ defined by:

$$\zeta_\nu(t) := \int_{S_t^{\delta_0}} |\partial_\rho \phi_\nu(t)|^2 \, dx,$$

so that $\|\zeta_\nu\|_{L_t^1[1,2]} \to 0$ by (3.0.3). Then, looking at the corresponding Hardy-Littlewood maximal functions:

$$\mathcal{M}\zeta_\nu(s) := \sup_{r > 0} \frac{1}{r} \int_{s-r}^{s+r} \zeta_\nu(t) dt,$$

the well-known maximal inequality of Hardy-Littlewood tells us that for any $\lambda > 0$:

$$|\{\mathcal{M}\zeta_\nu > \lambda\}| \lesssim \frac{1}{\lambda} \|\zeta_\nu\|_{L_t^1}.$$

Therefore taking a sequence $\lambda_\nu \sim \|\zeta_\nu\|_{L_t^1}^{1/2} \downarrow 0$ decaying slowly enough compared to $\|\zeta_\nu\|_{L_t^1}$, we can select a sequence of time slices $\{t_\nu\}_{\nu \in \mathbb{N}} \subset (1 + \delta_0, 2 - \delta_0)$ such that:

$$\mathcal{M}\zeta_\nu(t_\nu) \longrightarrow 0. \tag{3.3.2}$$

We should note here that this will not be quite the final sequence of time slices we will claim the soliton resolution on as we might need to perturb it a little at scales δ_ν.

From there, we have to study for each $i = 1, \ldots, I$, a sequence of wave maps obtained from ϕ_ν, upon translating by $(t_\nu, \varrho_i(t_\nu))$ and rescaling by δ_ν, which gives us by (3.2.3):

$$\widetilde{\phi}_{i,\nu}(\cdot) := \phi_\nu(t_\nu + \delta_\nu \cdot, \varrho_i(t_\nu) + \delta_\nu \cdot) \longrightarrow c_\phi \ \text{ on } \ ([-4,4] \times B_4) \setminus \varrho_i, \tag{3.3.3}$$

locally in $C_t^0(H_x^1) \cap C_t^1(L_x^2)$. Moreover from (3.3.2), denoting by X_i the unit constant time-like vector field pointing in the direction of the line ϱ_i, we have:

$$\left\| X_i \widetilde{\phi}_{i,\nu} \right\|_{L_{t,x}^2([-4,4] \times B_4)} \longrightarrow 0. \tag{3.3.4}$$

Proceeding as in Remark 2.2.3, we interpolate smoothly between $\widetilde{\phi}_{i,\nu}[0]$ and the constant initial data $(c_\phi, 0) \in T(\mathbb{S}^{n-1})$ on $B_4 \setminus B_3$, replacing the map $\widetilde{\phi}_{i,\nu}$ with a wave map $\phi_{i,\nu}$ agreeing with the latter on $[-\frac{3}{2}, \frac{3}{2}] \times B_{3/2}$ and constant outside B_6 (at most) for $t \in [-\frac{3}{2}, \frac{3}{2}]$ by finite speed of propagation. This introduces an error of asymptotically vanishing energy on this time interval, safely by (3.3.3). In fact, from the construction it is immediate that:

$$\phi_{i,\nu} - c_\phi \longrightarrow 0 \ \text{ in } \ C_t^0(L_x^2)[-\frac{3}{2}, \frac{3}{2}], \tag{3.3.5}$$

which improves to locally in $C_t^0(H_x^1) \cap C_t^1(L_x^2)$ away from ϱ_i, and we still have decay in a time-like direction:

$$\|X_i \phi_{i,\nu}\|_{L_{t,x}^2[-\frac{3}{2}, \frac{3}{2}]} \longrightarrow 0. \tag{3.3.6}$$

Let us fix a smooth time cut-off $\chi(t) \in C_0^\infty(-\frac{3}{2}, \frac{3}{2})$, identically 1 on $[-1, 1]$, so that we get now in position to apply Proposition 2.3.1, obtaining from (2.3.3) the following decomposition:

$$\chi \nabla_{t,x} \phi_{i,\nu} = \Theta_{i,\nu} + \Xi_{i,\nu}, \ \text{ with :}$$

$$\|\Theta_{i,\nu}\|_{L_{t,x}^2} \lesssim \|X_i \phi_{i,\nu}\|_{L_{t,x}^2[-\frac{3}{2}, \frac{3}{2}]} + \|\phi_{i,\nu} - c_\phi\|_{L_t^\infty(L_x^2)[-\frac{3}{2}, \frac{3}{2}]}, \tag{3.3.7}$$

$$\sum_{k \in \mathbb{Z}} \|P_k \Xi_{i,\nu}\|_{L_t^1(L_x^2)} \lesssim 1. \tag{3.3.8}$$

Furthermore, applying Lemma 2.3.2, we get from (2.3.18) a decomposition for second order time-like derivative of $\phi_{i,\nu}$:

$$\chi X_i^2 \phi_{i,\nu} = \Gamma_{i,\nu} + \Pi_{i,\nu},$$

where the first item is a linear combination of:

$$\sum_{k\in\mathbb{Z}} P_k \nabla_x [\Omega_x^{i,\nu}(P_{>k+10}\phi_{i,\nu})], \quad \sum_{k\in\mathbb{Z}} P_k [\Omega_x^{i,\nu}(P_{\leq k+10}\nabla_x\phi_{i,\nu})], \quad \text{and} \quad \Omega_{t,x}^{i,\nu}\nabla_{t,x}\phi_{i,\nu},$$

$$\text{with} \quad \Omega_\alpha^{i,\nu} := \phi_{i,\nu}\partial_\alpha\phi_{i,\nu}^\dagger - \partial_\alpha\phi_{i,\nu}\phi_{i,\nu}^\dagger,$$

while the second one satisfies (2.3.19):

$$\sum_{k\in\mathbb{Z}} 2^{-2k} \|P_k\Pi_{i,\nu}\|^2_{L^2_{t,x}[-\frac{3}{2},\frac{3}{2}]} \tag{3.3.9}$$

$$\lesssim (1 + \|X_i\phi_{i,\nu}\|_{L^2_{t,x}[-\frac{3}{2},\frac{3}{2}]}) \|X_i\phi_{i,\nu}\|_{L^2_{t,x}[-\frac{3}{2},\frac{3}{2}]}$$

$$+ (1 + \|\phi_{i,\nu} - c_\phi\|_{L^\infty_t(L^2_x)[-\frac{3}{2},\frac{3}{2}]}) \|\phi_{i,\nu} - c_\phi\|_{L^\infty_t(L^2_x)[-\frac{3}{2},\frac{3}{2}]}.$$

We note that the implicit constants, including the factors in the linear combination for $\Gamma_{i,\nu}$, depend only on the energy bound $\mathcal{E}$ from (2.3.1) and the distance δ_0 to the null boundary ∂C from (3.3.1), hence can be considered universal for the rest of the argument.

With this understood, we define non-negative functions $\vartheta_{i,\nu}$, $\xi_{i,\nu}$, $\zeta_{i,\nu}$, and $\pi_{i,\nu}$ for $i = 1, \ldots, I$ and $t \in [-1,1]$, setting:

$$\theta_{i,\nu}(t) := \|\Theta_{i,\nu}(t)\|^2_{L^2_x} \quad \text{with} \quad \|\theta_{i,\nu}\|_{L^1_t} \longrightarrow 0,$$

$$\xi_{i,\nu}(t) := \sum_{k\in\mathbb{Z}} \|P_k\Xi_{i,\nu}(t)\|_{L^2_x} \quad \text{with} \quad \|\xi_{i,\nu}\|_{L^1_t} = \sum_{k\in\mathbb{Z}} \|P_k\Xi_{i,\nu}(t)\|_{L^1_t(L^2_x)} \lesssim 1,$$

$$\zeta_{i,\nu}(t) := \|X_i\phi_{i,\nu}(t)\|^2_{L^2_x}, \quad \text{so that} \quad \|\zeta_{i,\nu}\|_{L^1_t} \longrightarrow 0,$$

as well as:

$$\pi_{i,\nu}(t) = \sum_{k\in\mathbb{Z}} 2^{-2k} \|P_k\Pi_{i,\nu}(t)\|^2_{L^2_x}, \quad \text{where} \quad \|\pi_{i,\nu}\|_{L^1_t} = \sum_{k\in\mathbb{Z}} 2^{-2k} \|P_k\Pi_{i,\nu}(t)\|^2_{L^2_{t,x}} \longrightarrow 0,$$

by (3.3.7) and (3.3.8), (3.3.4) and (3.3.5), and finally (3.3.9).

We will now choose a sequence of time slices where we uniformly control $\theta_{i,\nu}$ and have all of the other functions above asymptotically decaying. This will be used to prove decay of the weak Besov norm $\dot{B}_\infty^{1,2}$ on the neck regions, and ultimately get the energy collapsing there via the control on $\theta_{i,\nu}$. At the same time, to start this argument, we shall build first the weak bubble tree decomposition. To do so, one relies on the

small energy compactness result from Lemma 2.2.2 (which, for example, enables one to extract solitons from the standard concentration-compactness procedure). Hence, for that reason, we will need to control the maximal function $\mathcal{M}\zeta_{i,\nu}$ corresponding to $\|X_i\phi_{i,\nu}(t)\|_{L_x^2}^2$ as well.

Let us take $\lambda_\nu^{\theta_i} \sim \|\theta_{i,\nu}\|_{L_t^1}^{1/2} \downarrow 0$, $(\lambda_\nu^{\xi_i})^{-1}\|\xi_{i,\nu}\|_{L_t^1} < \epsilon$ for some arbitrarily small $\epsilon > 0$ to be fixed according to (3.3.10) below, as well as $\lambda_\nu^{\zeta_i} \sim \|\zeta_{i,\nu}\|_{L_t^1}^{1/2} \downarrow 0$ and $\lambda_\nu^{\pi_i} \sim \|\pi_{i,\nu}\|_{L_t^1}^{1/2} \downarrow 0$. Hence, applying Chebyshev's inequality and the maximal inequality of Hardy-Littlewood for $\mathcal{M}\zeta_{i,\nu}$, we get:

$$\sum_{i=1}^I \left(\left|\{\theta_{i,\nu} > \lambda_\nu^{\theta_i}\}\right| + \left|\{\xi_{i,\nu} > \lambda_\nu^{\xi_i}\}\right| + \left|\{\zeta_{i,\nu} > \lambda_\nu^{\zeta_i}\}\right| + \left|\{\pi_{i,\nu} > \lambda_\nu^{\pi_i}\}\right| \right. \qquad (3.3.10)$$
$$\left. + \left|\{\mathcal{M}\zeta_{i,\nu} > \lambda_\nu^{\zeta_i}\}\right| \right) < \frac{1}{10}.$$

Therefore, we can choose a sequence of time slices $\{t_\nu\}_{\nu\in\mathbb{N}} \subset [-\frac{1}{2}, \frac{1}{2}]$, that we may assume simply to be $t_\nu = 0$ upon translating the maps $\phi_{i,\nu}$ by $(t_\nu, \varrho_i(t_\nu))$ without changing notation for $\phi_{i,\nu}$ (and working on $[-\frac{1}{2}, \frac{1}{2}] \times B_6$), such that for all $i = 1, \ldots, I$ we have the following control:

$$\theta_{i,\nu}(0) \longrightarrow 0, \ \ \xi_{i,\nu}(0) \lesssim 1, \ \ \zeta_{i,\nu}(0) \longrightarrow 0, \ \ \pi_{i,\nu}(0) \longrightarrow 0, \qquad (3.3.11)$$

$$\text{and } \mathcal{M}\zeta_{i,\nu}(0) \longrightarrow 0.$$

These are the final time slices that we will consider and obtain the asymptotic decomposition on, as claimed in our main theorem. We start doing bubbling analysis on them just below. Here we just add the remark that, upon working in (3.3.10)-(3.3.11) with the maximal functions for $\theta_{i,\nu}$, $\xi_{i,\nu}$ and $\pi_{i,\nu}$ as well, it should be clear by end of the argument that we can also get the energy collapsing result for almost every time slice strictly within the lifespan of the fastest concentrating solitons.

3.3.2 The weak bubble tree decomposition

In the following lemma we present a preliminary version of the soliton decomposition. It is essentially the one that we aim towards from Theorem 1.2.2, but it contains errors that we shall call *necks* - those are wave maps on conformally degenerating annuli such that once localized in space they converge to a constant, but when considered on the whole annulus they might a priori carry a non-trivial amount of energy. Ruling out

such a scenario will be the last step in the proof of the main theorem.

We note that the proof of this lemma relies on a covering argument which goes back to at least Ding and Tian [6] and today is pretty standard in the literature on bubbling analysis of harmonic maps (and related areas, where some authors refer to it as *weak bubble tree convergence*). The lemma of course holds for any closed Riemannian manifold as a target.

Lemma 3.3.1. *Passing to a subsequence, there exists for each $i = 1, \ldots, I$ a collection of $J_i \lesssim_\varepsilon 1$ solitons $\omega_{j,i}$, $j = 1, \ldots, J_i$, with corresponding concentration points $a_{i,\nu}^j \in B_1$ converging to the origin, and scales $\lambda_{i,\nu}^j \downarrow 0$ satisfying the orthogonality relations:*

$$\frac{\lambda_{i,\nu}^j}{\lambda_{i,\nu}^{j'}} + \frac{\lambda_{i,\nu}^{j'}}{\lambda_{i,\nu}^j} + \frac{|a_{i,\nu}^j - a_{i,\nu}^{j'}|^2}{\lambda_{i,\nu}^j \lambda_{i,\nu}^{j'}} \longrightarrow \infty, \tag{3.3.12}$$

as $\nu \to \infty$ for j and j' distinct, such that:

$$\phi_{i,\nu}(\lambda_{i,\nu}^j t, a_\nu^j + \lambda_{i,\nu}^j x) \longrightarrow \omega_{j,i}(t, x) \ \ on \ \ \mathbb{R}^{2+1} \setminus \cup_q \varrho_q^{j,i},$$

locally in $C_t^0(H_x^1) \cap C_t^1(L_x^2)$ for a collection of at most $J_i - 1$ time-like geodesics $\varrho_q^{i,j}$ with direction X_i. Moreover, setting for any fixed positive constant $C > 0$:

$$\lambda_{\min,\nu} := C \cdot \min_{i,j} \left\{ \lambda_{i,\nu}^j \right\}, \ \ \nu \in \mathbb{N},$$

we have the following asymptotic decomposition holding for $t \in [-\lambda_{\min,\nu}, \lambda_{\min,\nu}]$:

$$\phi_{i,\nu}(t, x) - c_\phi = \sum_{j=1}^{J_i} \left(\omega_{j,i} \left(\frac{t}{\lambda_{i,\nu}^j}, \frac{x - a_{i,\nu}^j}{\lambda_{i,\nu}^j} \right) - \omega_{j,i}(\infty) \right) + \mathcal{N}_{i,\nu}(t, x) + o_{L_t^\infty(\dot{H}_x^1 \times L_x^2)}(1), \tag{3.3.13}$$

where $\mathcal{N}_{i,\nu}$ stands for the wave map $\phi_{i,\nu}$ restricted to a collection of $K_i \lesssim_\varepsilon 1$ sequences of degenerating annuli:

$$[-\frac{r_{i,\nu}^k}{2}, \frac{r_{i,\nu}^k}{2}] \times \left(B_{R_{i,\nu}^k}(x_{i,\nu}^k) \setminus B_{r_{i,\nu}^k}(x_{i,\nu}^k) \right) \subset [-\frac{1}{2}, \frac{1}{2}] \times B_3 \ \ with \ \ \lambda_{\min,\nu} \ll r_{i,\nu}^k \ll R_{i,\nu}^k, \tag{3.3.14}$$

such that we have:

$$\sup_{r_{i,\nu}^k \leq r \leq \frac{1}{2}R_{i,\nu}^k} \sup_{t \in [-\frac{r}{2}, \frac{r}{2}]} \mathcal{E}_{B_{2r}(x_{i,\nu}^k) \setminus B_r(x_{i,\nu}^k)}[\phi_{i,\nu}](t) \longrightarrow 0, \tag{3.3.15}$$

holding for each $k = 1, \ldots, K_i$.

Proof. Let us fix $i = 1, \ldots, I$, and suppress this subscript in the argument below to lighten the notation. In the same spirit, we also never change notation here whenever passing to a subsequence for $\{\phi_\nu\}_{\nu \in \mathbb{N}}$ while using Lemma 2.2.2 as it will be clear from the construction that we obtain in the end a countable cover of a suitable neighborhood of $\{t = 0\} \times B_3$ on which we can rely to build via the diagonal process a final subsequence that satisfies the claims of Lemma 3.3.1.

Pick a sequence of points $a_\nu^1 \in B_1$ with radii $\lambda_\nu^1 \downarrow 0$ such that:

$$\mathcal{E}_{B_{2\lambda_\nu^1}(a_\nu^1)}[\phi_\nu](0) = \epsilon_s. \qquad (3.3.16)$$

Note that such a concentration point is guaranteed to exist by the results of Section 3.1 and the compactness Lemma 2.2.2, and that by (3.3.3) any energy concentration point would have to converge to the origin.

Consider the sequence of balls $B_{2^k \lambda_\nu^1}(a_\nu^1)$ with k a positive integer, and choose the lowest $K_0 = K_0(\{a_\nu^1\}_{\nu \in \mathbb{N}}) \in \mathbb{N}$ such that the functions:

$$r_\nu : \overline{B_{2^{K_0-1}}(a_\nu^1)} \longrightarrow \mathbb{R}_{>0}$$

$$x \longmapsto r_\nu(x) := \sup\left\{r > 0 : \mathcal{E}_{B_r(x)}[\phi_\nu(\lambda_\nu^1 \cdot, a_\nu^1 + \lambda_\nu^1 \cdot)](0) \le \epsilon_s\right\}, \qquad (3.3.17)$$

which are continuous as the wave maps ϕ_ν are smooth, admit a collective positive lower bound $r' := \liminf_{x,\nu} r_\nu(x) > 0$ (assuming K_0 exists, the case when it does not is treated later when we describe convergence to solitons at infinity). As a preliminary step, relying on (3.3.11) and the compactness Lemma 2.2.2, we can obtain for the rescalings of the maps ϕ_ν at a_ν^1, upon passing to a subsequence, that:

$$\phi_\nu(\lambda_\nu^1 t, a_\nu^1 + \lambda_\nu^1 x) \longrightarrow \omega_1(t, x) \text{ on } \left[-\frac{r'}{3}, \frac{r'}{3}\right] \times B_{2^{K_0-1}},$$

in $C_t^0(H_x^1) \cap C_t^1(L_x^2)$ for some wave map ω_1 with regularity as in (2.2.9) and satisfying $X\omega_1 = 0$, with X standing for the constant time-like vector field X_i from (3.3.4). Therefore, the map ω_1 is part of a soliton.

The time interval $\left[-\frac{r'}{3}, \frac{r'}{3}\right]$ for the convergence above will be improved considerably below by recalling the methods from Section 3.1, see the proof of (3.3.20). Now, we shall proceed instead describing further ω_1 in space. Slightly abusing terminology, let us refer to ω_1 as a soliton already from here, bearing in mind that we will prove it is one shortly.

By construction, we can find at least one sequence of concentration points:

$$a_\nu^2 \in B_{2^{K_0}\lambda_\nu^1}(a_\nu^1) \setminus B_{2^{K_0-1}\lambda_\nu^1}(a_\nu^1), \tag{3.3.18}$$

bubbling off on the top of the soliton ω_1 in the sense that:

$$\mathcal{E}_{B_{2\lambda_\nu^2}(a_\nu^2)}[\phi_\nu](0) = \epsilon_s, \quad \lambda_\nu^2 \ll \lambda_\nu^1, \tag{3.3.19}$$

where it is quite important to note that we have an equality above, a fact that must hold by the compactness Lemma 2.2.2.

Let us consider a new sequence of concentration points satisfying (3.3.18) and (3.3.19) like $\{a_\nu^2\}_{\nu \in \mathbb{N}}$, in other words forming itself above the scales λ_ν^1 and converging, upon passing to a subsequence, in the closure of $B_{2^{K_0}\lambda_\nu^1}(a_\nu^1)$, so that it suffices to work in $B_{2^{K_0+1}\lambda_\nu^1}(a_\nu^1)$. There are of course uncountably many of those, given the existence of a single one, $\{a_\nu^2\}_{\nu \in \mathbb{N}}$, but we are going to consider equivalent all those for which the orthogonality condition (3.3.12) holds and pick only one representative per equivalence class. That is, if a sequence $\{a_\nu'\}_{\nu \in \mathbb{N}}$ satisfies (3.3.18) and (3.3.19) but in addition also has $\lambda_\nu' \sim \lambda_\nu^2$ with:

$$\frac{|a_\nu^2 - a_\nu'|}{\lambda_\nu^2} \lesssim 1,$$

then one can see that the maps $\phi_\nu(\lambda_\nu^2 t, a_\nu^2 + \lambda_\nu^2 x)$ and $\phi_\nu(\lambda_\nu' t, a_\nu' + \lambda_\nu' x)$ would converge on $[-2^{-1}, 2^{-1}] \times B_{2^{-1}}$, upon passing to a subsequence directly by Lemma 2.2.2, to the same soliton up to translation that we should denote by ω_2 as it was initially obtained from a_ν^2 once the procedure we are describing now for the soliton ω_1 is completed and applied to the soliton ω_2. Hence the sequence $\{a_\nu'\}_{\nu \in \mathbb{N}}$ should be discarded keeping $\{a_\nu^2\}_{\nu \in \mathbb{N}}$.

Given the orthogonality relations (3.3.12) holding between any two sequences of concentration points as above, we note that we are left with only finitely many possibilities, say $\{a_\nu^j\}_{\nu \in \mathbb{N}}$ with $j = 2, \ldots, J'$. This follows from the fact we are considering a sequence of functions $\{\nabla_{t,x}\phi_\nu\}_{\nu \in \mathbb{N}} \subset L_x^2$, bounded by the global energy control assumption (2.3.1), and with $\nabla_{t,x}\phi_\nu$ concentrating definite amounts of its L_x^2 norm, namely $\sqrt{\epsilon_s}$, note the *equality* in (3.3.19), at different frequency and/or spatial scales so that we can conclude that, since L_x^2 is a Hilbert space, we should have:

$$J' \lesssim \frac{1}{\epsilon_s}\mathcal{E},$$

which is a universal bound for us as desired.

The collection $\{a_\nu^j\}_{\nu \in \mathbb{N}}$, $j = 2, \ldots, J'$, gives rise to solitons ω_j, one for each j, by the same procedure as described for ω_1 and so from now on we should be running for each of them the same construction as we are currently considering for ω_1.

From the point of view of ω_1, we can subdivide the above collection of sequences of energy concentration points into disjoint families by considering the limit points $b_q^1 \in \overline{B_{2^{K_0}\lambda_\nu^1}(a_\nu^1)} \setminus B_{2^{K_0-1}\lambda_\nu^1}(a_\nu^1)$, indexed by $q = 1, \ldots, Q'$ for some integer $Q' \leq J'$, to which the sequences converge once rescaled by λ_ν^1. So for any $r > 0$ small but fixed, we have by Lemma 2.2.2:

$$\phi_\nu(\lambda_\nu^1 t, a_\nu^1 + \lambda_\nu^1 x) \longrightarrow \omega_1(t, x) \text{ on } [-\frac{r'}{3}, \frac{r'}{3}] \times \left(\overline{B_{2^{K_0}}} \setminus \cup_q B_r(b_q^1)\right),$$

in $C_t^0(H_x^1) \cap C_t^1(L_x^2)$ since the functions r_ν from (3.3.17) extended to $\overline{B_{2^{K_0}}} \setminus \cup_q B_r(b_q^1)$ admit a collective lower bound $r' := \liminf_{x,\nu} r_\nu(x) > 0$ (provided $r > 0$ is fixed of course as r' depends on it). Understanding the behavior of the maps ϕ_ν as $r \downarrow 0$ is linked to the convergence of ϕ_ν to solitons at the spatial infinity and this is when the neck domains enter into our picture. We shall discuss this straight after we finish the construction of the soliton ω_1 (and so for the other ones, ω_j above, in parallel).

Considering the annuli $\overline{B_{2^{K_0+k}\lambda_\nu^1}(a_\nu^1)} \setminus B_{2^{K_0+k-1}\lambda_\nu^1}(a_\nu^1)$ one after the other and studying as above whether there are new sequences of concentration points satisfying (3.3.19), upgrading the collection $\{a_\nu^j\}_{\nu \in \mathbb{N}}$, $j = 2, \ldots, J'$, accordingly upon checking the orthogonality relation (3.3.12) holds for each new member (we should not change the notation for the upgraded version), we must a reach an integer $K_1 = K_1(\{a_\nu^1\}_{\nu \in \mathbb{N}}, \epsilon_s, \mathcal{E}) \in \mathbb{N}$ such that for any $k \geq K_1$ the functions r_ν from (3.3.17) once considered on $\overline{B_{2^k}(a_\nu^1)} \setminus B_{2^{k-1}}(a_\nu^1)$ would admit a positive collective lower bound there. Note that this situation could have occurred without passing by the previous bubbling analysis induced by the existence of the integer K_0, e.g. if we would have picked up the fastest concentrating soliton initially for ω_1.

From there, we let $k \to \infty$ with $r \downarrow 0$ and fully construct the soliton ω_1 in the sense that we claim:

$$\phi_\nu(\lambda_\nu^1 t, a_\nu^1 + \lambda_\nu^1 x) \longrightarrow \omega_1(t, x) \text{ on } \mathbb{R}^{2+1} \setminus \cup_q \varrho_q^1, \tag{3.3.20}$$

locally in $C_t^0(H_x^1) \cap C_t^1(L_x^2)$ for a finite collection of geodesics ϱ_q^1, $q = 1, \ldots, Q'$, each passing through the corresponding point b_q^1, all with direction X, and such that $X\omega_1 = 0$ there. To prove (3.3.20), we note that by (3.3.11), used already above, we have for any

fixed bounded time interval the following decay estimate:

$$\int_{-s}^{s} \int_{\mathbb{R}^2} \left| X[\phi_\nu(\lambda_\nu^1 t, a_\nu^1 + \lambda_\nu^1 x)] \right|^2 dx dt = o(s), \tag{3.3.21}$$

and so denoting by Ψ the Lorentz boost taking ∂_t to X, if one considers the foliation induced by $\{\Psi(\{t\} \times \mathbb{R}^2)\}_{t \in \mathbb{R}}$ on the whole of Minkowski space $\mathbb{R}^{2+1}$ instead of the CMC foliation in the interior of the forward light cone as in Lemmata 3.1.1 and 3.1.2, the very same arguments would lead to the convergence claimed in (3.3.20). Let us present some details, setting $\varphi_\nu = \phi_\nu(\lambda_\nu^1 \cdot, a_\nu^1 + \lambda_\nu^1 \cdot)$.

Working on $\Psi^{-1}(\mathbb{R}^{2+1})$ we denote the coordinates there by $x^{\bar{\alpha}}$, or $(\bar{t}, \bar{x}_1, \bar{x}_2)$, and writing $\bar{\varphi}_\nu := \varphi_\nu \circ \Psi$ we get by the Lorentz invariance of smooth wave maps that the associated stress energy tensor $T_{\bar{\alpha}\bar{\beta}}[\bar{\varphi}_\nu]$ enjoys the conservation law $\partial^{\bar{\alpha}} T_{\bar{\alpha}\bar{\beta}}[\bar{\varphi}_\nu] = 0$. So, contracting $T[\bar{\varphi}_\nu]$ with the vector field $\chi(\bar{x})\partial_{\bar{t}}$, for some continuously differentiable test function χ with $\partial_{\bar{t}}\chi = 0$, and integrating the divergence of the Noether current $\partial^{\bar{\alpha}}((\chi(\bar{x})\partial_{\bar{t}}) P_{\bar{\alpha}})$ over the strip $\bar{t} \in [t, t + \lambda]$ for any $t \in \mathbb{R}$ and positive constant $\lambda > 0$ (similar considerations apply when $\lambda < 0$), we get by Stokes' theorem and the mentioned conservation law:

$$\int_{\{\bar{t}=t+\lambda\}} |\nabla_{\bar{t},\bar{x}}\bar{\varphi}_\nu|^2 \chi d\bar{x} - \int_{\{\bar{t}=t\}} |\nabla_{\bar{t},\bar{x}}\bar{\varphi}_\nu|^2 \chi d\bar{x}$$

$$= -2 \int_{[t,t+\lambda]\times\mathbb{R}_{\bar{x}}^2} \partial_{\bar{t}}\bar{\varphi}_\nu^\dagger (\partial_{\bar{x}_1}\bar{\varphi}_\nu \partial_{\bar{x}_1}\chi + \partial_{\bar{x}_2}\bar{\varphi}_\nu \partial_{\bar{x}_2}\chi) d\bar{t} d\bar{x}.$$

Hence, integrating the above identity over $t \in [t_0, t_1]$ for given $t_0, t_1 \in \mathbb{R}$, using the decay (3.3.21) we obtain:

$$\int_{[t_0,t_1]\times\mathbb{R}_{\bar{x}}^2} |\nabla_{\bar{t},\bar{x}}\bar{\varphi}_\nu|^2 \chi d\bar{t} d\bar{x} - \int_{[t_0+\lambda,t_1+\lambda]\times\mathbb{R}_{\bar{x}}^2} |\nabla_{\bar{t},\bar{x}}\bar{\varphi}_\nu|^2 \chi d\bar{t} d\bar{x} \longrightarrow 0, \tag{3.3.22}$$

analogously to (3.1.2) from Lemma 3.1.1. To use this asymptotic monotonicity formula to propagate small energy control, we note that we have $|\nabla_{\bar{t},\bar{x}}\bar{\varphi}_\nu| \sim |\nabla_{t,x}\varphi_\nu|$ with the implicit constant depending only on X, which is constant and fixed. Therefore, proceeding as in Lemma 3.1.2, given any point $y \in \mathbb{R}^2 \setminus \cup_q b_q^1$ and a positive constant $\eta > 0$, there exists a radius $r_1 = r_1(y, \eta) > 0$ such that:

$$\sup_{\nu \in \mathbb{N}} \sup_{t \in [-r_1, r_1]} \mathcal{E}_{B_{r_1}(y)}[\varphi_\nu](t) \leq \eta,$$

which leads to the control:

$$\sup_{\nu \in \mathbb{N}} \frac{1}{r_1} \int_{[-r_1, r_1] \times B_{r_1}(y)} |\nabla_{t,x} \varphi_\nu|^2 \, dt dx \lesssim \eta,$$

that in turn gives us, precomposing with Ψ and shrinking suitably the radius to $r_1 > r_2 \gtrsim r_1$:

$$\sup_{\nu \in \mathbb{N}} \frac{1}{r_2} \int_{[-r_2 + \bar{s}, r_2 + \bar{s}] \times B_{r_2}(\bar{y})} |\nabla_{\bar{t}, \bar{x}} \bar{\varphi}_\nu|^2 \, d\bar{t} d\bar{x} \lesssim \eta,$$

where $(\bar{s}, \bar{y}) := \Psi^{-1}(0, y)$. By the decay estimate (3.3.22), we get that given any $\lambda \in \mathbb{R}$:

$$\limsup_{\nu \in \mathbb{N}} \frac{1}{r_2} \int_{[-r_2 + \bar{s} + \lambda, r_2 + \bar{s} + \lambda] \times B_{r_2}(\bar{y})} |\nabla_{\bar{t}, \bar{x}} \bar{\varphi}_\nu|^2 \, d\bar{t} d\bar{x} \lesssim \eta,$$

and so going back to φ_ν by precomposing with Ψ^{-1}, shrinking further the radius to $r_2 > r_3 \gtrsim r_2$ we obtain by the pigeonhole principle, using the energy flux identity (1.2.2), the estimate:

$$\limsup_{\nu \in \mathbb{N}} \|\nabla_{t,x} \varphi_\nu\|_{L_t^\infty(L_x^2)\left(([-r_3, r_3] \times B_{r_3}(y)) + \lambda X\right)} \lesssim \eta,$$

for any given $\lambda \in \mathbb{R}$, viewing naturally $X \in \mathbb{R}^{2+1}$. All the implicit constants above being independent of η (and of λ, the dependence on which of our construction is hidden in the limsup), we can choose η small enough obtaining the small energy control for any fixed $\lambda \in \mathbb{R}$:

$$\limsup_{\nu \in \mathbb{N}} \|\nabla_{t,x} \varphi_\nu\|_{L_t^\infty(L_x^2)\left(([-r_3, r_3] \times B_{r_3}(y)) + \lambda X\right)} \leq \frac{1}{2} \epsilon_s, \tag{3.3.23}$$

with the radius $r_3 = r_3(y)$. Therefore, picking suitable collections of points $y \in \mathbb{R}^2 \setminus \cup_q b_q^1$ and constants $\lambda \in \mathbb{R}$, we construct a countable cover of $\mathbb{R}^{2+1} \setminus \cup_q \varrho_q^1$ such that relying on the estimates (3.3.21) and (3.3.23) we can apply Lemma 2.2.2 to get a subsequence via the diagonal process for which the local convergence claim (3.3.20) holds a desired.

Note that by construction ω_1 has energy bounded by $\mathcal{E}$, and so precomposing it with the Lorentz boost Ψ we get a steady in time finite energy harmonic map from $\mathbb{R}^2$ minus a finite set of points (note that the energy of this harmonic map will be smaller or equal to $\mathcal{E}[\omega_1]$, nothing travels faster than light!). By the regularity theory of Hélein [13] the latter has to be smooth and by the removable singularity theorem of Sacks and Uhlenbeck [25], it extends smoothly across the singular points. The outcome of this argument is therefore that ω_1 is a smooth finite energy wave map defined on the whole of $\mathbb{R}^{2+1}$ with $X\omega_1 = 0$, i.e. a genuine soliton as desired.

The same holds of course for the solitons ω_j, $j = 2, \ldots, J'$, but note that those do not of course constitute all the members of the decomposition (3.3.13) as parts of the maps ϕ_ν can get lost a priori at spatial infinity and in between the solitons we are considering. We shall address this issue now.

Consider the scales $\{\lambda_\nu^1\}_{\nu \in \mathbb{N}}$ corresponding to the soliton ω_1. Fix an arbitrary small $0 < \varepsilon < \epsilon_s$, then by the pigeonhole principle there exist an integer $K(\varepsilon) \geq K_1$ such that for any $k \in \mathbb{N}$ fixed:

$$\mathcal{E}_{B_{2^{K(\varepsilon)+k}\lambda_\nu^1}(a_\nu^1) \setminus B_{2^{K(\varepsilon)+k-1}\lambda_\nu^1}(a_\nu^1)}[\phi_\nu](0) < \varepsilon, \tag{3.3.24}$$

for all ν large enough. Suppose that there exist a sequence of smallest integer $k_\nu(\varepsilon) \geq K(\varepsilon)$, as ν gets large, such that the above inequality fails:

$$\mathcal{E}_{B_{2^{k_\nu(\varepsilon)+1}\lambda_\nu^1}(a_\nu^1) \setminus B_{2^{k_\nu(\varepsilon)}\lambda_\nu^1}(a_\nu^1)}[\phi_\nu](0) \geq \varepsilon,$$

and note that by construction we must have $k_\nu(\varepsilon) \to \infty$; then we have found a new soliton on the top of which our previous ω_1 is concentrating, that we should denote by $\omega_{J'+1}$ so that setting $\lambda_\nu^{J'+1} := 2^{k_\nu(\varepsilon)-1}\lambda_\nu^1$ we can apply directly Lemma 2.2.2, by the choice of $k_\nu(\varepsilon)$ and (3.3.24), to get:

$$\phi_\nu(\lambda_\nu^{J'+1}t, a_\nu^1 + \lambda_\nu^{J'+1}x) \longrightarrow \omega_{J'+1}(t,x) \text{ in } C_t^0(H_x^1) \cap C_t^1(L_x^2)([-\tfrac{1}{4}, \tfrac{1}{4}] \times (B_1 \setminus B_{\frac{1}{2}})),$$

$$\text{with } \mathcal{E}_{B_4 \setminus B_2}[\phi_\nu(\lambda_\nu^{J'+1}\cdot, a_\nu^1 + \lambda_\nu^{J'+1}\cdot)](0) \geq \varepsilon,$$

and the analysis we carried for ω_1 so far should also be applied to $\omega_{J'+1}$ now.

It should be clear that if no $k_\nu(\varepsilon)$ as above exist, i.e. (3.3.24) is not violated for any $k \in \mathbb{N}$ for ν large, then choosing $0 < \varepsilon < \epsilon_s$ small enough initially, by equality in (3.3.16) we must have been working with ω_1 and there should exist then a sequence of integers k_ν' such that $2^{k_\nu'}\lambda_\nu^1 \sim 1$ and (3.3.24) holding for any $k = 1, \ldots, k_\nu' - K(\varepsilon)$, with any $0 < \varepsilon' < \varepsilon$ for larger $k \geq k_\nu' - K(\varepsilon)$ by (3.3.3) as $\nu \to \infty$. The map $\omega_{J'+1}$ would be standing for the constant c_ϕ in this case.

For the other solitons ω_j, with $j \geq 2$, $k_\nu(\varepsilon)$ must exist and we could of course end up with ω_1, or also a constant (to which some authors refer to as a *ghost bubble*, i.e. a soliton on the top of which two or more non-constant solitons are concentrating but itself is constant) in which case we obviously do not consider this as a new soliton. This brings us to the final steps in the proof of Lemma 3.3.1.

In fact, in the above construction the constant $\varepsilon > 0$ could be arbitrarily small but

was initially fixed and we would like now to let it degenerate to 0. We claim that in fact we can put ourselves in a situation when for any smaller $0 < \varepsilon' < \varepsilon$ the choice of the integers $k_\nu(\varepsilon') \in \mathbb{N}$ is uniform in the sense that there exist positive integers $L(\varepsilon') \in \mathbb{N}$ independent of ν such that $k_\nu(\varepsilon') = k_\nu(\varepsilon) - L(\varepsilon')$, that is:

$$\sup_{K(\varepsilon') \leq k \leq k_\nu(\varepsilon) - L(\varepsilon')} \mathcal{E}_{B_{2^{k+1}\lambda_\nu^1}(a_\nu^1) \setminus B_{2^k \lambda_\nu^1}(a_\nu^1)}[\phi_\nu](0) < \varepsilon', \tag{3.3.25}$$

for ν large enough. If this were to fail for some $\varepsilon' > 0$, we could find a sequence of scales, that we denote by $\lambda_\nu^{J'+2}$, such that:

$$\mathcal{E}_{B_{2\lambda_\nu^{J'+2}}(a_\nu^1) \setminus B_{\lambda_\nu^{J'+2}}(a_\nu^1)}[\phi_\nu](0) > \varepsilon' \text{ and } \lambda_\nu^{J'+1} \ll \lambda_\nu^{J'+2} \ll \lambda_\nu^1, \tag{3.3.26}$$

and that would give rise to new non-constant solitons at scale $\lambda_\nu^{J'+2}$ or above, in which case we have to redefine ε as ε'. Note that we can have only finitely many non-constant solitons forming by the global energy bound (3.0.1) since those cannot have arbitrary small energy as this is not possible for harmonic 2-spheres, and by (3.3.26) they are asymptotically orthogonal in $\dot{H}_x^1 \times L_x^2$. Hence our procedure, applied to every single soliton we have found so far, detects all of the solitons in the claimed decomposition (3.3.13) and we are just left to characterize the regions in-between the domains of convergence to solitons as neck regions, but this can be obtained directly from (3.3.25) as follows.

Upon changing notation, by the above remarks we can assume that (3.3.25) holds. Now, we simply choose sequences $0 < r_\nu^1 \leq R_\nu^1$ tending to 0 slowly enough so that for any $\varepsilon > 0$ small enough:

$$\mathcal{E}_{B_{r_\nu^1}(a_\nu^1) \setminus B_{2^{K(\varepsilon)}\lambda_\nu^1}(a_\nu^1)}[\phi_\nu](0) \longrightarrow \mathcal{E}_{\mathbb{R}^2 \setminus B_{2^{K(\varepsilon)}}(0)}[\omega_1](0) \text{ and}$$

$$\mathcal{E}_{B_{\lambda_\nu^{J'+1}}(a_\nu^1) \setminus B_{R_\nu^1}(a_\nu^1)}[\phi_\nu](0) \longrightarrow \mathcal{E}_{B_1 \setminus \{0\}}[\omega_{J'+1}](0),$$

then by (3.3.25) there exits a sequence $\varepsilon_\nu^1 = \varepsilon_\nu^1(r_\nu^1, R_\nu^1) \downarrow 0$ such that:

$$\sup_{r_\nu^1 \leq r \leq \frac{1}{2}R_\nu^1} \sup_{t \in [-\frac{r}{2}, \frac{r}{2}]} \mathcal{E}_{B_{2r}(a_\nu^1) \setminus B_r(a_\nu^1)}[\phi_\nu](t) < \varepsilon_\nu^1.$$

If we know a priori that $r_\nu^1 \sim R_\nu^1$, then we can immediately absorb this part of the wave map ϕ_ν into the error term $o_{L_t^\infty(\dot{H}_x^1 \times L_x^2)}(1)$ in the decomposition (3.3.13) and there is no loss of energy between the considered solitons. Otherwise we should have $r_\nu^1 \ll R_\nu^1$, i.e.

the annulus is conformally degenerating, and this is precisely a neck in our terminology, as required. To prove Theorem 1.2.2 we must show that those terms can also be absorbed into $o_{\dot{H}^1_x \times L^2_x}(1)$ upon picking a suitable time slice, but that is the next and final step of the whole argument. So far we have established Lemma 3.3.1.   □

Remark 3.3.2. We note here that our techniques cannot say anything more about the decomposition beyond the scales $\{O(\lambda_{\min,\nu})\}_{\nu \in \mathbb{N}}$, which is a central issue to address if one were to try understanding the full soliton resolution conjecture.

Let us also remark that there is also quite some freedom in fixing the radii $R^k_{i,\nu}$ and $r^k_{i,\nu}$ defining the neck domain, as for any positive integer $\ell \in \mathbb{N}$ which can be arbitrarily large but fixed, we still have:

$$\sup_{2^{-\ell} r^k_{i,\nu} \leq r \leq 2^\ell R^k_{i,\nu}} \sup_{t \in [-\frac{r}{2}, \frac{r}{2}]} \mathcal{E}_{B_{2r}(x^k_{i,\nu}) \setminus B_r(x^k_{i,\nu})}[\phi_{i,\nu}](t) \longrightarrow 0,$$

which follows directly from the characterization (3.3.25) in the proof of Lemma 3.3.1 above.

3.3.3   Decay in the weak Besov norm for neck regions

Our aim now is to show energy collapsing for the necks $\mathcal{N}_{i,\nu}$, that is a decay to zero for the L^2_x norm of $\nabla_{t,x}\phi_\nu$ as $\nu \to +\infty$ on the degenerating annuli (3.3.14). We shall start by obtaining a decay in the weaker Besov $\dot{B}^{1,2}_\infty$ norm for $\mathcal{N}_{i,\nu}$, as consequence of the property (3.3.15), up to an error whose $\dot{H}^1_x$ norm is controlled by the L^2_x norm of $X\phi_\nu$ for some time-like vector field X that we will fix according to (3.3.11) later. This is the content of the following lemma.

Lemma 3.3.3. *Consider a sequence of smooth wave maps of bounded energy:*

$$\phi_\nu : [-2^{N_\nu + O(1)}, 2^{N_\nu + O(1)}] \times \mathbb{R}^2 \longrightarrow \mathbb{S}^{n-1}, \quad \|\nabla_{t,x}\phi_\nu\|^2_{L^\infty_t(L^2_x)} \leq \mathcal{E}, \tag{3.3.27}$$

obtained from Lemma 3.3.1 up to translating and rescaling, where we are given two sequences of positive integers $n_\nu, N_\nu \to +\infty$, $n_\nu \ll N_\nu$, such that the neck property holds on $B_{2^{N_\nu}} \setminus B_{2^{n_\nu}}$:

$$\sup_{n_\nu \leq \ell \pm O(1) \leq N_\nu} \|\nabla_{t,x}\phi_\nu\|_{L^\infty_t(L^2_x)\left([-2^{\ell-1}, 2^{\ell-1}] \times (B_{2^{\ell+1}} \setminus B_{2^\ell})\right)} \longrightarrow 0. \tag{3.3.28}$$

Moreover, we assume the maps are asymptotically steady in the direction of a constant

time-like vector field X, standing for one of the X_i's from (3.3.4) which we can take to be given by (2.3.2):

$$\|X\phi_\nu(0)\|_{L_x^2} \longrightarrow 0, \tag{3.3.29}$$

and the second order time-like derivatives satisfy:

$$\Pi_{X,\nu} := \operatorname{sech}^2(\zeta)X^2\phi_\nu - \Omega_\alpha^\nu \partial^\alpha \phi_\nu$$
$$+ \sum_{k\in\mathbb{Z}} P_k \left[\nabla_x \cdot (\Omega_{x,\beta}^\nu P_{>k+10}\phi_\nu) + \Omega_{x,\beta}^\nu \cdot P_{\leq k+10}\nabla_x\phi_\nu \right],$$

$$\sum_{k\in\mathbb{Z}} 2^{-2k} \|P_k\Pi_{X,\nu}(0)\|_{L_x^2}^2 \longrightarrow 0, \tag{3.3.30}$$

setting $\Omega_\alpha^\nu := \phi_\nu\partial_\alpha\phi_\nu^\dagger - \partial_\alpha\phi_\nu\phi_\nu^\dagger$ and $\Omega_{x,\beta}^\nu := (1-\beta^2)\Omega_{x_1}^\nu dx_1 + \Omega_{x_2}^\nu dx_2$. Both assumptions are justified by (3.3.11).

Then on the neck region, we can write for the map ϕ_ν:

$$\nabla_{t,x}\phi_\nu = \Upsilon_\nu \quad on \quad [-1,1] \times (B_{2^{N_\nu}} \setminus B_{2^{n_\nu}}),$$

see (3.3.36) in the proof, with $\Upsilon_\nu(t) \in C_0^\infty(B_{2^{N_\nu+1}} \setminus B_{2^{n_\nu-1}})$ for $t \in [-1,1]$ being of bounded energy $\|\Upsilon_\nu\|_{L_t^\infty(L_x^2)[-1,1]}^2 \lesssim \mathcal{E}$, and satisfying the following weak decay estimate on $t = 0$:

$$\sup_{k\in\mathbb{Z}} \|P_k\Upsilon_\nu(0)\|_{L_x^2} \longrightarrow 0.$$

The strategy of our argument is roughly to replace, by using the decay in the direction of the time-like vector field X, the sequence of wave maps on neck domains under consideration with another one, differing by an error of vanishing energy and converging locally to a constant on the neck domain with more regularity than $\dot{H}_x^1 \times L_x^2$ for ϕ_ν. However, because we need to obtain estimates that are uniform in time, working on very short intervals, we should not rely on the small energy regularity theory from Theorem 2.2.1 and the direct use of Fourier restriction spaces, as in the proof of the compactness result by Sterbenz and Tataru [30] (Proposition 5.1 there), but proceed directly via the wave maps equation (1.1.6) proving a weak $\dot{B}_\infty^{-1,2}$ decay estimate for its quadratic structure in the gradient at high frequency (without any null-structure involved, hence having target $\mathbb{S}^{n-1}$ is not specifically necessary for this part of the argument), and then using Lemma 2.3.2 to control the second order time-like derivatives (the latter though does involve the conservation law (1.1.9) for wave maps into spheres).

Proof. As usual, having the required control in a time-like direction, it is enough to

consider the spatial gradient only. Now working on the domain $[-1,1] \times (B_{2^{N_\nu}} \setminus B_{2^{n_\nu}})$, we note it being arbitrarily rough in time as $n_\nu, N_\nu \to +\infty$ degenerates. This is an additional difficulty, to be dealt with in the present proof, in comparison to the analogous estimate for harmonic maps, where ε-regularity is used on the domains $[-2^{\ell-1}, 2^{\ell-1}] \times (B_{2^{\ell+1}} \setminus B_{2^\ell})$ instead, see for the example the paper of Lin and Rivière [21] on page 188.

Before taking the main line of the argument, let us start with some preliminaries, fixing the decay rates for the assumptions of Lemma 3.3.3, that is sequences $\iota_\nu \downarrow 0$, $\sigma_\nu \downarrow 0$ and $\varepsilon_\nu \downarrow 0$ for which:

$$\sum_{k \in \mathbb{Z}} 2^{-2k} \left\| P_k \Pi_{X,\nu}(0) \right\|_{L_x^2}^2 \leq \iota_\nu^2, \tag{3.3.31}$$

$$\left\| X\phi_\nu(0) \right\|_{L_x^2} \leq \sigma_\nu, \tag{3.3.32}$$

$$\sup_{n_\nu \leq \ell \pm O(1) \leq N_\nu} \left\| \nabla_{t,x}\phi_\nu \right\|_{L_t^\infty(L_x^2)\left([-2^{\ell-1}, 2^{\ell-1}] \times (B_{2^{\ell+1}} \setminus B_{2^\ell})\right)} \leq \varepsilon_\nu, \tag{3.3.33}$$

corresponding to (3.3.30), (3.3.29) and (3.3.28) respectively. Next, we consider, for an arbitrary choice of integers ℓ_ν between n_ν and N_ν, the sequence of wave maps:

$$\phi_{\nu,\ell_\nu}(\cdot) := \phi_\nu(2^{\ell_\nu} \cdot) : [-2^{-4}, 2^{-4}] \times (B_{2^3} \setminus B_{2^{-3}}) \longrightarrow \mathbb{S}^{n-1}. \tag{3.3.34}$$

We build an extension ψ_{ν,ℓ_ν} of ϕ_{ν,ℓ_ν}, as in Remark 2.2.3, by smoothly interpolating on $(B_{2^{-2}} \setminus B_{2^{-3}}) \cup (B_{2^2} \setminus B_{2^2-1})$ between $\phi_{\nu,\ell_\nu}[0]$ and $(c_{\ell_\nu}, 0) \in T(\mathbb{S}^{n-1})$, for some suitably chosen sequence of constants $c_{\ell_\nu} = c_{\ell_\nu}(\phi_{\nu,\ell_\nu})$, solving the wave maps equation for ψ_{ν,ℓ_ν} with initial data of $\psi_{\nu,\ell_\nu}[0]$, such that scaling back and setting $\psi_\nu^{\ell_\nu}(\cdot) := \psi_{\nu,\ell_\nu}(2^{-\ell_\nu}\cdot)$, we have (denoting by 1_{ℓ_ν} the characteristic function of $B_{2^{\ell_\nu+1}} \setminus B_{2^{\ell_\nu-1}}$ over the time interval $[-2^{\ell_\nu-3}, 2^{\ell_\nu-3}]$):

$$\left\| \nabla_{t,x}\psi_\nu^{\ell_\nu} \right\|_{L_t^\infty(L_x^2)} \lesssim \varepsilon_\nu \text{ and } 1_{\ell_\nu}\phi_\nu = 1_{\ell_\nu}\psi_\nu^{\ell_\nu}, \tag{3.3.35}$$

by (3.3.33) and the finite speed of propagation property respectively.

From there, we construct a partition of unity over $[-1,1] \times (B_{2^{N_\nu}} \setminus B_{2^{n_\nu}})$ paralleling the Littlewood-Paley decomposition in frequency space. For the spatial directions, we recall the non-negative radial bump functions m_0 and $m_{\leq 0}$ used in the definition of the LP-projections P_0 and $P_{\leq 0}$, but which this time, we will use on the physical space setting:

$$\bar{m}_0(t,x) := m_0(|x|), \quad \bar{m}_\ell(t,x) := \bar{m}_0(2^{-\ell}t, 2^{-\ell}x),$$

$$\bar{m}_{\leq 0}(t,x) := m_{\leq 0}(|x|), \quad \bar{m}_{\leq \ell}(t,x) := \bar{m}_{\leq 0}(2^{-\ell}t, 2^{-\ell}x).$$

We get then the following "physical LP-decomposition":

$$\Upsilon_\nu := (\bar{m}_{\leq N_\nu} - \bar{m}_{\leq n_\nu - 1})\eta \nabla_x \phi_\nu = \sum_{\ell_\nu = n_\nu}^{N_\nu} \eta \bar{m}_{\ell_\nu} \nabla_x \phi_\nu. \tag{3.3.36}$$

where $\eta(t)$ stands for the rough cut-off to the time interval $[-1,1]$, and of course it is immediate that $\|\Upsilon_\nu\|^2_{L^\infty_t(L^2_x)[-1,1]} \lesssim \mathcal{E}$. Moreover we note that, recalling the extensions (3.3.35), we have $\eta \bar{m}_{\ell_\nu} \phi_\nu = \eta \bar{m}_{\ell_\nu} \psi_\nu^{\ell_\nu}$.

Writing $\phi_\nu^c := \phi_\nu - c_{\ell_\nu}$, for an arbitrary sequence of maps corresponding to (3.3.34), and similarly for ϕ_{ν,ℓ_ν}^c, together with the extensions $\psi_\nu^{\ell_\nu,c}$ and ψ_{ν,ℓ_ν}^c from (3.3.35) which become compactly supported by construction, we consider the commutator (denoting the cut-off functions by $\chi_{\ell_\nu} := \eta \bar{m}_{\ell_\nu}$):

$$\chi_{\ell_\nu} \nabla_x \phi_\nu = \nabla_x(\chi_{\ell_\nu} \phi_\nu^c) - (\nabla_x \chi_{\ell_\nu})\phi_\nu^c, \tag{3.3.37}$$

and start by treating the second term, for which we claim:

$$\|P_k[(\nabla_x \chi_{\ell_\nu})\phi_\nu^c]\|_{L^\infty_t(L^2_x)} \lesssim 2^{-|k+\ell_\nu|}\varepsilon_\nu, \tag{3.3.38}$$

for any $k \in \mathbb{Z}$. To see this, we rescale by 2^{ℓ_ν}. For high frequency scales $2^k \gtrsim 1$, we can use the extra regularity, the spatial derivative falling on the cut-off instead of the map, available from:

$$\left\|\nabla_x[(\nabla_x \bar{m}_0)\phi_{\nu,\ell_\nu}^c]\right\|_{L^\infty_t(L^2_x)} \lesssim \left\|(\nabla_x^2 \bar{m}_0)\psi_{\nu,\ell_\nu}^c\right\|_{L^\infty_t(L^2_x)} + \left\|(\nabla_x \bar{m}_0)\nabla_x \phi_{\nu,\ell_\nu}\right\|_{L^\infty_t(L^2_x)},$$

introducing the extensions ψ_{ν,ℓ_ν}^c, so that applying Poincaré's inequality in L^2_x for the first term, given the spatial localization of ψ_{ν,ℓ_ν}^c at any given time slice in the support of $\eta_{\ell_\nu}(\cdot) := \eta(2^{\ell_\nu}\cdot)$, we get by the finite band property (2.1.3) and the bound (3.3.35):

$$\left\|\eta_{\ell_\nu} P_k[(\nabla_x \bar{m}_0)\psi_{\nu,\ell_\nu}^c]\right\|_{L^\infty_t(L^2_x)} \lesssim 2^{-k}\varepsilon_\nu,$$

as desired. For low frequency scales $2^k \lesssim 1$, by Cauchy-Schwarz and Poincaré's inequalities, we have:

$$\left\|\eta_{\ell_\nu}(\nabla_x \bar{m}_0)\psi_{\nu,\ell_\nu}^c\right\|_{L^\infty_t(L^1_x)} \lesssim \left\|\eta_{\ell_\nu} \nabla_x \psi_{\nu,\ell_\nu}\right\|_{L^\infty_t(L^2_x)},$$

dropping $\nabla_x \bar{m}_0$, and so using Bernstein's inequality (2.1.4) we obtain here an exponen-

tial gain as well:
$$\left\|\eta_{\ell_\nu}P_k[(\nabla_x\bar{m}_0)\psi_{\nu,\ell_\nu}^c]\right\|_{L_t^\infty(L_x^2)} \lesssim 2^k\varepsilon_\nu,$$

by the energy bound (3.3.35). Hence, claim (3.3.38) follows.

We remark that, by the same argument, we get also control for the low frequencies of the first term $\nabla_x(\chi_{\ell_\nu}\phi_\nu^c)$ in the commutator:

$$\|P_k\nabla_x(\chi_{\ell_\nu}\phi_\nu^c)\|_{L_t^\infty(L_x^2)} \lesssim 2^{k+\ell_\nu}\varepsilon_\nu, \quad k \le -\ell_\nu + O(1), \tag{3.3.39}$$

and so it remains to treat now the main terms, that is the LHS above when $\ell_\nu \ge -k$, for which we should rely on the wave maps equation, the time-like control assumption (3.3.32), as well as the favorable decay (3.3.31) we already have.

Recalling the expression for the operator (2.3.9), we compute then:

$$\begin{aligned}
\Delta_{x,\beta}(\chi_{\ell_\nu}\phi_\nu^c) =&(\Delta_{x,\beta}\chi_{\ell_\nu})\phi_\nu^c + 2(1-\beta^2)(\partial_{x_1}\chi_{\ell_\nu})(\partial_{x_1}\phi_\nu) + 2(\partial_{x_2}\chi_{\ell_\nu})(\partial_{x_2}\phi_\nu) \\
&- 2\chi_{\ell_\nu}\mathrm{sech}^2(\zeta)\sinh(\zeta)\partial_{x_1}X\phi_\nu \\
&+ \chi_{\ell_\nu}(\mathrm{sech}^2(\zeta)X^2\phi_\nu - \Omega_\alpha^\nu\partial^\alpha\phi_\nu).
\end{aligned} \tag{3.3.40}$$

Let us treat first the smooth terms on the first line of (3.3.40), of which there are two types, $(\nabla_x^2\chi_{\ell_\nu})\psi_\nu^{\ell_\nu,c}$ and $\nabla_x\chi_{\ell_\nu}\nabla_{t,x}\psi_\nu^{\ell_\nu}$, the cut-off differentiated in a spatial direction, claiming for both the control:

$$\left\|\frac{\nabla_x}{\Delta_{x,\beta}}P_k[(\nabla_x^2\chi_{\ell_\nu})\psi_\nu^{\ell_\nu,c} + \nabla_x\chi_{\ell_\nu}\nabla_{t,x}\psi_\nu^{\ell_\nu}]\right\|_{L_t^\infty(L_x^2)} \lesssim 2^{-(k+\ell_\nu)}\varepsilon_\nu, \quad k \ge -\ell_\nu. \tag{3.3.41}$$

To show this, relying on Plancherel in L_x^2, we discard the Fourier multiplier $2^k\nabla_x\Delta_{x,\beta}^{-1}\widetilde{P}_k$ (where $\widetilde{P}_k = P_{k-1\le\cdot\le k+1}$), having symbol bounded uniformly in $k \in \mathbb{Z}$. Rescaling by 2^{ℓ_ν} we are brought to estimate for $k \ge O(1)$:

$$2^{-k}\left\|\eta_{\ell_\nu}[(\nabla_x^2\bar{m}_0)\psi_{\nu,\ell_\nu}^c + \nabla_x\bar{m}_0\nabla_{t,x}\psi_{\nu,\ell_\nu}]\right\|_{L_t^\infty(L_x^2)},$$

where the second term is directly seen to have the desired control by (3.3.33), whereas for the first one, given the spatial support of the extension ψ_{ν,ℓ_ν}^c, we apply Poincaré's inequality in L_x^2 as before, which allows us to conclude by (3.3.35).

The second line of (3.3.40) is an error term controlled thanks to the time-like decay

(3.3.32) we have. We first write:

$$\chi_{\ell_\nu} \nabla_x X \phi_\nu = \nabla_x (\chi_{\ell_\nu} X \phi_\nu) - (\nabla_x \chi_{\ell_\nu}) X \psi_\nu^{\ell_\nu},$$

and note that the second term here was already treated in (3.3.41), and so we just need to show:

$$\left\| \frac{\nabla_x^2}{\Delta_{x,\beta}} P_k \sum_{\ell_\nu = \max(-k, n_\nu)}^{N_\nu} (\chi_{\ell_\nu} X \phi_\nu)(0) \right\|_{L_x^2} \lesssim \sigma_\nu, \qquad (3.3.42)$$

but this follows at once by Plancherel in L_x^2, as the Fourier multiplier $\nabla_x^2 \Delta_{x,\beta}^{-1} P_k$ has a bounded symbol, dropping the cut-offs and relying on (3.3.32).

Finally, we shall consider the delicate second order time-like derivatives and the non-linear terms on the third line of (3.3.40). As was already required for (3.3.42), we restrict ourselves from now on to work exclusively over the time slice $t = 0$. And to lighten the notation, we shall not mention this explicitly anymore.

Thanks to the assumption (3.3.31), we have already partial control on them through $\Pi_{X,\nu}$, which however we need to localize to the neck region $B_{2^{N_\nu+1}} \setminus B_{2^{\max(-k,n_\nu)-1}}$. In doing so, we first note that since $\bar{m}_{\leq 0}$ was initially fixed spatially Schwartz, we have:

$$\|\nabla_x \widetilde{m}_{k,N_\nu}\|_{L_x^2} \lesssim 1, \text{ where } \widetilde{m}_{k,N_\nu} := \bar{m}_{\leq N_\nu} - \bar{m}_{\leq \max(-k,n_\nu)-1},$$

given that the above norm is scale invariant. Hence applying the Littlewood-Paley trichotomy to $\widetilde{m}_{k,N_\nu} \Pi_{X,\nu}$, we get:

$$
\begin{aligned}
P_k(\widetilde{m}_{k,N_\nu} \Pi_{X,\nu}) = P_k[& (P_{\leq k-7} \widetilde{m}_{k,N_\nu})(P_{k-3 \leq \cdot \leq k+3} \Pi_{X,\nu}) \\
& + (P_{k-3 \leq \cdot \leq k+3} \widetilde{m}_{k,N_\nu})(P_{\leq k-7} \Pi_{X,\nu}) \\
& + \sum_{k_1, k_2 \geq k-6 : |k_1 - k_2| \leq O(1)} (P_{k_1} \widetilde{m}_{k,N_\nu})(P_{k_2} \Pi_{X,\nu})].
\end{aligned}
$$

From there, using (3.3.31), we estimate the low-high interactions by:

$$2^{-k} \left\| (P_{\leq k-7} \widetilde{m}_{k,N_\nu})(P_{k-3 \leq \cdot \leq k+3} \Pi_{X,\nu}) \right\|_{L_x^2} \lesssim \|\widetilde{m}_{k,N_\nu}\|_{L_x^\infty} \iota_\nu,$$

the high-low ones by:

$$2^{-k} \left\| (P_{k-3 \leq \cdot \leq k+3} \widetilde{m}_{k,N_\nu}) \Big(\sum_{k_1 \leq k-7} P_{k_1} \Pi_{X,\nu} \Big) \right\|_{L_x^2} \lesssim \|\widetilde{m}_{k,N_\nu}\|_{L_x^\infty} \sum_{k_1 \leq k-7} 2^{-(k-k_1)} \iota_\nu,$$

whereas for the high-high cascade we have:

$$\sum_{k_1,k_2 \geq k-6 : |k_1-k_2| \leq O(1)} 2^{-k} \left\| \eta(P_{k_1}\widetilde{m}_{k,N_\nu})(P_{k_2}\Pi_{X,\nu}) \right\|_{L_x^2} \lesssim \left(\sum_{k_1} 2^{2k_1} \left\| P_{k_1}\widetilde{m}_{k,N_\nu} \right\|_{L_x^2}^2 \right)^{\frac{1}{2}} \iota_\nu,$$

where we have used Bernstein's inequality (2.1.4) passing to L_x^1, and then Cauchy-Schwarz with the fact that $k_1 = k_2 + O(1)$.

Putting those estimates together we get the required control for $\widetilde{m}_{k,N_\nu}\Pi_{X,\nu}$:

$$\left\| \frac{\nabla_x}{\Delta_{x,\beta}} P_k \sum_{\ell_\nu=\max(-k,n_\nu)}^{N_\nu} \chi_{\ell_\nu}\Pi_{X,\nu} \right\|_{L_x^2} \lesssim \iota_\nu, \tag{3.3.43}$$

by discarding the multiplier $2^k \nabla_x \Delta_{x,\beta}^{-1}\widetilde{P}_k$ and relying on the bounds for the cut-offs $\widetilde{m}_{k,N_\nu}$ discussed above.

We treat now the non-linear bulk left from Lemma 2.3.2, decomposing it into:

$$B_1^\nu := \sum_{k\in\mathbb{Z}} P_k \nabla_x \cdot (\Omega_{x,\beta}^\nu \phi_\nu^{>k+10}),$$

$$B_2^\nu := \sum_{k\in\mathbb{Z}} P_k(\Omega_{x,\beta}^\nu \cdot \nabla_x \phi_\nu^{\leq k+10}),$$

introducing the convenient notation $\phi_\nu^k := P_k\phi_\nu$ (also later $\phi_{\nu,\ell_\nu}^k := P_k\phi_{\nu,\ell_\nu}$ for the rescaled maps), etc. We want to treat this term perturbatively, as in elliptic regularity theory, and so we proceed claiming first the following $\dot{B}_\infty^{-1,2}$ estimate:

$$\left\| \frac{\nabla_x}{\Delta_{x,\beta}} P_k \sum_{\ell_\nu=\max(-k,n_\nu)}^{N_\nu} \chi_{\ell_\nu} B_i^\nu \right\|_{L_x^2}^2 \tag{3.3.44}$$

$$\lesssim \sum_{\jmath\geq 0} 2^{-\jmath} \sum_\ell \left\| \chi_{\ell_\nu} B_i^\nu \right\|_{L_x^1} \left\| \chi_{\ell_\nu+\jmath} B_i^\nu \right\|_{L_x^1},$$

where the sums are such that both ℓ_ν and $\ell_\nu + \jmath$ range between $\max(-k,n_\nu)$ and N_ν.

Discarding the Fourier multiplier $2^k \nabla_x \Delta_{x,\beta}^{-1}\widetilde{P}_k$ via Plancherel in L_x^2, we note the Littlewood-Paley projection P_k in front of the sum in (3.3.44) is crucial to handle the remaining factor 2^{-k}. But frequency localization induces spreading for the physical support by the uncertainty principle. And so, we are not allowed to use a square-summing trick relying on the finitely overlapping supports of $\chi_{\ell_\nu} B_i^\nu$. On the other

hand, this leakage is very much controllable given the fact that $k \geq -\ell_\nu + O(1)$, which corresponds to high frequency here.

More precisely, let us bound the LHS of (3.3.44) via:

$$2^{-2k} \sum_{\mu_\nu \geq \ell_\nu} \left| \int_{\mathbb{R}^2} [P_k(\chi_{\ell_\nu} B_i^\nu)] \, [P_k(\chi_{\mu_\nu} B_i^\nu)] \, dx \right|,$$

with both ℓ_ν and μ_ν ranging between $\max(-k, n_\nu)$ and N_ν. By the self-adjointness of P_k, the summand above can be estimated by:

$$\left\| \left[P_k^2(\chi_{\ell_\nu} B_i^\nu) \right] \chi_{\mu_\nu} B_i^\nu \right\|_{L_x^1}$$
$$\leq \left\| P_k^2(\chi_{\ell_\nu} B_i^\nu) \right\|_{L_x^\infty(\{|x| \sim 2^{\mu_\nu}\})} \left\| \chi_{\mu_\nu} B_i^\nu \right\|_{L_x^1}.$$

Now, looking at the convolution kernel for P_k^2, analogue to (2.1.1), we can estimate the first factor on the RHS above by:

$$\left\| P_k^2(\chi_{\ell_\nu} B_i^\nu) \right\|_{L_x^\infty(\{|x| \sim 2^{\mu_\nu}\})} \lesssim 2^{2k} 2^{-(\mu_\nu - \ell_\nu)} \left\| \chi_{\ell_\nu} B_i^\nu \right\|_{L_x^1},$$

for $\mu_\nu \geq \ell_\nu \geq -k$, a refined version of Bernstein's inequality (2.1.4). Hence, this leads us to estimate the LHS of (3.3.44) by:

$$\sum_{\mu_\nu \geq \ell_\nu} 2^{-(\mu_\nu - \ell_\nu)} \left\| \chi_{\ell_\nu} B_i^\nu \right\|_{L_x^1} \left\| \chi_{\mu_\nu} B_i^\nu \right\|_{L_x^1},$$

as required.

Given (3.3.44), we remark that summing one of the factors we get a universal bound. This follows from the global energy control (3.3.27) since, by the finitely overlapping supports of $\chi_{\ell_\nu} B_i^\nu$:

$$\sum_{\ell_\nu} \left\| \chi_{\ell_\nu} B_i^\nu \right\|_{L_x^1} \lesssim \left\| B_i^\nu \right\|_{L_x^1},$$

and in fact we have the stronger control:

$$\sum_{k \in \mathbb{Z}} \left\| P_k B_1^\nu \right\|_{L_x^1} + \left\| \left(\sum_{k \in \mathbb{Z}} |P_k B_2^\nu|^2 \right)^{\frac{1}{2}} \right\|_{L_x^1} \lesssim \mathcal{E}, \tag{3.3.45}$$

where for the former we have:

$$\|P_k B_1^\nu\|_{L_x^1} \lesssim \sum_{k_1,k_2 \geq k+5:|k_1-k_2| \leq O(1)} 2^{-(k_2-k)} \left\|P_{k_1} \Omega_{x,\beta}^\nu\right\|_{L_x^2} \left\|\nabla_x \phi_\nu^{k_2}\right\|_{L_x^2},$$

applying initially the finite band property (2.1.3), and then once again for $\phi_\nu^{k_2}$, and this can be summed over $k \in \mathbb{Z}$ using discrete Cauchy-Schwarz in $k_1 = k_2 + O(1)$. Whereas for the latter, we note that by the Littlewood-Paley trichotomy:

$$\begin{aligned}
P_k(\Omega_{x,\beta}^\nu \cdot \nabla_x \phi_\nu^{\leq k+10}) = P_k[&P_{\leq k-7}(\Omega_{x,\beta}^\nu) \cdot \nabla_x \phi_\nu^{k-3 \leq \cdot \leq k+3} \\
+ &P_{k-3 \leq \cdot \leq k+3}(\Omega_{x,\beta}^\nu) \cdot \nabla_x \phi_\nu^{\leq k-7} \\
+ &\sum_{k_1,k_2 \sim k} P_{k_1}(\Omega_{x,\beta}^\nu) \cdot \nabla_x \phi_\nu^{k_2}],
\end{aligned}$$

and so the first two terms correspond to paraproducts, already localized to $|\xi| \sim 2^k$, and therefore their sum in $k \in \mathbb{Z}$ lies in the homogeneous Hardy space $\dot{F}_2^{0,1}$ with bound $O(\mathcal{E})$, and for the last term the stronger estimate in $\dot{B}_1^{0,1}$ with bound $O(\mathcal{E})$ as for B_1^ν holds, since the sum under P_k is finite and we can apply the discrete Cauchy-Schwarz inequality.

Hence, rescaling by 2^{ℓ_ν} and setting $B_i^{\nu,\ell_\nu}(\cdot) = 2^{2\ell_\nu} B_i^\nu(2^{\ell_\nu}\cdot)$, to obtain decay for (3.3.44) it suffices to prove:

$$\sup_{n_\nu \leq \ell_\nu \leq N_\nu} \left\|\bar{m}_0 B_i^{\nu,\ell_\nu}\right\|_{L_x^1} \leq o(\mathcal{E}). \tag{3.3.46}$$

This is a direct manifestation of the perturbative nature of quadratic non-linearities on neck regions, thanks to the local energy decay (3.3.33). In our case, the argument is however slightly more involved because our product structure is non-local. This represents however a minor technicality only, and we shall treat this analogously to the previous instances of physical support leakage.

Let us introduce two auxiliary parameters. Setting $\Omega_{x,\beta}^{\nu,\ell_\nu}(\cdot) := 2^{\ell_\nu}\Omega_{x,\beta}^\nu(2^{\ell_\nu}\cdot)$, by the local energy estimate (3.3.33), we can find sequences $\kappa_\nu \to +\infty$ and $\tilde{\varepsilon}_\nu \downarrow 0$ such that:

$$\left\|\bar{m}_{-10 \leq \cdot \leq \kappa_\nu} \Omega_{x,\beta}^{\nu,\ell_\nu}\right\|_{L_x^2} \leq \tilde{\varepsilon}_\nu,$$

where we use the convention $\bar{m}_{k_1 \leq \cdot \leq k_2} := \bar{m}_{\leq k_2} - \bar{m}_{\leq k_1-1}$, and similarly for $\bar{m}_{\geq k_1} := 1 - \bar{m}_{\leq k_1-1}$. Let us first treat the annulus determined so, and then the outer and inner regions separately.

For the annulus we can discard the cut-off $\bar{m}_0$. Regarding B_1^{ν,ℓ_ν}, we have:

$$
\sum_{k\in\mathbb{Z}}\left\|P_k\nabla_x\cdot(\bar{m}_{-10\leq\cdot\leq\kappa_\nu}\Omega_{x,\beta}^{\nu,\ell_\nu}\phi_{\nu,\ell_\nu}^{>k+10})\right\|_{L_x^1}
$$
$$
\lesssim\sum_{k\in\mathbb{Z}}\sum_{k_1,k_2\geq k+5:|k_1-k_2|\leq O(1)}2^{-(k_1-k)}\left\|P_{k_1}(\bar{m}_{-10\leq\cdot\leq\kappa_\nu}\Omega_{x,\beta}^{\nu,\ell_\nu})\right\|_{L_x^2}\left\|\nabla_x\phi_\nu^{k_2}\right\|_{L_x^2},
$$

where we have used the finite band property (2.1.3) as usual, and we control this by $O(\tilde{\varepsilon}_\nu\mathcal{E}^{\frac{1}{2}})$ relying on the discrete Cauchy-Schwarz and $k_1=k_2+O(1)$, which is acceptable for (3.3.46). For B_2^{ν,ℓ_ν}, we use Littlewood-Paley trichotomy as previously to get:

$$
\left\|\sum_{k\in\mathbb{Z}}P_k(\bar{m}_{-10\leq\cdot\leq\kappa_\nu}\Omega_{x,\beta}^{\nu,\ell_\nu}\cdot\nabla_x\phi_{\nu,\ell_\nu}^{\leq k+10})\right\|_{L_x^1}
$$
$$
\lesssim\left\|\sup_{k_1\in\mathbb{Z}}|P_{\leq k_1-7}(\bar{m}_{-10\leq\cdot\leq\kappa_\nu}\Omega_{x,\beta}^\nu)|\right\|_{L_x^2}\cdot\left\|(\sum_{k_2\in\mathbb{Z}}|\nabla_x\phi_{\nu,\ell_\nu}^{k_2-3\leq\cdot\leq k_2+3}|^2)^{\frac{1}{2}}\right\|_{L_x^2}
$$
$$
+\left\|(\sum_{k_1\in\mathbb{Z}}|P_{k_1-3\leq\cdot\leq k_1+3}(\bar{m}_{-10\leq\cdot\leq\kappa_\nu}\Omega_{x,\beta}^\nu)|^2)^{\frac{1}{2}}\right\|_{L_x^2}\cdot\left\|\sup_{k_2\in\mathbb{Z}}|\nabla_x\phi_{\nu,\ell_\nu}^{\leq k_2-7}|\right\|_{L_x^2}
$$
$$
+\sum_{k\in\mathbb{Z}}\sum_{k_1,k_2\sim k}\left\|P_{k_1}(\bar{m}_{-10\leq\cdot\leq\kappa_\nu}\Omega_{x,\beta}^\nu)\right\|_{L_x^2}\cdot\left\|\nabla_x\phi_{\nu,\ell_\nu}^{k_2}\right\|_{L_x^2},
$$

and relying on the Littlewood-Paley square function estimate for the first two terms, and simply the discrete Cauchy-Schwarz for the last, we can bound the above by $O(\tilde{\varepsilon}_\nu\mathcal{E}^{\frac{1}{2}})$ again. Therefore this is a permissible contribution to (3.3.46).

Now we treat the error terms. First, let us consider the outer region defined by the cut-off $\bar{m}_{>\kappa_\nu}$. Writing:

$$
\left\|\bar{m}_0\sum_{k\in\mathbb{Z}}P_kB\right\|_{L_x^1}\lesssim\|\bar{m}_0\|_{L_x^1}\sum_{k\in\mathbb{Z}}\|P_kB\|_{L_x^\infty(\{2^{-1}\leq|x|\leq2\})}, \tag{3.3.47}
$$

we proceed, first for:
$$
B:=\sum_{k\in\mathbb{Z}}P_k\nabla_x\cdot(\bar{m}_{>\kappa_\nu}\Omega_{x,\beta}^{\nu,\ell_\nu}\phi_{\nu,\ell_\nu}^{>k+10}),
$$

by considering the convolution kernel for the Fourier multiplier $\nabla_xP_kP_{k'}$, with $k=k'+O(1)$, which gives:

$$
\|P_kB\|_{L_x^\infty(\{2^{-1}\leq|x|\leq2\})}\lesssim_N\frac{2^{3k}}{(1+2^k2^{\kappa_\nu})^N}\left\|\bar{m}_{>\kappa_\nu}\Omega_{x,\beta}^{\nu,\ell_\nu}\phi_{\nu,\ell_\nu}^{>k'+10}\right\|_{L_x^1},
$$

for any positive integer $N \in \mathbb{N}$, bearing in mind the physical support of $\bar{m}_{>\kappa_\nu} \Omega_{x,\beta}^{\nu,\ell_\nu} \phi_{\nu,\ell_\nu}^{>k+10}$. Using this estimate, for high frequency scales, we choose $N = 3$, getting the following bound for the sum in $k \geq 0$ from (3.3.47) :

$$2^{-3\kappa_\nu} \sum_{k \geq 0} 2^{-k} \left\| \bar{m}_{>\kappa_\nu} \Omega_{x,\beta}^{\nu,\ell_\nu} \right\|_{L_x^2} \sum_{k_1 > k+10} 2^{-(k_1-k)} \left\| \nabla_x \phi_{\nu,\ell_\nu}^{k_1} \right\|_{L_x^2},$$

by the finite band property (2.1.3) for ϕ_{ν,ℓ_ν}. This is immediately seen to be $o(\mathcal{E})$ as $\kappa_\nu \to +\infty$, hence this contribution is acceptable. For the low frequency scales, if we set $N = 1$ above, we have for the sum over $k < 0$ in (3.3.47):

$$2^{-\kappa_\nu} \sum_{k < 0} 2^k \left\| \bar{m}_{>\kappa_\nu} \Omega_{x,\beta}^{\nu,\ell_\nu} \right\|_{L_x^2} \sum_{k_1 > k+10} 2^{-(k_1-k)} \left\| \nabla_x \phi_{\nu,\ell_\nu}^{k_1} \right\|_{L_x^2} \lesssim o(\mathcal{E}),$$

as desired, so the contribution of the outer region is controlled for B_1^{ν,ℓ_ν}. Regarding B_2^{ν,ℓ_ν}, we have to control (3.3.47) with:

$$B := \sum_{k \in \mathbb{Z}} P_k \big(\bar{m}_{>\kappa_\nu} \Omega_{x,\beta}^{\nu,\ell_\nu} \cdot \nabla_x \phi_{\nu,\ell_\nu}^{\leq k+10} \big).$$

Proceeding similarly to the above, we look at the convolution kernel of $P_k P_{k'}$, with $k = k' + O(1)$, and given the spatial support of $\bar{m}_{>\kappa_\nu} \Omega_{x,\beta}^{\nu,\ell_\nu} \phi_{\nu,\ell_\nu}^{\leq k+10}$, we get the analogous estimate for $N \in \mathbb{Z}$:

$$\left\| P_k B \right\|_{L_x^\infty(\{2^{-1} \leq |x| \leq 2\})} \lesssim_N \frac{2^{2k}}{(1 + 2^k 2^{\kappa_\nu})^N} \left\| \bar{m}_{>\kappa_\nu} \Omega_{x,\beta}^{\nu,\ell_\nu} \cdot \nabla_x \phi_{\nu,\ell_\nu}^{\leq k'+10} \right\|_{L_x^1},$$

so that choosing $N = 3$ when $k \geq 0$, and $N = 1$ if $k < 0$ as previously, yields the control for (3.3.47):

$$2^{-\kappa_\nu} \Big(\sum_{k \in \mathbb{Z}} 2^{-|k|} \Big) \left\| \bar{m}_{>\kappa_\nu} \Omega_{x,\beta}^{\nu,\ell_\nu} \right\|_{L_x^2} \left\| \nabla_x \phi_{\nu,\ell_\nu} \right\|_{L_x^2} \lesssim o(\mathcal{E}),$$

as desired, and this completes the treatment of the contribution to (3.3.46) of the outer region.

Finally, we need to study the contribution of the interior region defined by the support of $\bar{m}_{<-10}$, that we note being at a definite amount of distance from the support of $\bar{m}_0$. First, we remark that we have:

$$\left\| \bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \right\|_{H_x^{-1}} \longrightarrow 0, \tag{3.3.48}$$

and to see this, we start by getting an extension φ_ν of $\phi_{\nu,\ell_\nu}|_{B_1}$, equal to a suitably chosen constant $c = c(\{\phi_{\nu,\ell_\nu}\}_{\nu\in\mathbb{N}})$, such that by the construction of the sequence of wave maps and the covering in Lemma 3.3.1, we have $\varphi_\nu^c := \varphi_\nu - c$ vanishing strongly in supercritical spaces:

$$\|\varphi_\nu^c\|_{H_x^s} \longrightarrow 0, \quad s < 1. \tag{3.3.49}$$

To establish (3.3.48) it is enough to consider $\widetilde{\varphi}_\nu \nabla_x \varphi_\nu$, where $\widetilde{\varphi}_\nu := \bar{m}_{<-10}\varphi_\nu$. For low frequencies:

$$\|P_{\leq 0}(\widetilde{\varphi}_\nu \nabla_x \varphi_\nu)\|_{L_x^2} \lesssim \|\widetilde{\varphi}_\nu\|_{L_x^\infty} \left\|P_{\leq O(1)}\varphi_\nu^c\right\|_{L_x^2}$$
$$+ \sum_{k_1,k_2\geq O(1):|k_1-k_2|\leq O(1)} \|\nabla_x P_{k_1}\widetilde{\varphi}_\nu\|_{L_x^2} \|P_{k_2}\varphi_\nu^c\|_{L_x^2},$$

where for the first term we have used (2.1.2) to discard ∇_x, and for the second we passed initially to L_x^1 applying (2.1.4), and then transferred ∇_x from φ_ν^c to $\widetilde{\varphi}_\nu$ via (2.1.3). Both items are acceptable by (3.3.49). For high frequencies, we apply precisely the same argument, but with a slightly more refined Littlewood-Paley trichotomy decomposition:

$$2^{-k}\|P_k(\widetilde{\varphi}_\nu \nabla_x \varphi_\nu)\|_{L_x^2} \lesssim \|P_{\leq k-7}\widetilde{\varphi}_\nu\|_{L_x^\infty} \|P_{k-3\leq\cdot\leq k+3}\varphi_\nu^c\|_{L_x^2}$$
$$+ \|P_{k-3\leq\cdot\leq k+3}\nabla_x\widetilde{\varphi}_\nu\|_{L_x^2} \|P_{\leq k-7}\varphi_\nu^c\|_{L_x^2}$$
$$+ 2^{-\frac{k}{2}} \sum_{k_1,k_2\geq k-6:|k_1-k_2|\leq O(1)} \|\nabla_x P_{k_1}\widetilde{\varphi}_\nu\|_{L_x^2} 2^{\frac{k_2}{2}} \|P_{k_2}\varphi_\nu^c\|_{L_x^2},$$

where for the first term we applied (2.1.2) and for the other two we passed first to L_x^1 via (2.1.3), then used Cauchy-Schwarz, from where for the second term we used (2.1.2) for $P_{\leq k-7}\varphi_\nu^c$ and (2.1.3) for $\widetilde{\varphi}_\nu$ transferring ∇_x from one to the other, whereas for the third term this transfer of ∇_x happened at once via (2.1.3) since $k_1 = k_2 + O(1)$, and then multiplied $P_{k_2}\varphi_\nu^c$ simply by $2^{-k_2/2}2^{k_2/2}$ which led to the exponential gain $2^{-k/2}$ in front of the sum since $k_2 \geq k + O(1)$. Square-summing the above estimate over $k > 0$, and applying discrete Cauchy-Schwarz for the third item, gives an acceptable bound by (3.3.49), therefore we have claim (3.3.48).

With this understood, we can control the contribution of the inner region to (3.3.46)

for the low frequencies. Given any positive integer $K > 0$, we have regarding B_1^{ν,ℓ_ν}:

$$\sum_{k \leq K} \left\| P_k \nabla_x \cdot (\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \phi_{\nu,\ell_\nu}^{>k+10}) \right\|_{L_x^1}$$

$$\lesssim \sum_{k \leq K} 2^k \sum_{k_1,k_2 \geq k+5: |k_1-k_2| \leq O(1)} 2^{-k_2} \left\| P_{k_1}(\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu}) \right\|_{L_x^2} \left\| \nabla_x \phi_\nu^{k_2} \right\|_{L_x^2}$$

$$\lesssim \sum_{k \leq O(1)} \left(\sum_{k+5 \leq k_1,k_2 \leq O(1): |k_1-k_2| \leq O(1)} 2^{-(k_2-k)} \left\| P_{k_1}(\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu}) \right\|_{L_x^2} \left\| \nabla_x \phi_\nu^{k_2} \right\|_{L_x^2} \right)$$

$$+ 2^K \sum_{k_1,k_2 \geq O(1): |k_1-k_2| \leq O(1)} 2^{-k_1} \left\| P_{k_1}(\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu}) \right\|_{L_x^2} \left\| \nabla_x \phi_\nu^{k_2} \right\|_{L_x^2},$$

which is $o(\mathcal{E})$ for the first term and $o_K(\mathcal{E})$ for the second by (3.3.48). Analogously, looking at B_2^{ν,ℓ_ν} we get:

$$\left\| \sum_{k \leq K} P_k(\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \cdot \nabla_x \phi_{\nu,\ell_\nu}^{\leq k+10}) \right\|_{L_x^1}$$

$$\lesssim \left\| \sum_{k \leq K} P_k[P_{\leq K+O(1)}(\bar{m}_{<-10} \Omega_{x,\beta}^\nu) \cdot \nabla_x \phi_{\nu,\ell_\nu}^{k-3 \leq \cdot \leq k+3}] \right\|_{L_x^1}$$

$$+ \left\| \sum_{k \leq K} P_k[P_{k_1-3 \leq \cdot \leq k_1+3}(\bar{m}_{<-10} \Omega_{x,\beta}^\nu) \cdot \nabla_x \phi_{\nu,\ell_\nu}^{\leq k_2-7}] \right\|_{L_x^1}$$

$$\lesssim \left\| P_{\leq K+O(1)}(\bar{m}_{<-10} \Omega_{x,\beta}^\nu) \right\|_{L_x^2} \cdot \left\| \left(\sum_{k_2 \leq K} |\nabla_x \phi_{\nu,\ell_\nu}^{k_2-3 \leq \cdot \leq k_2+3}| \right)^{\frac{1}{2}} \right\|_{L_x^2}$$

$$+ \left\| \left(\sum_{k_1 \leq K} |P_{k_1-3 \leq \cdot \leq k_1+3}(\bar{m}_{<-10} \Omega_{x,\beta}^\nu)|^2 \right)^{\frac{1}{2}} \right\|_{L_x^2} \cdot \left\| \sup_{k_2 \leq K} |\nabla_x \phi_{\nu,\ell_\nu}^{\leq k_2-7}| \right\|_{L_x^2},$$

and this is again controlled by $o_K(\mathcal{E})$ via (3.3.48). Therefore, for both contributions, we can choose a sequence of integers $K_\nu \to +\infty$, together with decaying constants $\varsigma_\nu \downarrow 0$, such that:

$$\sum_{k \leq K_\nu} \left\| P_k \nabla_x \cdot (\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \phi_{\nu,\ell_\nu}^{>k+10}) \right\|_{L_x^1} + \left\| \sum_{k \leq K_\nu} P_k(\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \cdot \nabla_x \phi_{\nu,\ell_\nu}^{\leq k+10}) \right\|_{L_x^1} \leq \varsigma_\nu,$$

and this yields the decay of slowly growing frequencies for the inner region, as desired. Note that the cut-off $\bar{m}_0$ has not played any role in the above argument. However,

for the high frequencies $k > K_\nu$, having $\bar{m}_0$ will be crucial as we are going to pass by (3.3.47) as before, first with:

$$B := \sum_{k > K_\nu} P_k \nabla_x \cdot (\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \phi_{\nu,\ell_\nu}^{>k+10}).$$

Considering the convolution kernel for $\nabla_x P_k P_{k'}$, with $k = k' + O(1)$, as previously, we estimate:

$$\|P_k B\|_{L_x^\infty(\{2^{-1} \le |x| \le 2\})} \lesssim \frac{2^{3k}}{(1+2^k)^3} \left\| \bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \phi_{\nu,\ell_\nu}^{>k'+10} \right\|_{L_x^1},$$

noting the fixed positive distance of the physical support of $\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \phi_{\nu,\ell_\nu}^{>k+10}$ to the annulus $\{2^{-1} \le |x| \le 2\}$. Using this, we can bound (3.3.47) in this case by:

$$\sum_{k > K_\nu} 2^{-k} \left\| \bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \right\|_{L_x^2} \sum_{k_1 > k+10} 2^{-(k_1-k)} \left\| \nabla_x \phi_{\nu,\ell_\nu}^{k_1} \right\|_{L_x^2} \lesssim 2^{-K_\nu} \mathcal{E},$$

which is certainly acceptable, given that $K_\nu \to +\infty$. Finally, the last contribution to treat is when:

$$B := \sum_{k > K_\nu} P_k (\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \cdot \nabla_x \phi_{\nu,\ell_\nu}^{\le k+10}),$$

in (3.3.47), and here we proceed in complete analogy to the above, getting the following estimate:

$$\|P_k B\|_{L_x^\infty(\{2^{-1} \le |x| \le 2\})} \lesssim \frac{2^{2k}}{(1+2^k)^3} \left\| \bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \cdot \nabla_x \phi_{\nu,\ell_\nu}^{\le k'+10} \right\|_{L_x^1},$$

by looking at the convolution kernel of $P_k P_{k'}$, with $k = k' + O(1)$, and the location of spatial support of $\bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \cdot \nabla_x \phi_{\nu,\ell_\nu}^{\le k+10}$ with respect to the annulus $\{2^{-1} \le |x| \le 2\}$. This in turn, yields the following control for (3.3.47):

$$\sum_{k > K_\nu} 2^{-k} \left\| \bar{m}_{<-10} \Omega_{x,\beta}^{\nu,\ell_\nu} \right\|_{L_x^2} \|\nabla_x \phi_{\nu,\ell_\nu}\|_{L_x^2} \lesssim 2^{-K_\nu} \mathcal{E},$$

which, as noted above, is permissible. That concludes the treatment of the contribution of the inner region, and therefore we have obtained claim (3.3.46).

In the end, going back to the physical Littlewood-Paley decomposition (3.3.36) and expressing the time derivative ∂_t via X and ∂_{x_1} using expression (2.3.2), we have for

any $k \in \mathbb{Z}$:

$$\left\|P_k[(\bar{m}_{\leq N_\nu} - \bar{m}_{\leq n_\nu - 1})\nabla_{t,x}\phi_\nu](0)\right\|_{L_x^2} \lesssim \sum_{\ell \in \mathbb{Z}} 2^{-|k+\ell|}\varepsilon_\nu + \sigma_\nu + \iota_\nu + o(\mathcal{E}) \longrightarrow 0,$$

where the first sum arises from the low frequencies (3.3.39) and the regular part involving spatial derivatives falling on the cut-offs from (3.3.38) and (3.3.41), the second term comes from errors having good time-like control (3.3.42), the third arises from treating the higher-order time like derivative in (3.3.43), and finally the last term is due to the perturbative $\dot{B}_\infty^{-1,2}$ estimate of the non-linearity for the wave maps equation at high frequency (3.3.44), combined with (3.3.45) and (3.3.46).

Lemma 3.3.3 is proved. □

3.3.4   Completion of the proof

We are now at the concluding stage of the proof of Theorem 1.2.4, for which, going back to the weak bubble tree decomposition (3.3.13), we must show that the energy of the necks $\mathcal{N}_{i,\nu}$ is asymptotically vanishing as $\nu \to +\infty$. Recall that those are provided with corresponding neck domains, that is the conformally degeneration annuli from (3.3.14), so that setting:

$$\phi_{\nu,x_{i,\nu}^k}(t,x) := \phi_{i,\nu}(\lambda_{\min,\nu}t, x_{i,\nu}^k + \lambda_{\min,\nu}x),$$

we can apply Lemma 3.3.3, by (3.3.15) and (3.3.11), to write:

$$\nabla_{t,x}\phi_{\nu,x_{i,\nu}^k} = \Upsilon_{\nu,x_{i,\nu}^k} \quad \text{on} \ [-1,1] \times (B_{\lambda_{\min,\nu}^{-1}R_{i,\nu}^k} \setminus B_{\lambda_{\min,\nu}^{-1}r_{i,\nu}^k}),$$

where $\Upsilon_{\nu,x_{i,\nu}^k}$ is supported on $[-1,1] \times (B_{2\lambda_{\min,\nu}^{-1}R_{i,\nu}^k} \setminus B_{2^{-1}\lambda_{\min,\nu}^{-1}r_{i,\nu}^k})$ with

$$\left\|\Upsilon_{\nu,x_{i,\nu}^k}\right\|_{L_t^\infty(L_x^2)[-1,1]} \lesssim 1,$$

and satisfying the decay:

$$\sup_{k \in \mathbb{Z}}\left\|P_k\Upsilon_{\nu,x_{i,\nu}^k}(0)\right\|_{L_x^2} \longrightarrow 0.$$

Recalling (3.3.11), we also have:

$$\left\|\Theta_{\nu,x_{i,\nu}^k}(0)\right\|_{L_x^2} \longrightarrow 0,$$

where $\Theta_{\nu,x_{i,\nu}^k}(t,x) := \lambda_{\min,\nu}\Theta_{i,\nu}(\lambda_{\min,\nu}t, x_{i,\nu}^k + \lambda_{\min,\nu}x)$, together with:

$$\sum_{k\in\mathbb{Z}} \left\| P_k\Xi_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2} \lesssim 1,$$

where $\Xi_{\nu,x_{i,\nu}^k}(t,x) := \lambda_{\min,\nu}\Xi_{i,\nu}(\lambda_{\min,\nu}t, x_{i,\nu}^k + \lambda_{\min,\nu}x)$.

From there, we can estimate the energy at time $t = 0$ on a neck region by:

$$\left\| \nabla_{t,x}\phi_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2(B_{\lambda_{\min,\nu}^{-1}R_{i,\nu}^k}\setminus B_{\lambda_{\min,\nu}^{-1}r_{i,\nu}^k})}^2 \lesssim \left| \int_{\mathbb{R}^2} \Upsilon_{\nu,x_{i,\nu}^k}(0)\Xi_{\nu,x_{i,\nu}^k}(0)dx \right|$$

$$+ \left| \int_{\mathbb{R}^2} \Upsilon_{\nu,x_{i,\nu}^k}(0)\Theta_{\nu,x_{i,\nu}^k}(0)dx \right| + o(1),$$

which we bound by:

$$(\sup_{k\in\mathbb{Z}} \left\| P_k\Upsilon_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2}) \sum_{k\in\mathbb{Z}} \left\| P_k\Xi_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2} + \left\| \Upsilon_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2} \left\| \Theta_{\nu,x_{i,\nu}^k}(0) \right\|_{L_x^2} + o(1),$$

and by the previous estimates this tends to 0 as $\nu \to +\infty$. Theorem 1.2.4 is proved.

Chapter 4

Conclusion: Future directions

We close the thesis with some thoughts and reflections about some of the challenges lying on the road towards the full Soliton Resolution Conjecture, that the author is most interested in. For the sake of simplicity, we will restrict our discussion to the case of finite time blow up.

The most immediate question that arises, of course, is regarding the phenomena of null concentration. In fact, taking the comparison to the parabolic counterpart of the theory, the harmonic map heat flow treated by Ding and Tian [6], the analogue of our results would be weaker as we move away from the equivariant set up. Nevertheless, the techniques to deal with this concentration and the whole strategy are actually already there. Let us say a few words about this now.

In fact, the analogue of our results would also be weaker than those of Duyckaerts, Jia, Kenig and Merle [7], who treated the critical focusing wave equation $\Box u = u^5$ on $\mathbb{R}^{3+1}$. In this work, null concentration was proved to vanish in the energy space. The argument advanced by the authors there seems to give not only hope, but also the ingredients.

They argued by contradiction. Assuming that some energy was escaping into the null boundary, away from the solitons, they could construct a whole "channel of energy", going backward along the null boundary up to the initial data. The latter of course being fixed, concentration of energy was excluded, and that led to the desired contradiction. The channel of energy itself is in reality only proved to exist for free waves with initial data well concentrated on a boundary of a disc, with the angular and the L null energy sufficiently small. This is precisely the initial data we construct in Theorem 1.2.3. After the authors obtain decay in a suitable dispersive norm for the part of the solution that approaches the boundary of light cone, analogously to what we do in Section 3.2, they

get the latter arbitrarily close to the free wave in the energy space, and this then leads to the desired construction.

As one guesses now, this is precisely what we were trying to mimic with our decomposition (1.2.4). However, the perturbation theory for the focusing wave equation is considerably easier than for wave maps (partly because of the bi-linear structure in the gradient in the non-linearity, and partly because of the gauge fixing issues). We hit therefore two important obstacles. The first one is that we cannot work with a free wave as our basis, but with a gauge co-variant one. So in particular, the channels of energy (or at least some approximated version of them) must be extended to that case. In that regard, having the frequency support degenerating for the anti-symmetric matrices should help. And secondly, one needs to be able to extend the solution well beyond the interval we are working on currently, all the way up to the initial data. That is our challenge for now!

There are a couple of other problems related to SRC, that the author would love to carry on thinking about. One concerns building examples of winding behavior and non-uniqueness of bubbles, as was beautifully achieved by Peter Topping in [36] for the harmonic map heat flow. This is an ongoing project which, however, requires a lot of background, as developed in the ground breaking work of Raphaël and Rodnianski [22], going well beyond techniques we have mastered so far.

Finally, there is the mysterious problem of obtaining compensation type estimates for wave maps, and more basically for harmonic maps, into general closed Riemannian manifolds. We have in fact spent the largest part of our DPhil thinking about that... but maybe that's a good point to stop the story there, as far as this thesis is concerned.

Bibliography

[1] D. Christodoulou and A. S. Tahvildar-Zadeh. On the regularity of spherically symmetric wave maps. *Comm. Pure Appl. Math.*, 46, no. 7:1041–1091, 1993.

[2] R. Côte. On the Soliton Resolution for Equivariant Wave Maps to the Sphere. *Comm. Pure Appl. Math.*, 68, no. 11:1946–2004, 2015.

[3] R. Côte. Corrigendum: On the Soliton Resolution for Equivariant Wave Maps to the Sphere. *Comm. Pure Appl. Math.*, 69, no. 4:609–612, 2016.

[4] R. Côte, C. Kenig, A. Lawrie, and W. Schlag. Characterization of large energy solutions of the equivariant wave map problem: I. *Amer. J. Math.*, 137, no.1:139–207 (New version on arXiv:1209.3682v2), 2015.

[5] R. Côte, C. Kenig, A. Lawrie, and W. Schlag. Characterization of large energy solutions of the equivariant wave map problem: II. *Amer. J. Math.*, 137, no.1:209–250 (New version on arXiv:1209.3684v2), 2015.

[6] W. Ding and G. Tian. Energy identity for a class of approximate harmonic maps from surfaces. *Comm. Anal. Geom.*, no. 3:543–554, 1995.

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

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

[9] T. Duyckaerts, C. Kenig, and F. Merle. Scattering profile for global solutions of the energy-critical wave equation. arXiv:1601.02107, 2016.

[10] A. Freire, S. Müller, and M. Struwe. Weak convergence of wave maps from (1 + 2)-dimensional Minkowski space to Riemannian manifolds. *Invent. Math.*, 130, no.3:589–617, 1997.

[11] L. Grafakos. *Modern Fourier Analysis*. Graduate Texts in Mathematics, vol. 250, 3rd edition, Springer, New York, 2014.

[12] R. Grinis. Quantization of Time-Like Energy for Wave Maps into Spheres. *Comm. Math. Phys.*, doi:10.1007/s00220-016-2766-9, arXiv:1601.07452, 2016.

[13] F. Hélein. *Harmonic maps, conservation laws and moving frames.* Cambridge Tracts in Mathematics, 150, CUP, Cambridge, 2002.

[14] H. Jia. Soliton resolution along a sequence of times with dispersive error for type II singular solutions to focusing energy critical wave equations. arXiv:1510.00075, 2015.

[15] H. Jia and C. Kenig. Asymptotic decomposition for semilinear wave and equivariant wave map equations. arXiv:1503.06715, 2015.

[16] J. Krieger. Global regularity of wave maps from $\mathbb{R}^{2+1}$ to $\mathbb{H}^2$. Small energy. *Commun. Math. Phys.*, 250, no.3:507–580, 2004.

[17] J. Krieger and W. Schlag. *Concentration Compactness for Critical Wave Maps.* EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2012.

[18] J. Krieger, W. Schlag, and D. Tataru. Renormalization and blow up for charge one equivariant critical wave maps. *Invent. Math.*, 171, no.3:543–615, 2008.

[19] P. Laurain and T. Rivière. Angular energy quantization for linear elliptic systems with anti-symmetric potentials and applications. *Anal. PDE.*, 7, no.1:1–41, 2014.

[20] F. H. Lin. Gradient estimates and blow-up analysis for stationary harmonic maps. *Ann. of Math.*, (2) 149:785–829, 1999.

[21] F. H. Lin and T. Rivière. Energy quantization for harmonic maps. *Duke Math J.*, 111:177–193, 2002.

[22] P. Raphaël and I. Rodnianski. Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang-Mills problems. *Publ. Math. Inst. Hautes Etudes Sci.*, 155, no.1:1–122, 2012.

[23] T. Rivière. Conservation laws for conformally invariant variational problems. *Invent. Math.*, 168:1–22, 2007.

[24] I. Rodnianski and J. Sterbenz. On the formation of singularities in the critical O(3) σ-model. *Ann. of Math.*, 172, no.1:187–242, 2010.

[25] J. Sacks and K. Uhlenbeck. The existence of minimal immersions of 2-spheres. *Annals of Math.*, 113:1–24, 1981.

[26] J. Shatah and M. Struwe. *Geometric Wave Equations.* Courant Lecture Notes in Mathematics 2, 1998.

[27] J. Shatah and A. S. Tahvildar-Zadeh. Regularity of harmonic maps from the minkowski space into rotationally symmetric manifolds. *Comm. Pure Appl. Math.*, 45, no. 8:947–971, 1992.

[28] L. Simon. Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems. *Ann. of Math.*, 118, no.3:525–571, 1983.

[29] J. Sterbenz and D. Tataru. Energy dispersed large data wave maps in 2 + 1 dimensions. *Comm. Math. Phys.*, 298, no. 1:139–230, 2010.

[30] J. Sterbenz and D. Tataru. Regularity of wave-maps in dimension 2 + 1. *Comm. Math. Phys.*, 298, no. 1:231–264, 2010.

[31] T. Tao. Global regularity of wave maps I. Small critical Sobolev norm in high dimension. *Internat. Math. Res. Notices*, no. 6:299–328, 2001.

[32] T. Tao. Global regularity of wave maps II. Small energy in two dimensions. *Comm. Math. Phys.*, 224, no. 2:443–544, 2001.

[33] T. Tao. Global regularity of wave maps III-VII. arXiv:0805.4666, arXiv:0806.3592, arXiv:0808.0368, arXiv:0906.2833, arXiv:0908.0776, 2009.

[34] D. Tataru. Rough solutions for the wave maps equation. *Amer. J. Math.*, 127, no. 2:293–377, 2005.

[35] M. Taylor. *Tools for PDE. Pseudodifferential operators, paradifferential operators, and layer potentials.* Mathematical Surveys and Monographs, 81, A.M.S., Providence, RI, 2000.

[36] P. Topping. Winding behaviour of finite-time singularities of the harmonic map heat flow. *Math. Zeit.*, 247:279–302, 2004.