

# GLOBAL REGULARITY AND SINGULARITY DEVELOPMENT FOR WAVE MAPS

J. KRIEGER

## 1. THE WAVE MAPS PROBLEM

Let  $(M, g)$  be a Riemannian manifold, and denote the standard Minkowski space by  $\mathbb{R}^{n+1}$ ,  $n \geq 1$ , equipped with the Minkowski metric  $m_{\alpha\beta} = m^{\alpha\beta} = \text{diag}(-1, 1, \dots, 1)$ . A map  $u : \mathbb{R}^{n+1} \rightarrow (M, g)$  is called a *wave map* provided it is formally critical with respect to the Lagrangian action functional<sup>1</sup>

$$\mathcal{L}(u) := \int_{\mathbb{R}^{n+1}} \langle \partial_\alpha u, \partial^\alpha u \rangle_g d\sigma, \quad \partial^\alpha = m^{\alpha\beta} \partial_\beta$$

Here  $d\sigma = \prod_{\alpha=0}^n dx_\alpha$  denotes the standard volume element on  $\mathbb{R}^{n+1}$ . In local coordinates, one finds the equation

$$(1.1) \quad \square u^i + \Gamma_{jk}^i \partial_\alpha u^j \partial^\alpha u^k = 0, \quad \square = \partial_\alpha \partial^\alpha$$

We shall refer to this as the local coordinate formulation of the wave maps condition. Another possibility is to isometrically embed the target into a Euclidean space  $M \hookrightarrow \mathbb{R}^k$ , and then work in terms of the ambient coordinates. This is seen to lead to the condition

$$(1.2) \quad \square u \perp TM_u$$

which leads to

$$(1.3) \quad \square u^i + S_{jk}^i(u) \partial_\alpha u^j \partial^\alpha u^k = 0$$

in terms of the ambient coordinates, where  $S_{jk}^i$  is the 2nd fundamental form of the embedding  $M \hookrightarrow \mathbb{R}^k$ .

Finally, in case the target is parallelizable<sup>2</sup>, and assuming  $\{e_i\}_{i=1}^k$  is a global orthonormal frame, we can introduce the family of functions  $\phi_\alpha^i$  via the relations

$$u_*(\partial_\alpha) = \phi_\alpha^i e_i$$

Then one deduces a divergence-curl system for the quantities  $\phi_\alpha^i$  of the form

$$(1.4) \quad \partial_\alpha \phi_\beta^i - \partial_\beta \phi_\alpha^i = C_{jk}^i \phi_\alpha^j \phi_\beta^k$$

$$(1.5) \quad \partial_\alpha \phi^{i\alpha} = -\Gamma_{jk}^i \phi_\alpha^j \phi^{k\alpha}$$

Here we have the relations  $C_{jk}^i = \Gamma_{jk}^i - \Gamma_{kj}^i$ , while the  $\Gamma_{jk}^i$  are the Christoffel symbols with respect to the frame  $\{e_i\}$ , i. e.  $\nabla_{e_j} e_k = \Gamma_{jk}^i e_i$ . While the former relation (1.4) arises simply from the fact that the  $\phi_\alpha^i$  represent the pushforward of the  $\partial_\alpha$

---

<sup>1</sup>Throughout, the Einstein summation convention is in force

<sup>2</sup>In a lot of cases, one may reduce to this situation, for example by finding an isometric embedding of  $M$  into an  $\mathbb{R}^k$  equipped with a metric which renders the embedding totally geodesic; see [8]

under  $u$ , the latter relation (1.5) encodes the wave maps condition. We shall refer to this system as the *intrinsic formulation* of wave maps.

In some sense, the intrinsic formulation is the most natural one, as it exemplifies the *gauge invariance* inherent in the problem: choosing a gauge means fixing an orthonormal frame  $\{e_i\}_{i=1}^k$  for  $TM$ . Choosing this frame judiciously may improve the features of the equations. Indeed, this gauge freedom was exploited first for the elliptic analogue of wave maps, harmonic maps (for example [17]), and then in [11], and implicitly in [61], [62], which then led to explicit applications in [27], [42], [50], [33].

### Examples

- (i) let  $M = \mathbb{R}$ . A wave map  $u : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$  is a free wave  $\square u = 0$ .
- (ii) let  $M = S^{k-1} \subset \mathbb{R}^k$  equipped with the standard metric. Then the wave maps condition can be cast in terms of ambient coordinates as follows:

$$\square u = -uu^t_{,\alpha} u^{\alpha}, \quad u \in \mathbb{R}^k$$

Here  $u \in \mathbb{R}^k$  is a column vector, and  $u^t$  its transpose.

- (iii) Now let  $M = \mathbf{H}^2$ , the hyperbolic plane with metric  $dh = \frac{dx^2 + dy^2}{y^2}$ ,  $y > 0$ . Choose the frame  $e_1 = -y \frac{d}{dx}$ ,  $e_2 = -y \frac{d}{dy}$ . Then the wave maps condition, cast in terms of the associated family  $\phi_{\alpha}^i$ ,  $\alpha = 0, 1, 2$ ,  $i = 1, 2$ , becomes

$$(1.6) \quad \partial_{\beta} \phi_{\alpha}^1 - \partial_{\alpha} \phi_{\beta}^1 = \phi_{\alpha}^1 \phi_{\beta}^2 - \phi_{\alpha}^2 \phi_{\beta}^1, \quad \partial_{\beta} \phi_{\alpha}^2 - \partial_{\alpha} \phi_{\beta}^2 = 0$$

$$(1.7) \quad \partial_{\alpha} \phi^{1\alpha} = -\phi_{\alpha}^1 \phi^{2\alpha}, \quad \partial_{\alpha} \phi^{2\alpha} = \phi_{\alpha}^1 \phi^{1\alpha}$$

The basic question one would like to answer is the *Cauchy problem* for wave maps: given initial data  $(u, \partial_t u) : \mathbb{R}^n \rightarrow M \times TM$ , is there a global-in-time wave map extending them? If not, explain the breakdown.

As posed the question is too vague. Indeed, a first issue to answer is what are the minimal requirements on regularity of the data to even obtain unique local solutions. It is by now well understood that there is a satisfactory local well-posedness theory provided the data are of regularity  $(u, \partial_t u) \in H^s \times H^{s-1}$ ,  $s > \frac{n}{2}$ ,  $n \geq 2$ . See [24], [28]. Moreover, there are examples demonstrating that uniqueness of (distributional) solutions fails provided the data are of class  $H^s$ ,  $s < \frac{n}{2}$ , see [1]. On the positive side, global distributional solutions of energy class ( $H^1$ ) have been constructed in the case  $n = 2$  in [41].

Solutions of the optimal  $C^{\infty}$  smoothness are also called *classical wave maps*. We also observe here in passing that the Cauchy problem for the case  $n = 1$  is well understood (see [16], [60]) and globally well-posed, and so we omit this case from now on. Thus a more precise formulation of our main problem is as follows: *given initial data  $(u, \partial_t u) \in H^s \times H^{s-1}$ , for some  $s > \frac{n}{2}$ ,  $n \geq 2$ , decide if they can be globally extended to a wave map.*

In spite of a lot of recent progress, this question is largely open at this time.

### Motivation for studying wave maps

-(a) Particle physics: The wave maps problem appears in [12] with target  $SU(2) = S^3$  and domain  $\mathbb{R}^{3+1}$ , also referred to as nonlinear sigma model.

-(b) Einstein's equations: following Choquet-Bruhat and Moncrief [7], consider a space-time 4 manifold  $M = M_1 \times \mathbb{R}$ , where  $M_1$  is a principal  $U(1)$ -bundle over a surface  $\Sigma$ , with projection  $\pi : M \rightarrow \Sigma \times \mathbb{R}$ . Then consider metrics on  $M$  of the Kaluza-Klein form

$$g = \pi^*(\tilde{g}) + \pi^*(e^{2\gamma})\Theta^2$$

for some  $U(1)$ -connection 1-form  $\Theta = A_\alpha dx^\alpha$ ,  $A_\alpha \in u(1) = i\mathbb{R}$ . Choosing suitable local coordinates we may write  $\Theta = dx^3 + \sum_{\alpha=0}^2 A_\alpha dx^\alpha$ , where  $x_{1,2}$  are coordinates on  $\Sigma$ . Then the curvature is given by  $F := d\Theta = F_{\alpha\beta} dx^\alpha dx^\beta$ , and Einstein's equations may be seen to imply that defining<sup>3</sup>  $e^{3\gamma}(*F) = E$ , we have  $dE = 0$ . If for example  $\Sigma = \mathbb{R}^2$  topologically, we have  $E = d\omega$  for some function  $\omega$ . Then it may be seen that the map

$$(\gamma, \omega) : (\Sigma \times \mathbb{R}, \tilde{g}) \longrightarrow (\mathbf{H}^2, h)$$

is a wave map. The metric  $\tilde{g}$  in turn is driven by the wave map. This suggests as a model problem wave maps  $u : \mathbb{R}^{2+1} \longrightarrow \mathbf{H}^2$ . The case  $n = 2$  is also referred to as the *energy critical case*, as explained below.

(c) Wave maps are the natural hyperbolic analogues of the much studied *harmonic map heat flow*, which in local coordinates is described by

$$\partial_t u^i = \Delta u^i + \sum_{\alpha=1}^n \Gamma_{jk}^i(u) \partial_\alpha u^j \partial^\alpha u^k$$

## 2. WELL-POSEDNESS TYPOLOGY FOR NONLINEAR WAVE EQUATIONS

Consider a general problem of the form

$$(2.1) \quad \square u = N(u, \nabla u), \quad (u, \partial_t u)|_{t=0} = (u_0, u_1)$$

for some smooth  $N(., .)$ . Of course, wave maps in local coordinates falls into this category. Assume that the set of solutions  $u(t, x)$  is invariant under the scaling transformation  $u(t, x) \rightarrow \lambda^\alpha u(\lambda t, \lambda x)$ . For wave maps, for example, we have  $\alpha = 0$ . Then one introduces the *critical Sobolev index*  $s_c = \frac{n}{2} - \alpha$ . Observe that the norm

$$\|u_0\|_{\dot{H}^{s_c}} + \|u_1\|_{\dot{H}^{s_c-1}}$$

is left invariant under the re-scaling. We are interested in (i) *local well-posedness*, and, having established (i), (ii) *global well-posedness*. In particular, we shall use the following definitions:

**Definition 2.1.** (*strong local well-posedness*) We call the problem (2.1) *strongly locally well-posed* in the Sobolev space  $H^s(\mathbb{R}^n)$ , which we recall is the space of all tempered distributions  $f$  satisfying  $\int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 (1 + |\xi|^2)^s d\xi < \infty$ , provided the following holds: given data  $(u_0, u_1) \in H^s \times H^{s-1}$ , there is a time interval<sup>4</sup>  $[0, T]$ ,  $T = T(\|u_0\|_{H^s} + \|u_1\|_{H^{s-1}}) > 0$  and a unique distributional solution  $u(t, x) \in C([0, T], H^s) \cap C^1([0, T], H^{s-1})$  of (2.1) on  $[0, T] \times \mathbf{R}^n$ , which depends continuously on the data, and preserves higher regularity of the data.

<sup>3</sup>Here  $*F = \frac{1}{2} \epsilon_{\alpha\beta\gamma} F^{\beta\gamma} dx^\alpha$  for the antisymmetric symbol  $\epsilon_{\alpha\beta\gamma}$ .

<sup>4</sup>It is understood that  $T(\cdot)$  is a decreasing function of its argument.

**Definition 2.2.** (*weak global well-posedness for small data*) We call (2.1) *weakly globally well-posed for small data*  $(u_0, u_1)$  in  $\dot{H}^{s_c} \times \dot{H}^{s_c-1}$ , provided there exists some  $\epsilon > 0$  such that for all  $(u_0, u_1) \in H^s \times H^{s-1}$ ,  $s > s_c$  with  $\|u_0\|_{\dot{H}^{s_c}} + \|u_1\|_{\dot{H}^{s_c-1}} < \epsilon$ , there exists a unique global solution  $u(t, x) \in C((-\infty, \infty), \dot{H}^s) \cap C^1((-\infty, \infty), H^{s-1})$ .

The general method for establishing strong local well-posedness is by means of iteration in a suitable Banach space, while the global (weak) well-posedness at the critical Sobolev index follows usually by proving a priori bounds.

The significance of the critical Sobolev index becomes clear on account of the following

### Expected Optimal Well-posedness:

#### -Scaling constraint

-*Supercritical:* If  $s < s_c$ , one expects ill-posedness in  $H^s$ . Assume for example  $s_c > 0$ , and there exist data  $(u_0, u_1) \in H^s \times H^{s-1}$ ,  $s < s_c$ , such that the corresponding solution exists on a maximal interval  $[0, T)$ ,  $T < \infty$ . By finite propagation speed we may assume that they are compactly supported. The data  $(u_{0,i}, u_{1,i})$ , where  $(v_{0,i}, v_{1,i})(x) = (2^{\alpha i} v_0(2^i x), 2^{(\alpha+1)i} v_1(2^i x))$ ,  $i \geq 1$ , result in solutions with maximal interval of existence  $[0, 2^{-i} T)$ . Then letting  $(a_0, a_1) := \sum_{i \geq 1} T_i(u_{0,i}, u_{1,i})$  for a suitable translation operator  $T_i$ ,  $T_i(v(x)) := v(x + b_i)$ , chosen such that the corresponding solutions do not overlap on their interval of existence, we have

$$\left\| \sum_{i \geq 1} T_i(u_{0,i}, u_{1,i}) \right\|_{\dot{H}^s \times \dot{H}^{s-1}} \leq \sum_{i \geq 1} 2^{i(s-s_c)} < \infty$$

for any  $s < s_c$ . Hence we have  $\left\| \sum_{i \geq 1} T_i(u_{0,i}, u_{1,i}) \right\|_{H^s} < \infty$ ,  $s < s_c$ , but the solution corresponding to  $\sum_{i \geq 1} T_i(u_{0,i}, u_{1,i})$  cannot exist on any interval  $[0, T_0)$ ,  $T_0 > 0$ .

-*Critical:* If  $s = s_c$ , one expects weak global well-posedness for data small in the critical space  $\dot{H}^{s_c} \times \dot{H}^{s_c-1}$ . Note that re-scaling does not change  $\|u_0\|_{\dot{H}^{s_c} \times \dot{H}^{s_c-1}}$ , so a local well-posedness result translates into a global well-posedness result. Strong local well-posedness at the critical level is not expected due to negative results in special cases, see e. g. [60], where failure of uniformly continuous dependence on the data is shown. Thus the interval of existence  $[0, T)$  is expected to depend on the profile of the data, and not just  $\|(u_0, u_1)\|_{\dot{H}^{s_c} \times \dot{H}^{s_c-1}}$ .

-*Subcritical:* If  $s > s_c$ , one expects strong local well-posedness in  $H^s \times H^{s-1}$ , for 'well behaved equations'. However, for generic equations, this intuition fails, see [39]; in particular, one may have to impose  $s > s_1 > s_c$ . It is correct for geometric wave equations with null-structure such as WM [24], YM [31], MKG [40].

Note that  $s_c = \frac{n}{2}$  for wave maps in the local coordinate formulation.

**-Energy constraint**

Now assume that our problem of type (2.1) also admits an energy type conserved functional, i. e. a quantity  $E[u] \gtrsim \|u\|_{\dot{H}^{s_0}} + \|u_t\|_{\dot{H}^{s_0-1}}$  which is preserved under the flow.

**Examples:** -For the equation  $\square u = |u|^{p-1}u$ ,  $p \geq 1$ , on  $\mathbb{R}^{n+1}$ , we may put

$$E[u] := \int_{\mathbb{R}^n} \left( \frac{1}{2} [u_t^2 + |\nabla_x u|^2] + \frac{1}{p+1} |u|^{p+1} \right) dx,$$

whence we have  $E[u] \gtrsim \|u\|_{\dot{H}^1}$ .

-For the wave maps problem, if we embed isometrically  $M \hookrightarrow \mathbb{R}^N$ , we may define

$$E[u] := \int_{\mathbb{R}^n} (|u_t|^2 + |\nabla_x u|^2) dx$$

Alternatively, using assuming  $TM$  parallelizable and  $\phi_\alpha^i$  as in the gauge invariant formulation of wave maps, we may define

$$E[u] = \sum_{i,\alpha} \int_{\mathbb{R}^n} (\phi_\alpha^i)^2 dx$$

Then one distinguishes between the following cases:

-*Energy subcritical  $s_c < s_0$ : one expects global well-posedness, provided strong local well-posedness in the full subcritical range, or also just for some  $s_c < s < s_0$ .* Assume for example  $s_0 \geq 1$ . By finite propagation speed, to control the evolution on any finite time interval we may assume the data to be compactly supported. Then a priori control over  $\|u\|_{\dot{H}^{s_0}} + \|u_t\|_{\dot{H}^{s_0-1}}$  in conjunction with Sobolev's embedding and Hoelder's inequality results in an a priori bound on  $\|u\|_{L^2} + \|u_t\|_{L^2}$  on fixed time slices. Hence we get a priori bounds on  $\|u\|_{H^s} + \|u_t\|_{H^{s-1}}$ ,  $s \leq s_0$ , which allows us to extend the local solutions to global ones.

-*Energy critical  $s_c = s_0$ . Global well-posedness hinges on fine structure of equation.* Assume that weak global well-posedness at the critical Sobolev regularity obtains. Then finite propagation speed reduces establishing global well-posedness (in the sense of regularity preservation above the critical regularity) to excluding an energy concentration scenario.

**Examples:** -The semilinear energy critical equation  $\square u = u^5$  on  $\mathbb{R}^{3+1}$  is globally well-posed, see e. g. [56], [14], [49]. Proof of this fact relies on establishing a Morawetz type inequality, i. e. a priori estimate for a suitable space-time average of  $|u|$ .

-The wave maps problem becomes energy critical in dimension  $n = 2$ . Then the global behavior hinges on the geometric structure of the target. We have that suitable wave maps  $u : \mathbb{R}^{2+1} \rightarrow S^2$  break down in finite time, see [37], [45] and below, while it is conjectured that we have global well-posedness when the target is  $\mathbf{H}^2$ , see e. g. [22].

-*Energy supercritical  $s_c > s_0$ . No global well-posedness for generic large data expected.* This is under the assumption that there is no 'better energy type functional'

which provides a priori control for  $\|u\|_{\dot{H}^s}$ , some  $s > s_c$  in the local well-posedness range. Indeed, it appears that for all examples falling under this heading no positive well-posedness results for large data are known.

**Example:** The wave maps equation with  $n \geq 3$  is energy supercritical. Explicit blow up examples with a rather eclectic list of targets can be found in [47], [6].

### 3. THE LOCAL EXISTENCE THEORY FOR WAVE MAPS.

Recall the equation (1.1), which will suffice for the purposes of constructing a local solution, provided we further assume the data to map into a single coordinate chart. Schematically, the equation is thus of the form  $\square u = \Gamma(u)(\nabla u)^2$ ,  $(u, u_t)|_{t=0} = (u_0, u_1)$ . This crude formulation suffices for strong local well-posedness in  $H^s$  for the range  $s > \frac{n}{2} + 1$ ; recall that scaling is  $s_c = \frac{n}{2}$ . To see this, recall the Sobolev embedding  $H^{\frac{n}{2}+\epsilon}(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n)$ , whence

$$\|\Gamma(u)(\nabla u)^2\|_{H^{\frac{n}{2}+\epsilon}(\mathbb{R}^n)} \lesssim \|\nabla u\|_{H^{\frac{n}{2}+\epsilon}(\mathbb{R}^n)}^2, \epsilon > 0$$

by using the fractional Leibnitz' rule. Then using that

$$u(t, x) = S(t)u_0 + U(t)u_1 + \int_0^t U(t-s)[\Gamma(u)(\nabla u)^2](s)ds,$$

where  $S(t)u_0$ ,  $U(t)u_1$  are the free propagators for initial data  $(u_0, 0)$ ,  $(0, u_1)$ , respectively, as well as the energy inequality

$$\|S(t)u_0 + U(t)u_1\|_{\dot{H}^s(\mathbb{R}^n)} \lesssim \|u_0\|_{\dot{H}^s(\mathbb{R}^n)} + \|u_1\|_{\dot{H}^{s-1}(\mathbb{R}^n)}, \epsilon \in \mathbb{R},$$

we can run a Banach iteration on sufficiently small time intervals to get a unique fixed point:  $u(t, x) \in C([0, T], H^{\frac{n}{2}+1+\epsilon}(\mathbb{R}^n)) \cap C^1([0, T], H^{\frac{n}{2}+\epsilon}(\mathbb{R}^n))$ ,  $\epsilon > 0$ ,  $T > 0$  sufficiently small.

The main issue for the local theory then is how to improve this argument to require as little smoothness as possible. Recall from the preceding section that we expect the optimal local well-posedness threshold to be  $s_c = \frac{n}{2}$ . Indeed, a well-known example of Nirenberg, see e. g. [53], [51], shows that this cannot be improved in general: consider the scalar problem

$$(3.1) \quad \square u = -(u_t)^2 + \sum_{i=1}^n (\partial_{x_i} u)^2$$

Note that using the coordinates  $(\log r, \theta)$  on  $\mathbb{R}^2 \setminus \{0\}$  with polar coordinates  $(r, \theta)$ , the map  $(u, 0) : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^2 \setminus \{0\}$  is then a wave map. This is due to the fact that (3.1) implies  $\square(e^u) = 0$ . Now choose a free wave  $v$  on  $\mathbb{R}^{n+1}$ ,  $n \geq 2$ , such that  $\|v(0, \cdot)\|_{L_x^\infty} \ll 1$  but such that  $\|v(t, \cdot)\|_{L_x^\infty} > 1$ , and  $\sup_{t \in \mathbb{R}} \|v(t, \cdot)\|_{H^{\frac{n}{2}}} < \epsilon$ , for arbitrarily small  $\epsilon > 0$ . Then  $u = -\log(1 - v)$  is defined as long as  $v < 1$ , but becomes singular in finite time. Indeed, by rescaling  $u(t, x) \rightarrow u(\lambda t, \lambda x)$ ,  $\lambda > 1$  (which will not increase the  $\|\cdot\|_{H^{\frac{n}{2}}}$ -norm), we can achieve breakdown on arbitrarily small time intervals, even when restricting to small  $H^{\frac{n}{2}}$ -norm. This shows that we cannot expect (1.1) to be strongly locally well-posed at the critical level(or below)<sup>5</sup>. It turns out that the counterexample provided above is sharp, due to the following

<sup>5</sup>One may object here that the issue comes from the geodesic incompleteness of the target  $\mathbb{R}^2 \setminus \{0\}$ . However, there are examples of failure of local well-posedness below the critical regularity for target  $S^2$ , see [1].

theorem of Klainerman-Machedon [24] and Klainerman-Selberg [28]: we formulate it as follows (the original work only referred to wave maps confined to single coordinate charts)

**Theorem 3.1.** *Let  $(M, g)$  be uniformly isometrically embeddable<sup>6</sup> into a Euclidean space. Then the wave maps problem is strongly locally well-posed in  $H^{\frac{n}{2}+\epsilon}$ , any  $\epsilon > 0$ ,  $n \geq 2$ . Here the norm  $\|u\|_{H^{\frac{n}{2}+\epsilon}}$  is defined by using ambient coordinates.*

We outline here roughly the procedure for obtaining this theorem. Observe that the preceding argument established *fixed time slice estimates* on  $\nabla u$ , even though the logic of the preceding argument only presupposes  $L^1([0, T], H^{\frac{n}{2}+\epsilon})$  control of the nonlinearity  $\Gamma(u)(\nabla u)^2$ . In effect, it is more natural to work with  $L^2$  based spaces, on account of Plancherel's theorem. We then seek spaces  $X, Y$  (the latter to hold the nonlinearity) of functions defined (possibly locally) on space-time and satisfying the following requirements:

**Abstract function space requirements:**

- (i)  $X \subset C([0, T], H^{\frac{n}{2}+\epsilon}) \cap C^1([0, T], H^{\frac{n}{2}-1+\epsilon})$ ,  $\epsilon > 0$
- (ii)  $X \times X \subset X$
- (iii)  $\nabla_{x,t} X \times \nabla_{x,t} X \subset Y$ ,  $Y \times X \subset Y$
- (iv)  $\square^{-1} Y \subset X$
- (v)  $S(t)[H^{\frac{n}{2}+\epsilon}] \subset X$ ,  $U(t)[H^{\frac{n}{2}-1+\epsilon}] \subset X$ .

It turns out that these requirements cannot quite work. First, due to explicit examples in [39], the equation  $\square u = u_t^2$  on  $\mathbb{R}^{3+1}$  is ill-posed in  $H^2$ . Wave maps has the benefit of exhibiting a special algebraic cancellation structure, following Klainerman [20], called a *null-structure*. Hence (iii) needs to be replaced by the requirement

$$(iii') \quad Q_0(X, X) \subset Y, \quad Y \times X \subset Y, \quad \text{where } Q_0(u, v) = \sum_{\nu=0}^n \partial_\nu u \partial^\nu v.$$

Further, the fact that free waves correspond to measures supported on the light cone on the Fourier side lead to difficulties when working with  $L^2(\mathbb{R}^{n+1})$  based functions. Hence using a trick of Bourgain, working on a finite time interval  $[0, T]$ ,  $T > 0$ , one introduces a smooth cutoff  $\chi_T(t)$  supported on a dilate of  $[0, T]$  and satisfying  $\chi_T|_{[0, T]} = 1$ . Then we require<sup>7</sup>

- (iv')  $\chi_T \square^{-1} Y \subset X$
- (v')  $\chi_T S(t)[H^{\frac{n}{2}+\epsilon}] \subset X$ ,  $\chi_T U(t)[H^{\frac{n}{2}-1+\epsilon}] \subset X$ .

<sup>6</sup>This means we assume that there exists an isometric embedding such that the 2nd fundamental form has bounded derivatives with respect to geodesic coordinates on  $M$ . Thus one has (1.3), which has the same structure as (1.1).

<sup>7</sup>The implied constants depend on  $T$ .

Function spaces  $X, Y$  satisfying the requirements (i), (ii), (iii'), (iv'), (v'), were introduced in [24], in fashion similar to the  $X^{s,b}$  or Bourgain spaces introduced in the context of the nonlinear Schrodinger equation. The idea is to introduce a family of spaces analogous to Sobolev spaces but with weights on the Fourier side reflecting the geometry of the light cone.

**Definition:** Define  $X^{s,b}$  to be the completion of  $\mathcal{S}(\mathbb{R}^{n+1})$  with respect to the norm

$$\|u\|_{X^{s,b}}^2 := \int_{\mathbb{R}^{n+1}} (1 + [|\xi| + |\tau|]^2)^s (1 + ||\tau| - |\xi||^2)^b |\tilde{u}|^2(\tau, \xi) d\tau d\xi$$

Here we denote the space-time Fourier transform by<sup>8</sup>

$$\tilde{u}(\tau, \xi) := \int_{\mathbb{R}^{n+1}} u(t, x) e^{-i(t\tau + x \cdot \xi)} dt dx$$

Then we have the theorem of Klainerman-Machedon and Klainerman-Selberg [24], [28], [29]:

**Theorem 3.2.** *Let  $b > \frac{1}{2}$ ,  $s = \frac{n}{2} + \epsilon$ ,  $n \geq 2$ . Then the embeddings (i), (ii), (iii'), (iv'), (v') are true.*

One version of the proof proceeds by first showing it for truncated free waves as inputs<sup>9</sup> and then using the following **superposition principle for  $X^{s,b}$** :

**Lemma 3.3.** *Let  $u \in X^{s,b}$ ,  $b > \frac{1}{2}$ . Then we can write*

$$u = \int_{\mathbb{R}} u_a da \text{ with } \int_{\mathbb{R}} \|u_a\|_{L_t^\infty H_x^s} da \lesssim \|u\|_{X^{s,b}},$$

where each function  $u_a = e^{iat} \phi_a$ , with  $\square \phi_a = 0$ .

*Proof.* This is easily seen by writing

$$\tilde{u}(\tau, \xi) = \int_{\mathbb{R}} \delta(\tau - |\xi| - a) \tilde{u}(|\xi| + a, \xi) da, \quad \tilde{u}_a := \delta(\tau - |\xi| - a) \tilde{u}(|\xi| + a, \xi)$$

and then, using  $b > \frac{1}{2}$  and the Cauchy-Schwarz inequality,

$$\begin{aligned} & \int_{\mathbb{R}} \left( \int_{\mathbb{R}^n} (1 + |\xi|)^{2s} |\tilde{u}|^2(|\xi| + a, \xi) d\xi \right)^{\frac{1}{2}} da \\ &= \int_{\mathbb{R}} (1 + |a|)^{-b} \left( \int_{\mathbb{R}^n} (1 + |\xi|)^{2s} (1 + |a|)^{2b} |\tilde{u}|^2(|\xi| + a, \xi) d\xi \right)^{\frac{1}{2}} da \\ &\leq \left( \int_{\mathbb{R}} (1 + |a|)^{-2b} da \right)^{\frac{1}{2}} \left( \int_{\mathbb{R}^n} (1 + |\xi|)^{2s} (1 + |a|)^{2b} |\tilde{u}|^2(|\xi| + a, \xi) d\xi da \right)^{\frac{1}{2}}, \end{aligned}$$

whence  $\int_{\mathbb{R}} \|\tilde{u}(|\xi| + a, \xi)\|_{H^s} da \lesssim \|u\|_{X^{s,b}}$ .  $\square$

The role of the null-structure is easily seen by noting the identity  $2\partial_\nu u \partial^\nu v = \square(uv) - (\square u)v - u(\square v)$ . Hence one essentially reduces the proof of theorem 3.2 to establishing (i), (ii), and the 2nd part of (iii'). We refer to [29] for the details. The local solution is then constructed by a simple iteration procedure in the Banach space  $X$ . Another version of the proof proceeds via (space-time) frequency localization as well as Strichartz estimates, as detailed in the next section, which deals with well-posedness at the critical level. See e. g. [29].

<sup>8</sup>By contrast, we denote the spatial Fourier transform  $\int_{\mathbb{R}^n} e^{-ix \cdot \xi} f(x) dx$  by  $\hat{f}(\xi)$ .

<sup>9</sup>Thus one considers for example  $Q_0(u, v)$  where  $u, v$  are free waves

## 4. SMALL DATA GLOBAL EXISTENCE THEORY

One strategy toward global existence is using Klainerman's method of commuting vector fields, which is well adapted to equations exhibiting a null-form structure, such as wave maps. See e. g. [21], [53]. The difficulty with this approach is that the smallness requirement for the data is too stringent to use it as stepping stone toward the large data global existence problem in the pivotal energy critical case  $n = 2$ . Indeed, for energy critical problems (like the 2 + 1-dimensional wave maps), the most natural approach, see [8], is to first establish weak global well-posedness for small data at the energy level and then show, using a Morawetz type monotonicity statement, that the energy cannot concentrate inside a backward light cone. This suffices due to the following

**Lemma 4.1.** *Assume that the Cauchy problem for wave maps  $u : \mathbb{R}^{2+1} \rightarrow M$  is weakly globally well-posed for data of small energy. Also, assume that the energy cannot concentrate in light cones. This means that for every light cone<sup>10</sup>  $C_{t_0, x_0}^\pm = \{(t, x) | t - t_0 = \pm |x - x_0|\}$ , we have  $\lim_{t \rightarrow t_0} \int_{t \times \mathbb{R}^n \cap C_{t_0, x_0}^\pm} [|u_t|^2 + |\nabla u|^2] dx = 0$ , see comments in the proof below. Then the problem is globally well-posed in the sense of global regularity preservation for data in  $H^s \times H^{s-1}$ ,  $s > 1$ .*

*Proof.* Assume that the wave map breaks down at time  $0 < t = T < \infty$ , and exists for  $t < T$ . We may truncate the data at time  $t = 0$  to compact support. We then claim that if  $\epsilon > 0$  is such that  $\|u_0\|_{\dot{H}^1} + \|u_1\|_{L^2} < \epsilon$  ensures global regularity preservation, there exists a time  $T_0 < T$  such that

$$\int_{|x-x_0| \leq 2(T-T_0)} [|u_t|^2 + |\nabla u|^2](T_0, x) dx < \epsilon, \forall x_0 \in \mathbb{R}^n$$

Here we assume  $M \hookrightarrow \mathbb{R}^k$  via isometric embedding. Indeed, this follows easily from a compactness argument and finite propagation speed, as well as the local energy inequality. By global weak well-posedness for small data<sup>11</sup>, we may extend  $u$  of the same regularity as the initial data to the cone  $|x - x_0| = 2(T - T_0) - (t - T_0)$ ,  $T_0 \leq t \leq 2T - T_0$ . The solutions constructed on the cones as  $x_0$  varies agree on the overlap by finite propagation speed, and hence we can extend the solution past the purported blow up time, contradicting the assumption.  $\square$

**4.1. Strong global wellposedness in the critical Besov space.** A first step in this direction was taken by Tataru, who proved the following result, [66], [67]:

**Theorem 4.2.** *(Tataru) The problem (1.1) is globally strongly well-posed for small data  $(u_0, u_1)$  in the Besov space  $\dot{B}^{\frac{n}{2}, 1} \times \dot{B}^{\frac{n}{2}-1, 1}$ ,  $n \geq 2$ .*

The definition of the spaces  $\dot{B}^{\frac{n}{2}, 1}$  involves Littlewood-Paley frequency localizers as follows: choose a nonnegative  $\chi \in C_0^\infty$  with support in  $[\frac{1}{4}, 4]$  and such that  $\sum_{k \in \mathbb{Z}} \chi(\frac{x}{2^k}) = 1$  for all  $x \in \mathbb{R}_{>0}$ . Then we define

$$P_k f(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \chi\left(\frac{\xi}{2^k}\right) \hat{f}(\xi) e^{ix \cdot \xi} d\xi,$$

<sup>10</sup>Here the choice of the sign depends on what side of the plane  $t = t_0$  the wave map exists, at least locally.

<sup>11</sup>Strictly speaking, one has to truncate the solution on the slice  $t = T_0$  to define it on all of  $\mathbb{R}^n$  with small energy. This may be done upon choosing the embedding  $M \hookrightarrow \mathbb{R}^k$  to have bounded image, for example, see [43]

where  $\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)e^{ix \cdot \xi} dx$ . Note that provided  $f \in C_0^\infty(\mathbb{R}^n)$ , we have  $f = \sum_{k \in \mathbf{Z}} P_k f$ . One then defines  $\|\phi\|_{\dot{B}^{\frac{n}{2},1}(\mathbb{R}^n)} := \sum_{k \in \mathbf{Z}} 2^{\frac{nk}{2}} \|P_k \phi\|_{L^2(\mathbb{R}^n)}$  for smooth compactly supported  $\phi$ , say, and denotes by  $\dot{B}^{\frac{n}{2},1}$  the closure. We immediately observe the important embedding  $\dot{B}^{\frac{n}{2},1}(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n)$ . In particular, enforcing a smallness condition on the initial data in the Besov space ensures that the wave map will remain in a single coordinate patch, whence one may work globally in time with the formulation (1.1). The crux toward establishing Tataru's result in high dimensions, i. e.  $n \geq 4$ , consists in *frequency localization* and exploitation of *Strichartz estimates*, which are quite useful in a high dimensional setting. We recall here, see for example [51], [53]

**Theorem 4.3.** (*Strichartz estimates in high dimensions*) *Assume  $\square u = 0$ ,  $(u, u_t)|_{t=0} = (u_0, u_1)$ ,  $u \in C^\infty(\mathbb{R}^{n+1})$  and compactly supported on fixed time slices. Then we have for  $n \geq 4$*

$$\sup_{\frac{1}{p} + \frac{n-1}{2q} \leq \frac{n-1}{4}, p \geq 2} \|P_0 u\|_{L_t^p L_x^q(\mathbb{R}^{n+1})} \lesssim \|u_0\|_{L^2} + \|u_1\|_{L^2}$$

Now consider the nonlinearity in (1.1), which we simplify to  $2Q_0(u, u) = \square(u^2) - 2(\square u)u$ . As the free wave propagator respects frequency localization, one first tries to control the frequency localized constituents of the nonlinearity, i. e. the expressions  $P_k[\square(u^2) - 2(\square u)u]$ . For this one needs to find a suitable Banach space to iterate in. The idea here is to place each frequency localized component of  $u$  in a Banach space which is compatible with the scaling underlying the equation. A natural candidate for such a space is provided by the space

$$X = \left\{ \phi \mid \sum_{k \in \mathbf{Z}} \sup_{\frac{1}{p} + \frac{n-1}{2q} \leq \frac{n-1}{4}, p \geq 2} 2^{[\frac{1}{p} + \frac{n}{q}]k} \|P_k \phi\|_{L_t^p L_x^q} < \infty \right\}$$

The problem here is that this space is not well-behaved with respect to applying the operator  $\square$ . The solution here is to shrink the space by imposing a weak type of  $X^{s,b}$  control in addition: introduce the dyadic space-time frequency localizers  $Q_j$  via  $\widetilde{Q_j \phi} = \chi(\frac{\|\tau\| - \|\xi\|}{2^j}) \tilde{\phi}(\tau, \xi)$ ,  $j \in \mathbf{Z}$ . Hence again  $\sum_{j \in \mathbf{Z}} Q_j \phi = \phi$  provided  $\phi \in C_0^\infty(\mathbb{R}^{n+1})$ , say. Then our new space  $X$  is defined to be

$$X = \left\{ \phi \mid \sum_{k \in \mathbf{Z}} \sup_{\frac{1}{p} + \frac{n-1}{2q} \leq \frac{n-1}{4}, p \geq 2} 2^{[\frac{1}{p} + \frac{n}{q}]k} \|P_k \phi\|_{L_t^p L_x^q} < \infty \right\} \\ \cap \left\{ \phi \mid \sum_{k \in \mathbf{Z}} \sup_{j \in \mathbf{Z}} 2^{\frac{j}{2} + \frac{nk}{2}} \|P_k Q_j \phi\|_{L_t^2 L_x^2} < \infty \right\}$$

The correct space to place the nonlinearity in is then for example the following (see [66]) :

$$Y = \left\{ \phi \mid \sum_{k \in \mathbf{Z}} [2^{(\frac{n}{2}-1)k}] \|P_k \phi\|_{L_t^1 L_x^2} < \infty \right\}$$

Then one can show the embeddings<sup>12 13 14</sup>

$$\square IX \times X \subset Y, \square^{-1}Y \subset X, I = \sum_{k \in \mathbf{Z}} P_k Q_{<k}$$

These in conjunction with the algebra estimates  $X \times X \subset X$ ,  $X \times Y \subset Y$  and the structure of  $Q_0$  suffice to set up an iteration scheme for small data. To see how the embedding  $\square IX \times X \subset Y$  can hold, let us consider a simplified situation in which both factors are of unit frequency. Thus  $Iu = P_0 Q_{\leq 0} u$ ,  $v = P_0 v$ . Then

$$\sum_{k \in \mathbf{Z}} 2^{(\frac{n}{2}-1)k} \|P_k[\square Iuv]\|_{L_t^1 L_x^2} \lesssim \|\square Iu\|_{L_t^2 L_x^2} \|v\|_{L_t^2 L_x^\infty} \lesssim \|u\|_X \|v\|_X$$

Here one uses that the summation over  $k$  is effectively only over  $k \leq O(1)$ , due to our frequency support assumptions. The fact that  $\square^{-1}Y \subset X$  is partly due to the Strichartz estimates. We refer for a detailed proof of Theorem 4.2 in case  $n \geq 4$  to [66].

The cases  $n = 2, 3$ , treated in [67], are significantly harder, and involve a different kind of decomposition of  $u$  into travelling waves. We shall discuss this in more detail below.

**4.2. Weak global well-posedness in the critical Sobolev space.** The first result establishing weak global well-posedness for small energy data in the case  $n = 2$  (as well as  $n = 3, 4$ ) was obtained by Tao in [62], in the case when the target  $M = S^k$ ,  $k \geq 1$ . This was preceded by [61] which dealt with weak global well-posedness for data small in  $\dot{H}^{\frac{n}{2}}$ ,  $n \geq 5$ . The challenge in improving Tataru's result to data in  $\dot{H}^{\frac{n}{2}}$  lies in dealing with what one may refer to as the *summation problem*: recall that the strategy for proving Theorem 4.2 was to control the frequency localized components of the nonlinearity and then invoke the Besov structure to sum over all frequencies<sup>15</sup> involved. However, when working with Sobolev spaces, one can only assume *square summability* of the various frequency components, and hence summing over all possible frequency interactions becomes quite challenging. Tao's method takes advantage of the intrinsic Gauge freedom of the problem, albeit in somewhat complicated form. This was clarified in [27], [42] and especially [50], where Tao's result was extended to a much more general class of target manifolds and the case  $n \geq 4$ . As hinted at in the preceding section, the case  $n \geq 4$  is significantly simpler than the case  $n = 3$  and especially the case  $n = 2$ , as the dispersive effects of free waves become weaker in lower dimensions and hence estimates from the linear theory become increasingly scarce, forcing one to exploit more and more of the fine structure of the problem. Indeed, in [50] it became clear that in dimensions  $n \geq 4$ , one does not have to invoke the somewhat difficult  $X^{s,b}$ -space framework, but Strichartz estimates alone suffice. We cite

<sup>12</sup>The operator  $\square^{-1}$  refers to the solution of the inhomogeneous problem given by the Duhamel parametrix

<sup>13</sup>One uses the fact that  $P_k(L_t^1 L_x^2) \subset \dot{X}_k^{0, -\frac{1}{2}} = \{\phi | \sup_{j \in \mathbf{Z}} 2^{-\frac{j}{2}} \|P_k Q_j \phi\|_{L_t^2 L_x^2} < \infty\}$

<sup>14</sup>We define  $Q_{<k} := \sum_{j < k} Q_j$ . The point of the operator  $I$  is to ensure restriction to the hyperbolic region where the distance to the light cone is at most comparable to the spatial frequency. In the opposite case, expanding the null-form  $\partial_\nu u \partial^\nu u$  as above may be harmful.

<sup>15</sup>The only delicate issue then is to deal with destructive resonance phenomena, i. e. two high frequency waves resulting in a cascade of low frequencies. This turns out to be fairly simple.

**Theorem 4.4.** [27] ( $n \geq 5$ , boundedly parallelizable), [42] ( $n \geq 4$ , compact symmetric space target), [50] ( $n \geq 4$ , uniformly isometrically embeddable targets) Let  $n \geq 4$  and let  $(M, g) \hookrightarrow \mathbb{R}^k$  be uniformly isometrically embeddable into a Euclidean space<sup>16</sup>. Then wave maps  $u : \mathbb{R}^{n+1} \rightarrow M$  are weakly globally well-posed for data small in  $\dot{H}^{\frac{n}{2}}$ . More precisely, this statement applies to the equation for  $u$  obtained in terms of the ambient coordinates on  $\mathbb{R}^k$ , (1.3).

The method of Klainerman-Rodnianski [27], Nahmod-Stefanov-Uhlenbeck [42] and Shatah and Struwe in [50] exploits the Gauge freedom as in Tao's breakthrough work [61], [62], but from the point of view of the intrinsic formulation (1.4), (1.5). Observe that from there one may infer a system of wave equations for the  $\phi_\alpha^i$ , which may be seen to be of the form

$$\square \phi_\alpha^i = -2\Gamma_{jk}^i \phi_\nu^k \partial^\nu \phi_\alpha^j + \text{''}\phi^3\text{''}$$

Using Strichartz estimates alone and without introduction of a null-form structure, the quadratic term on the right cannot be estimated. The trick then consists in realizing that the product  $A_{j\nu}^i := -2\Gamma_{jk}^i \phi_\nu^k$  is effectively a skew-symmetric (in  $i, j$ ) connection form. This is of course a reflection of the fact that  $\{e_i\}$  is an orthonormal frame. Now one changes the Gauge, i. e. replaces  $\{e_i\}$  by  $\{\tilde{e}_i\} := \{ge_i\}$  for a suitable  $SO(k)$ -valued function  $g$  (with  $k$  the dimension of  $M$ ), and  $(A_{j\nu}^i)_{ij} =: A_\nu$  by  $\tilde{A}_\nu := g^{-1} \partial_i g + g^{-1} A_\nu g$ . Then the corresponding equation for  $\tilde{\phi}_\alpha^i$  becomes

$$\square \tilde{\phi}_\alpha^i = \tilde{A}_{j\nu}^i \partial^\nu \tilde{\phi}_\alpha^j + \text{''}\tilde{\phi}^3\text{''}$$

Imposing the Coulomb condition  $\sum_{i=1}^n \partial_i \tilde{A}_i = 0$ , and using the crucial fact that *the curvature of the original connection*  $A_\nu = (A_{j\nu}^i)_{ij}$ ,  $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha + [A_\alpha, A_\beta]$ , is quadratic in  $\phi$  thanks to (1.4), one can solve an elliptic system of equations for the  $\tilde{A}_\alpha$ , resulting schematically in  $\tilde{A} = \nabla^{-1}(\phi^2)$ , where  $\nabla^{-1}$  stands for operators of the form  $\nabla \Delta^{-1}$ . This Gauge condition is of course quite classical, and was treated in the seminal paper [69] by Karen Uhlenbeck. Thus the new leading term in the nonlinearity is now of the form

$$\nabla \phi \nabla^{-1}(\phi^2),$$

which is trilinear, and assuming the factors  $\nabla \phi$ ,  $\nabla^{-1}(\phi^2)$  to be of the same frequency, of essentially cubic form  $\phi^3$  (we are careless here about the distinction of  $\tilde{\phi}$ ,  $\phi$ ). One can now easily build a Banach space  $X$  based on Strichartz estimates alone and of essentially the kind displayed in the preceding subsection, with

$$\|\nabla \phi \nabla^{-1}(\phi^2)\|_{L_t^1 \dot{H}_x^{\frac{n}{2}-1}} \lesssim \|\phi\|_X^3,$$

as long as  $n \geq 4$ , and Strichartz estimates allow one to close a bootstrapping argument. We refer to [50] for details, but note that the use of Lorentz spaces there is easily avoidable due to the  $L_t^2 L_x^6$ -endpoint for Strichartz estimates in  $n = 4$  dimensions.

The extension of theorem 4.4 to the significantly more difficult cases  $n = 2, 3$  and target different than  $S^k$  was achieved in [33], [35], [68]. We state

<sup>16</sup>This means we assume that there exists an isometric embedding such that the 2nd fundamental form has bounded derivatives with respect to geodesic coordinates on  $M$  of all orders

**Theorem 4.5.** [33], [35], [68]. *Let the target  $M$  be uniformly isometrically embeddable into a Euclidean space or let  $M$  be the hyperbolic plane. Then the Cauchy problem for wave maps originating on  $\mathbb{R}^{n+1}$ ,  $n = 2, 3$ , is weakly globally well-posed for data of small critical Sobolev norm.*

We explain here the method of [33], where the Gauge freedom becomes particularly transparent<sup>17</sup> when the target is  $\mathbf{H}^2$ . We state the following

**Theorem 4.6.** [33] *The Cauchy problem for wave maps  $u : \mathbb{R}^{2+1} \rightarrow \mathbf{H}^2$  is globally well-posed for data of small energy. In particular, given smooth data  $(u_0, u_1) : \mathbb{R}^2 \rightarrow \mathbf{H}^2 \times T\mathbf{H}^2$ , which, upon writing  $u_0(x) = (\mathbf{x}, \mathbf{y})(x)$ ,  $\mathbf{y} > 0$ ,  $u_1(x) = (\tilde{\mathbf{x}}, \tilde{\mathbf{y}})(x)$  (where we use the identification  $T_{(\mathbf{x}, \mathbf{y})}\mathbf{H}^2 = \mathbb{R}^2$  via  $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \rightarrow \tilde{\mathbf{x}} \frac{\partial}{\partial \mathbf{x}} + \tilde{\mathbf{y}} \frac{\partial}{\partial \mathbf{y}}$ ) satisfy the inequality*

$$\int_{\mathbb{R}^2} \left( \sum_{i=1,2} \left[ \frac{\partial_i \mathbf{x}}{\mathbf{y}} \right]^2(x) + \left( \sum_{i=1,2} \left[ \frac{\partial_i \mathbf{y}}{\mathbf{y}} \right]^2(x) + \left[ \frac{\tilde{\mathbf{x}}}{\mathbf{y}} \right]^2(x) + \left[ \frac{\tilde{\mathbf{y}}}{\mathbf{y}} \right]^2(x) \right) dx < \epsilon$$

for  $\epsilon > 0$  sufficiently small, there exists a unique global in time smooth wave map  $u : \mathbb{R}^{2+1} \rightarrow \mathbf{H}^2$  extending the data globally in time; thus  $(u, u_t)|_{t=0} = (u_0, u_1)$ .

The proof departs from the intrinsic formulation in terms of the global orthonormal frame  $\{e_1, e_2\} := \{-\mathbf{y} \frac{\partial}{\partial \mathbf{x}}, -\mathbf{y} \frac{\partial}{\partial \mathbf{y}}\}$ , where we use the identification  $\mathbf{H}^2 = \{(\mathbf{x}, \mathbf{y}) | \mathbf{y} > 0\}$ . Then we have (1.6), (1.7), which can be used to deduce a system of wave equations for the quantities  $\phi_\alpha^i$ , of the following schematic form:

$$\square \phi_\alpha^1 = -2\phi_\nu^1 \partial^\nu \phi_\alpha^2 + \text{''}\phi^3\text{''}, \quad \square \phi_\alpha^2 = +2\phi_\nu^1 \partial^\nu \phi_\alpha^1 + \text{''}\phi^3\text{''},$$

where the fine structure of the terms  $\text{''}\phi^3\text{''}$  has been suppressed. Given the scarcity of Strichartz estimates in  $2+1$  dimensions (recall that one can only control norms  $\|u\|_{L_t^p L_x^q}$  where  $\frac{1}{p} + \frac{n-1}{2q} \leq \frac{n-1}{4}$ , which for  $n = 2$  translates into  $p \geq 4$ ), there is no hope to control the quadratic terms on the right hand side using Strichartz estimates alone. Indeed, there is no hope to place these in the energy space  $L_t^1 L_x^2$  even at the fixed frequency  $k = 0$ .

One then passes into the Coulomb Gauge following the procedure detailed above in the case  $n \geq 4$ . Specifically, we introduce the complex valued functions

$$\psi_\alpha = \psi_\alpha^1 + i\psi_\alpha^2 := (\phi_\alpha^1 + i\phi_\alpha^2) e^{-i\Delta^{-1} \sum_{j=1,2} \partial_j \phi_j^1}$$

These quantities then satisfy a remarkable divergence curl system of the following form:

$$(4.1) \quad \partial_\alpha \psi_\beta - \partial_\beta \psi_\alpha = i\psi_\beta \Delta^{-1} \sum_{j=1,2} \partial_j (\psi_\alpha^1 \psi_j^2 - \psi_\alpha^2 \psi_j^1) - i\psi_\alpha \Delta^{-1} \sum_{j=1,2} \partial_j (\psi_\beta^1 \psi_j^2 - \psi_\beta^2 \psi_j^1)$$

$$(4.2) \quad \partial_\nu \psi^\nu = i\psi^\nu \Delta^{-1} \sum_{j=1,2} \partial_j (\psi_\nu^1 \psi_j^2 - \psi_\nu^2 \psi_j^1).$$

<sup>17</sup>One is aided by the fact that the Gauge group  $U(1)$  is abelian in this case.

From here one can deduce a system of wave equations as follows:

$$\begin{aligned}
(4.3) \quad \square \psi_\alpha &= i\partial^\beta [\psi_\alpha \Delta^{-1} \sum_{j=1}^2 \partial_j [\psi_\beta^1 \psi_j^2 - \psi_\beta^2 \psi_j^1]] \\
&\quad - i\partial^\beta [\psi_\beta \Delta^{-1} \sum_{j=1}^2 \partial_j [\psi_\alpha^1 \psi_j^2 - \psi_\alpha^2 \psi_j^1]] \\
&\quad + i\partial_\alpha [\psi_\nu \Delta^{-1} \sum_{j=1}^2 \partial_j [\psi^{1\nu} \psi_j^2 - \psi^{2\nu} \psi_j^1]].
\end{aligned}$$

The virtue of these equations is that the right hand side is roughly speaking of the form  $\nabla[\psi \nabla^{-1}[\psi^2]]$ , and hence of cubic form, which in some sense provides more room for estimates. Nonetheless, finding a space  $X$  holding the functions  $\psi_\alpha$  as well as a space  $Y$  holding the nonlinearity still appears extraordinarily challenging, and the strategy in [33] uses a refined version of the spaces in [62], which in turn are based on Tataru's travelling wave decompositions in [67]. The rough idea behind these spaces is as follows: using the schematic form  $\nabla[\psi \nabla^{-1}[\psi^2]]$  for the nonlinearity for now, localize it to frequency  $k = 0$ , i. e. replace it by  $P_0 \nabla[\psi \nabla^{-1}[\psi^2]]$ . Ideally, assuming control over a Strichartz type norm  $\|P_k \psi\|_{L_t^2 L_x^\infty}$ , we could at least in principle try to place this expression into  $L_t^1 L_x^2$  by using an estimate of the form

$$\|P_0 \nabla[\psi \nabla^{-1}[\psi^2]]\|_{L_t^1 L_x^2} \lesssim \|\psi\|_{L_t^2 L_x^\infty} \|\nabla^{-1}[\psi^2]\|_{L_t^2 L_x^2}$$

Here the 2nd term could be estimated, at least when restricted to frequencies  $\geq 0$ , by

$$\|P_{\geq 0} \nabla^{-1}[\psi^2]\|_{L_t^2 L_x^2} \lesssim \|\psi\|_{L_t^2 L_x^\infty} \|\psi\|_{L_t^\infty L_x^2}$$

This procedure unfortunately fails for several reasons. Most prominently, in  $n = 2$  spatial dimensions, there is no hope to recover control over  $\|\psi\|_{L_t^2 L_x^\infty}$  using Strichartz estimates, as the best such estimate available is for  $\|\psi\|_{L_t^1 L_x^\infty}$ , see theorem 4.3.

In spite of this, it turns out and is well-known (e. g. [70]) that in certain situations, *it is possible to obtain control over  $\|\psi_1 \psi_2\|_{L_t^2 L_x^2}$ , provided  $\psi_1, \psi_2$  are free waves satisfying a certain angular separation condition for their space-time frequency supports*. Indeed, letting  $\psi_{1,2}$  be free waves, say at frequency 0, write  $\psi_1 = \widetilde{\mu_1}$ ,  $\psi_2 = \widetilde{\mu_2}$ , where  $\widetilde{\cdot}$  denotes space-time Fourier transform, and  $\mu_{1,2} = \delta(\tau - |\xi|) f_1(\xi)$ ,  $\mu_2 = \delta(\tau - |\xi|) f_2(\xi)$ , are measures supported on the light cone. Then, assuming the angular supports of  $f_{1,2}$  to be separated, we have  $\widetilde{\psi_1 \psi_2} = \mu_1 * \mu_2$ , whence  $\|\widetilde{\psi_1 \psi_2}\|_{L_t^1 L_x^1} \leq \|f_1\|_{L^1} \|f_2\|_{L^1}$ ,  $\|\widetilde{\psi_1 \psi_2}\|_{L_t^\infty L_x^\infty} \lesssim \|f_1\|_{L^\infty} \|f_2\|_{L^\infty}$ , and from here one gets  $\|\psi_1 \psi_2\|_{L_t^2 L_x^2} \lesssim \|\psi_1\|_{L_t^\infty L_x^2} \|\psi_2\|_{L_t^\infty L_x^2}$  by interpolation and Plancherel's theorem.

The issue then becomes how to deal with bilinear interactions  $\psi_1 \psi_2$  where no angular separation condition for the Fourier support of the factors is fulfilled. The standard way to deal with such a situation is to exploit an additional cancellation stemming from a null-form structure in the equation at hand, for example one of type  $Q_{ij}(\psi_1, \psi_2) = \partial_i \psi_1 \partial_j \psi_2 - \partial_j \psi_1 \partial_i \psi_2$ . However, such a structure is not apparent in (4.3).

To remedy this situation, one resorts to a Hodge type decomposition, which works

due to (4.1), see [33]. This trick is inspired due to a similar procedure in [26]. Specifically, one writes  $\psi_\nu = R_\nu \psi + \chi_\nu$  where  $R_\nu = \partial_\nu \sqrt{-\Delta}^{-1}$ ,  $\nu = 0, 1, 2$ , and we impose the vanishing divergence condition  $\sum_{i=1,2} \partial_i \chi_i = 0$ . From this one easily deduces relations of the schematic form  $\chi = \nabla^{-1}[\psi \nabla^{-1}[\psi^2]]$ . Now one substitutes  $\psi_\nu = R_\nu \psi + \chi_\nu$  into (4.3). The null-structure emerges if one only substitutes gradient type terms  $R_\nu \psi$ . On the other hand, the strategy implemented in [33] is that substituting at least one term  $\chi_\nu$  for  $\psi_\nu$ , one has 'more room' for estimates due to the inherently multilinear structure of  $\chi$ .

On a more technical level, the above intuition needs to be translated to a context in which the factors  $\psi_{1,2}$  etc are no longer free waves. A reasonable procedure is to use function spaces whose elements are built up from pieces which in some way behave like free waves. The breakthrough paper [67], which also provided the basis for much of the harmonic analysis in [62], achieves this by decompositions into functions replicating the behavior of travelling waves. Specifically, the intuition again comes from free waves: let  $\square u = 0$ ,  $(u, u_t)|_{t=0} = (0, g)$ , in the context of  $\mathbb{R}^{2+1}$ . Then we have

$$\begin{aligned} u(t, x) &= c \int_0^{2\pi} \int_0^\infty \frac{e^{it|\xi|} - e^{-it|\xi|}}{2|\xi|} e^{ix \cdot \xi} \hat{g}(|\xi|, \omega) |\xi| d|\xi| d\omega \\ &= \int_0^{2\pi} a_\omega(t + x \cdot \omega) d\omega - \int_0^{2\pi} b_\omega(-t + x \cdot \omega) d\omega \end{aligned}$$

where we define  $a_\omega(s) = c \int_0^\infty e^{is|\xi|} \hat{g}(|\xi|, \omega) d|\xi|$ , and similarly for  $b_\omega(s)$ . Note that each of the functions  $(t, x) \rightarrow a_\omega(t + x \cdot \omega)$ ,  $(t, x) \rightarrow b_\omega(-t + x \cdot \omega)$ , is a travelling wave. Furthermore, from Plancherel's theorem and Holder's inequality, we can conclude that if the support of  $\hat{g}(\xi)$  is contained in the region  $|\xi| \sim 1$ , we have

$$\int_0^{2\pi} \|a_\omega(\cdot)\|_{L^2_s} d\omega \lesssim \|g\|_{L^2_s}, \quad \int_0^{2\pi} \|b_\omega(\cdot)\|_{L^2_s} d\omega \lesssim \|g\|_{L^2_s}$$

Introduce the 'null-frame coordinates'

$$t_\omega := \frac{1}{\sqrt{2}}(1, \omega) \cdot (t, x), \quad x_\omega = (t, x) - \frac{t_\omega}{\sqrt{2}}(1, \omega),$$

Then if we abuse notation and write  $a_\omega(t, x) := a_\omega(t + x \cdot \omega)$  etc, we have just observed that  $\int_0^{2\pi} \|a_\omega(t, x)\|_{L^2_{t_\omega} L^\infty_{x_\omega}} d\omega \lesssim \|g\|_{L^2}$ , and similarly for  $b_\omega$ . In other words, free waves can be written as superpositions of travelling waves which satisfy the preceding relations. To move beyond the context of free waves and yet retain the property of decomposability into functions behaving in some sense like travelling waves, one introduces an atomic Banach space. More specifically, for reasons which will become apparent soon, one introduces a Banach space for each angular sector  $\kappa \subset S^1$ , as follows: define  $PW[\kappa]$  to be the atomic Banach space whose atoms are functions  $\psi$  satisfying  $\inf_{\omega \in \kappa} \|\psi\|_{L^2_{t_\omega} L^\infty_{x_\omega}} \leq 1$ . This means that we define the norm  $\|\cdot\|_{PW[\kappa]}$  as follows: assuming  $\psi \in C_0^\infty(\mathbb{R}^n)$

$$\|\psi\|_{PW[\kappa]} := \inf_{\int_\kappa \phi_\omega d\omega = \psi} \int_\kappa \|\phi_\omega\|_{L^2_{t_\omega} L^\infty_{x_\omega}},$$

where the infimum is over all representations  $\int_\kappa \phi_\omega d\omega = \psi$  with  $\omega \rightarrow C_0^\infty(\mathbb{R}^{n+1})$  continuous, say. The norms  $\|\cdot\|_{PW[\kappa]}$  are effectively our substitute for the missing

$\|\cdot\|_{L_t^2 L_x^\infty}$ .

Now consider a product  $\phi_1 \phi_2$  where the factors have angularly separated Fourier support, say along sectors  $\kappa_1, \kappa_2 \subset S^1$ . Also, suppose that the distance is much larger than  $|\kappa_2|$ . Assuming control over  $\|\phi_1\|_{PW[\kappa_1]}$ , we see that it suffices to control a norm such as  $\sup_{\omega \notin 2\kappa_2} \|\phi_2\|_{L_{t\omega}^\infty L_{x\omega}^2}$  in order to be able to bound  $\|\phi_1 \phi_2\|_{L_t^2 L_x^2}$ . That a slight modification of this indeed works is suggested by the following lemma, which tests this ansatz by looking at essentially free waves again:

**Lemma 4.7.** *Let  $\psi$  be a temporally truncated free wave with compact spatial support on fixed time slices in the context of  $\mathbb{R}^{2+1}$ . Thus we have  $\psi(t, x) = \chi(t) \tilde{\psi}(t, x)$  where  $\square \tilde{\psi} = 0$ , and  $\chi(\cdot)$  smoothly localizes to a compact time interval  $[-T, T]$ , say. Letting  $P_{0, \kappa}$  denote a Fourier multiplier which localizes smoothly to (logarithmic) frequency  $\sim 0$  and angular sector  $\kappa$ , and further  $Q_{< -2l-10}^+$  a space-time frequency multiplier smoothly restricting to distance  $\leq 2^{-2l-10}$  to the light cone,  $l \in \mathbf{N}$ , as well as the upper half space  $\tau > 0$ , we have the inequality*

$$\sup_{\omega \notin 2\kappa} \text{dist}(\omega, \kappa) \|P_{0, \kappa} Q_{< -2l-10}^+ \psi\|_{L_{t\omega}^\infty L_{x\omega}^2} \lesssim \|\psi\|_{\dot{X}_0^{1, \frac{1}{2}, 1}} = \sum_{j \in \mathbf{Z}} 2^{\frac{j}{2}} \|Q_j \psi\|_{L_t^2 L_x^2}$$

The latter quantity is bounded by the energy of  $\tilde{\psi}$ .

*Proof.* Note that for  $j < -2l - 10$ , using Hoelder's and Young's inequality

$$\|P_{0, \kappa} Q_j^+ \psi\|_{L_{t\omega}^\infty L_{x\omega}^2} \leq \|P_{0, \kappa} Q_j^+ \psi\|_{L_{\xi\omega}^2 L_{\tau\omega}^1} \lesssim [\text{dist}(\kappa, \omega)]^{-1} 2^{\frac{j}{2}} \|Q_j \psi\|_{L_t^2 L_x^2}$$

To see this, one needs to obtain a restriction on the interval for  $\tau_\omega = \frac{1}{\sqrt{2}}(1, \omega) \cdot (\tau, \xi)$  on the space time Fourier support of  $P_{0, \kappa} Q_j^+ \psi$ : note that for fixed  $\xi_\omega = (\tau - \frac{\tau_\omega}{\sqrt{2}}, \xi - \frac{\tau_\omega \omega}{\sqrt{2}}) := (\xi_\omega^1, \xi_\omega^2)$ , one has

$$2^j \sim |(\xi_\omega^1 + \frac{\tau_\omega}{\sqrt{2}})^2 - |\xi_\omega^2 + \frac{\tau_\omega \omega}{\sqrt{2}}|^2| = |(\xi_\omega^1)^2 - |\xi_\omega^2|^2 + \frac{\tau_\omega}{\sqrt{2}}(\tau - \xi \cdot \omega)|$$

Further, we can bound from below

$$|\tau - \xi \cdot \omega| = |\tau - |\xi| + |\xi| - \xi \cdot \omega| \gtrsim [\text{dis}(\omega, \kappa)]^2$$

on account of  $|\tau - |\xi|| \ll 2^{-2l} \leq [\text{dis}(\omega, \kappa)]^2$ . This implies that  $\tau_\omega$  is restricted to an interval of length  $\sim (\text{dis}(\omega, \kappa))^{-2} 2^j$ .  $\square$

The preceding observations render the functional framework below plausible. We need to define it carefully to state the precise null-form estimate, theorem 4.8, which is at the heart of the paper [33]. First, we define (using notation from [62])

$$\|\psi\|_{NFA[\kappa]^*} := \sup_{\omega \notin 2\kappa} \text{dist}(\omega, \kappa) \|\psi\|_{L_{t\omega}^\infty L_{x\omega}^2}$$

Then we define

$$\|\psi\|_{S[k, \kappa]} := \|\psi\|_{L_t^\infty L_x^2} + 2^{-\frac{k}{2}} |\kappa|^{-\frac{1}{2}} \|\psi\|_{PW[\kappa]} + \|\psi\|_{NFA[\kappa]^*}$$

**The spaces:** For every integer  $l < -10$ , subdivide  $S^1$  into a uniformly finitely overlapping collection  $K_l$  of caps  $\kappa$  of diameter  $2^l$ . Also, for every integer  $\lambda$  with  $-10 \geq \lambda \geq l$ , we subdivide the angular sector  $\{\xi \in \mathbf{R}^2 \mid \frac{\xi}{|\xi|} \in \kappa, |\xi| \sim 2^k\}$  into a uniformly finitely overlapping collection  $C_{k, \kappa, \lambda}$  of slabs  $R$  of width  $2^{k+\lambda}$ . We introduce various localization operators associated with these regions: for each  $\kappa \in K_l$ , choose a smooth cutoff  $a_\kappa : S^1 \rightarrow \mathbf{R}_{\geq 0}$  supported on a dilate of  $\kappa$ . These

are to be chosen such that  $\sum_{\kappa \in K_l} a_\kappa = 1$ . We also introduce cutoffs  $m_R(\cdot) : \mathbf{R}_{>0} \rightarrow \mathbf{R}_{\geq 0}$  such that the cutoff  $m_R(|\xi|)a_\kappa(\frac{\xi}{|\xi|})$  localizes to a dilate of the slab  $R$ . Also, we require that  $\sum_{R \in C_{k,\kappa,\lambda}} m_R(|\xi|) = \chi_0(\frac{|\xi|}{2^k})$ . We have the associated pseudo differential operator  $\tilde{P}_R\psi$ :

$$\widehat{\tilde{P}_R\psi}(t, \xi) = m_R(|\xi|)a_\kappa\left(\frac{\xi}{|\xi|}\right)\hat{\psi}(\xi)$$

We also have the  $\Psi$ DO's  $P_{k,\kappa}$  associated with multiplier  $a_\kappa(\frac{\xi}{|\xi|})\chi_0(\frac{|\xi|}{2^k})$ . Then, define

$$(4.4) \quad \begin{aligned} \|\psi\|_{S[k]} := & \|\psi\|_{L_t^\infty L^2} + \|\psi\|_{\dot{X}_k^{0, \frac{1}{2}, \infty}} + \|\psi\|_{\dot{X}_k^{-\frac{1}{2}, 1, 2}} \\ & + \sup_{\pm} \sup_{l < -10} \sup_{-10 \geq \lambda \geq l} |\lambda|^{-1} \left( \sum_{\kappa \in K_l} \sum_{R \in C_{k,\kappa,\lambda}} \|\tilde{P}_R Q_{<k+2l}^\pm \psi\|_{S[k, \pm\kappa]}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Here we use the notation  $\|\phi\|_{\dot{X}_k^{a,b,c}} := 2^{ak} [\sum_{j \in \mathbf{Z}} (2^{jb} \|Q_j \phi\|_{L_t^2 L_x^2})^c]^{\frac{1}{c}}$ . This is the norm used to control the frequency- $k$  piece of  $\psi$ .

Next, let  $N[k]$  be the atomic Banach space whose atoms are Schwartz functions  $F \in \mathcal{S}(\mathbf{R}^{2+1})$  with spatial Fourier support contained in the region  $|\xi| \sim 2^k$  and

- (1)  $\|F\|_{L_t^1 \dot{H}^{-1}} \leq 1$  and  $F$  has modulation  $< 2^{k+100}$ .
- (2)  $F$  is at modulation<sup>18</sup>  $\sim 2^j$  and satisfies  $\|F\|_{L_t^2 L_x^2} \leq 2^{\frac{j}{2}} 2^k$ .
- (3)  $F$  satisfies  $\|F\|_{\dot{X}_k^{-\frac{1}{2}, -1, 2}} \leq 1$ .
- (4) There exists an integer  $l < -10$ , and Schwartz functions  $F_\kappa$  with Fourier support in the region

$$\{(\tau, \xi) \mid \pm \tau > 0, \quad |\tau| - |\xi| \leq 2^{k-2l-100}, \quad \frac{\xi}{|\xi|} \in \pm\kappa\}$$

with the properties

$$F = \sum_{\kappa \in K_l} F_\kappa, \quad \left( \sum_{\kappa \in K_l} \|F_\kappa\|_{NFA[\kappa]}^2 \right)^{\frac{1}{2}} \leq 2^k$$

In the last inequality,  $NFA[\kappa]$  denotes the dual of  $NFA[\kappa]^*$  (the completion of  $\mathcal{S}(\mathbf{R}^{2+1})$  with respect to  $\|\cdot\|_{NFA[\kappa]^*}$ ) used in the definition of  $S[k, \kappa]$ : Thus  $NFA[\kappa]$  is the atomic Banach space whose atoms  $F$  satisfy

$$\frac{1}{\text{dist}(\omega, \kappa)} \|F\|_{L_{t\omega}^1 L_{x\omega}^2} \leq 1$$

for some  $\omega \notin 2\kappa$ .

We can now state the core null-form estimate of [33], which is the crux for estimating the right hand side of (4.3). Note that the expressions estimated below, up to the frequency localizations, are obtained by substituting the gradients  $R_\nu \psi$  for  $\psi_\nu$  on the right of (4.3):

<sup>18</sup>By this mean that  $|\tau| - |\xi| \sim 2^j$  on the space-time Fourier support of  $F$ . The term 'modulation' is adopted from [62].

**Theorem 4.8.** [33] *Let  $I = \sum_{k \in \mathbf{Z}} P_k Q_{<k+100}$ . Then there exists numbers  $\delta_{1,2} > 0$  such that*

(4.5)

$$\begin{aligned} & \|\partial^\beta P_k [R_\alpha P_{k_1} \psi_1 \Delta^{-1} \sum_{j=1}^2 \partial_j I [R_\beta P_{k_2} \psi_2 R_j P_{k_3} \psi_3 - R_\beta P_{k_3} \psi_3 R_j P_{k_2} \psi_2]] \\ & + \partial_\alpha P_k [R_\beta P_{k_1} \psi_1 \Delta^{-1} \sum_{j=1}^2 \partial_j I [R^\beta P_{k_2} \psi_2 R_j P_{k_3} \psi_3 - R_j P_{k_2} \psi_2 R^\beta P_{k_3} \psi_3]]\|_{N[k]} \\ & \leq C 2^{\delta_1 \min\{-\min\{k_1-k, k_2-k, k_3-k\}, 0\}} \prod_i 2^{\delta_2 \min\{\max_{j \neq i} \{k_i-k, k_i-k_j\}, 0\}} \prod_l \|P_{k_l} \psi_l\|_{S[k_l]}, \end{aligned}$$

(4.6)

$$\begin{aligned} & \|P_k \partial^\beta [R_\beta P_{k_1} \psi_1 \Delta^{-1} \sum_j \partial_j I [R_\alpha P_{k_2} \psi_2 R_j P_{k_3} \psi_3 - R_j P_{k_2} \psi_2 R_\alpha P_{k_3} \psi_3]]\|_{N[k]} \\ & \leq C 2^{\delta_1 \min\{-\min\{k_1-k, k_2-k, k_3-k\}, 0\}} \prod_i 2^{\delta_2 \min\{\max_{j \neq i} \{k_i-k, k_i-k_j\}, 0\}} \prod_l \|P_{k_l} \psi_l\|_{S[k_l]}. \end{aligned}$$

We observe that this theorem effectively deals with the *summation problem*, as summing over the indices  $k_i$  for fixed  $k$  is reduced to adding the squares of  $\|P_{k_i} \psi\|_{S[k_i]}$ , as is easily verified. Full details are to be found in [33].

## 5. APPROACHING THE LARGE DATA PROBLEM IN THE CRITICAL DIMENSION $n = 2$ AND HYPERBOLIC TARGET

With the result of the preceding section as well as lemma 4.1 in hand, global regularity for wave maps  $u : \mathbb{R}^{2+1} \rightarrow \mathbf{H}^2$  with large data will follow if we establish a non-concentration result for the energy inside light cones. Unfortunately, there does not seem to be a straightforward procedure for this at the present time, and the most promising strategy may be a variant of *Bourgain's induction on energy method*, [4]: the idea here is to not only show global regularity preservation, but furthermore obtain global control over a suitable space-time norm of the wave map, in particular ensuring a type of asymptotic growth control. To implement Bourgain's method, one first needs to establish a good perturbation theory for wave maps, of the following kind. We let  $u : [-T, T] \times \mathbb{R}^2 \rightarrow \mathbf{H}^2$ ,  $T \in (0, \infty]$ , a smooth wave map, with derivative components  $\phi$  with respect to the standard frame and  $\psi$  in the Coulomb Gauge:

### Perturbative setup:

Find a family of norms  $\|\psi\|_{X([-T, T] \times \mathbb{R}^2)}$ , scaling like  $\|\cdot\|_{L_t^\infty L_x^2}$ , and increasing with respect to  $T$ , with the properties that

- (i)  $\|\psi\|_{X([- \infty, \infty] \times \mathbb{R}^2)} < \infty$  whenever  $\psi$  smooth and the energy is small enough.
- (ii)  $\sup_{T > 0} \|\psi\|_{X([-T, T] \times \mathbb{R}^2)} < \infty$  and smooth data implies  $\psi$  globally smooth.
- (iii) assuming  $\|\psi\|_{X([- \infty, \infty] \times \mathbb{R}^2)} < \infty$ , there exists an open neighborhood  $U$  of the data  $(u_0, u_1)$  in the energy topology such that for a wave map  $(\tilde{u}, \tilde{\phi}, \tilde{\psi})$  with data  $(\tilde{u}_0, \tilde{u}_1) \in U$ , we have  $\|\tilde{\psi}\|_{X([- \infty, \infty] \times \mathbb{R}^2)} < \infty$ .

**Goal:** show that  $\forall E_0 \geq 0$ ,  $\sup_{E[u] \leq E_0} \|\psi\|_{X([-\infty, \infty] \times \mathbb{R}^2)} = C(E_0) < \infty$ . Here the supremum is taken over all smooth wave maps  $u : \mathbb{R}^{2+1} \rightarrow \mathbf{H}^2$  of energy  $\leq E_0$ .

Note that the perturbative setup ensures that the set of energies for which  $C(E_0) < \infty$  is nonempty and open. Hence if the Goal is false, there exists a least energy for which  $C(E_0) = \infty$ . Ideally, one wants to show that there exists a wave map of energy  $E_0$  for which  $\sup_{T > 0} \|\psi\|_{X([-T, T] \times \mathbb{R}^2)} = \infty$ . This wave map being a least energy blow up solution has to be extremely special, and one hopes to be able to rule out such solutions. More specifically, the next section will discuss the ‘bubbling off of a harmonic map’ scenario in a symmetric situation, but a modification of which is believed to also apply in the general case, see the discussion below. In particular, the non-existence of finite energy harmonic maps from  $\mathbb{R}^2$  to  $\mathbf{H}^2$  should then rule out blow up for target  $\mathbf{H}^2$ .

While there is no result yet outlining a general perturbation theory as in the above setup, the functional framework detailed before should suffice for this task. Moreover, due to the abelian nature of the Gauge group, the Coulomb Gauge can be canonically constructed for target  $\mathbf{H}^2$  even for large energy wave maps. The more general case of targets  $M$  of dimension  $k \geq 3$  and hence nonabelian Gauge group offers the additional challenge of constructing canonical Gauges substitutable for the Coulomb Gauge, which is no longer canonically constructible for large data then. In the case of negative curvature, a promising candidate for such a Gauge has been announced in [64].

Finally, we mention a partial large energy perturbative result, [36], which is based on the fact, to be discussed in the next section, that the Cauchy problem for radial data is globally well-posed, [8], [9]: we have the

**Theorem 5.1.** [36] *Let  $(u_0, u_1) : \mathbb{R}^2 \rightarrow \mathbf{H}^2 \times T\mathbf{H}^2$  be smooth, compactly supported spherically symmetric Cauchy data. Then for any  $\sigma > 0$ , there exists  $\epsilon > 0$  such that for all no longer necessarily spherically symmetric smooth initial data  $(\tilde{u}_0, \tilde{u}_1) : \mathbb{R}^2 \rightarrow \mathbf{H}^2 \times T\mathbf{H}^2$  which are  $\epsilon$ -close to  $(u_0, u_1)$  in the  $H^{1+\sigma}$ -topology, one has global existence.*

The proof of this theorem exploits additional information about the pointwise decay of radial wave maps into geodesically convex targets, which was derived in [9]. Part of the challenge consists in proving that the large energy radial wave maps can be bounded with respect to the norms discussed above. A further crucial ingredient in the proof is control over a certain range of *subcritical Sobolev norms* for the radial wave map, which is a priori only globally bounded in energy. For details, we refer to [36].

## 6. IMPOSING SYMMETRY: RADIAL AND EQUIVARIANT WAVE MAPS IN THE CASE $n = 2$ .

**6.1. Radial wave maps.** A wave map  $u : \mathbb{R}^{2+1} \rightarrow M$  is called radial, provided that it only depends on  $r$  if we equip  $\mathbb{R}^2$  with polar coordinates  $(r, \theta)$ . In this case, we have the following theorem due to Christodoulou-Tahvildar-Zadeh and Struwe:

**Theorem 6.1.** [8], [58], [59] *Let  $(u_0, u_1) : \mathbb{R}^2 \rightarrow M \times TM$  be smooth radial data,  $(M, g)$  a smooth Riemannian manifold which is either compact or satisfies a*

*suitable geodesic convexity type condition. Then there exists a unique smooth and global-in-time wave map  $u : \mathbb{R}^{2+1} \longrightarrow M$  extending  $(u_0, u_1)$ .*

This theorem was first proved in [8] for targets satisfying a type of geodesic convexity condition, and then later relaxed to general targets in [59], by means of a careful blow up analysis. We shortly explain the outline of the argument in [8], which of course parallels the strategy for the general case explained before:

- (i) Establish a small data global well-posedness result.
- (ii) Show that an energy concentration scenario is impossible.

An advantage of the radially assumption is that if a singularity forms (whence the energy concentrates in a light cone), this can only possibly occur at the spatial origin  $r = 0$ . This is on account of the conservation of energy. Furthermore, assuming an energy concentration scenario in a backward light cone centered at the space-time origin  $((t, r) = (0, 0))$ , which we may always arrange), pointwise estimates on certain components the energy momentum tensor become possible. The latter is a family of functions  $T_{\mu\nu}$ ,  $0 \leq \mu, \nu \leq 2$ , associated with the wave map  $u$  as follows:

$$T_{\mu\nu} = g_{ij}(\partial_\mu u^i \partial_\nu u^j - \frac{1}{2} m_{\mu\nu} \partial_\alpha u^i \partial^\alpha u^j), \quad m_{\mu\nu} = \text{Minkowski metric}$$

Note that the wave maps condition implies the vanishing divergence equation

$$(6.1) \quad \nabla^\nu T_{\mu\nu} = 0,$$

where we stress the use of covariant differentiation as this is to be valid for arbitrary systems of coordinates on  $\mathbb{R}^{2+1}$ , in particular the polar ones  $(t, r, \theta)$ . In the radial context, the equations (6.1) can be integrated along characteristics to yield pointwise bounds inside the light cone and away from the singularity, which then imply the vanishing of the kinetic part of the energy in an averaged sense, see [8]:

$$(6.2) \quad \lim_{T \rightarrow 0} \frac{1}{|T|} \int_{0 < |t| < |T|} \int_{|x| \leq |t|} |u_t|^2 dx = 0$$

Furthermore, one can show that the the energy has to concentrate in arbitrarily narrow neighborhoods of the axis  $r = 0$ :

$$(6.3) \quad \lim_{|t| \rightarrow 0} \int_{|t| > |x| > \lambda |t|} [|u_t|^2 + |u_r|^2] dx = 0 \quad \forall \lambda > 0$$

Assuming a kind of geodesic convexity assumption, one can bound the spatial part of the local energy  $\int_{|x| < |t|} |u_r|^2 dx$  in terms of the temporal part  $\int_{|x| < |t|} |u_t|^2$ , which together with (6.2) and the monotonicity of the local energy implies the impossibility of an energy concentration scenario in the radial case.

This argument is extended in [58], [59] to the context of general compact smooth targets. The argument in [58] is especially intuitive, and parallels developments detailed below in the equivariant case: one uses (6.2), as well as (6.3), to infer, using suitable rescalings of the wave map, existence of a nontrivial finite energy harmonic map  $u_0 : \mathbb{R}^2 \longrightarrow S^k$ , which also needs to be radial. This however is impossible, ruling out blow up.

We also mention here the paper [9], which provides detailed pointwise asymptotics of radial wave maps, under a geodesic convexity type assumption on the target.

This work provides the basis for theorem 6.1. The interesting issue remains as to whether the solutions constructed in [58], [59], satisfy similar asymptotics.

**6.2. Equivariant wave maps.** Assume that the target  $M$  admits a smooth  $S^1$  action,  $\rho : S^1 \rightarrow \text{Isom}(M)$ . A wave map  $u : \mathbb{R}^{2+1} \rightarrow M$  is called equivariant with respect to this action, provided we have

$$(6.4) \quad u(t, \omega x) = \rho(\omega)u(t, x), \quad \forall \omega \in S^1$$

Here  $S^1$  acts on  $\mathbb{R}^2$  in the canonical fashion as rotations. Paralleling the developments for radial wave maps in [8], we have the following result by Shatah-Tahvildar-Zadeh

**Theorem 6.2.** [48] *Let the target  $(M, g)$  be a warped product manifold satisfying a suitable geodesic convexity condition. Then equivariant wave maps  $u : \mathbb{R}^{2+1} \rightarrow M$  with smooth data stay globally regular.*

We refer to [48] for a detailed statement, as well as further results, including optimal local well-posedness results which are now subsumed by the more general theory expounded in section 4.

Of particular interest is the case of equivariant wave maps  $u : \mathbb{R}^{2+1} \rightarrow S^2$ , which does not satisfy the hypotheses of the preceding theorem. We let  $S^1$  act on  $S^2$  by means of rotations around the  $z$ -axis via  $\rho(\omega) = k\omega$ ,  $k \in \mathbf{Z} \setminus \{0\}$ ,  $\omega \in S^1$ . Fixing a  $k$ , the wave map is then determined in terms of the polar angle, and becomes a scalar equation on  $\mathbb{R}^{1+1}$  as follows:

$$-u_{tt} + u_{rr} + \frac{1}{r}u_r = k^2 \frac{\sin(2u)}{2r^2}$$

The case  $k = 1$  in particular is called *co-rotational*, and has aroused a lot of interest, due to numerical experiments [3], [18], , which suggested development of singularities within finite time. A rigorous result establishing the development of singularities [37] will be discussed below.

Recall that every homotopy class of maps  $u : S^2 \rightarrow S^2$  has a harmonic representative. Indeed, in terms of polar coordinates on  $S^2 = \mathbb{R}^2 \cup \{\infty\}$  as well as spherical coordinates on  $S^2$  these are given by

$$Q_k : (r, \theta) \rightarrow (2 \arctan(r^k), k\theta),$$

where  $k \in \mathbf{Z} \cong \pi_2(S^2)$ . Alternatively, one may view  $Q_k$  as the composition of  $z^k : \mathbb{C} \cup \{\infty\} \rightarrow \mathbb{C} \cup \{\infty\}$  with the (essentially) stereographic projection  $Q_1$ . Each of these harmonic maps is of course a *static wave map*.

We observe that due to weak global well-posedness for small data as well as energy conservation, a singularity in this context can only form at  $r = 0$ . A general 'bubbling off' theorem is due to Struwe in the co-rotational case  $k = 1$ :

**Theorem 6.3.** [57] *Let the wave map  $u : \mathbb{R}^{2+1} \rightarrow S^2$  be co-rotational and become singular at  $(t, r) = (0, 0)$ . Then there exists a sequence of times  $t_i \rightarrow 0$ , as well as scaling parameters  $\lambda_i$  with  $|\lambda_i| \rightarrow \infty$ , and further  $|\lambda_i t_i| \rightarrow \infty$ , such that we have*

$$(6.5) \quad u(t_i, r) = Q_1(\lambda_i r) + \epsilon(t_i, r)$$

where the local energy of  $\epsilon(t_i, \cdot)$ ,  $E_{loc}(\epsilon(t_i, \cdot))$ , converges to 0. Explicitly, this means

$$\lim_{i \rightarrow \infty} \int_{r < |t_i|} \left( \frac{1}{2} [\epsilon_t^2(t_i, r) + \epsilon_r^2(t_i, r)] + \frac{\sin^2(\epsilon(t_i, r))}{r^2} \right) r dr = 0$$

In other words, if a co-rotational wave map  $u : \mathbb{R}^{2+1} \rightarrow S^2$  develops a singularity, there exists a sequence of times  $t_i$  approaching the singularity such that, upon restricting the wave map to the light cone centered at the singularity, it can be decoupled into a re-scaled harmonic map  $Q_1(\lambda_i r)$ , as well as a radiation part, which converges to zero in the energy topology. Furthermore, the re-scaling is strictly faster than self-similar, whence the energy concentrates in a cuspidal region  $r < \frac{1}{\lambda_i}$  (on each time slice  $t = t_i$ ). Note that the fact that  $|\lambda_i t_i| \rightarrow \infty$  is a reflection of (6.3), which can also be proven for equivariant wave maps, see [48], by the same technique as in the radial case.

Note that Struwe's theorem does not make an assertion about the actual occurrence of singularity formation, nor about the rate at which  $|\lambda_i| \rightarrow \infty$ . For example, it is an open issue whether the above decomposition (6.5) holds on each time sufficiently close to the blow up.

**6.3. Consequences for non-symmetric wave maps in the critical dimension.** Theorem 6.3 indicates that singularity formation for wave maps may be tied to an inherent lack of compactness, due to the scaling invariance. Removing the equivariance enlarges this lack of compactness, by adding translations, as well as Lorentz boosts, for example of the form

$$(t, x_1, x_2) \rightarrow \left( \frac{ct - x_1}{\sqrt{c^2 - 1}}, \frac{t - cx_1}{\sqrt{c^2 - 2}}, x_2 \right), c > 1$$

Fortunately, unlike for example in the context of the critical defocusing Schrödinger equation on  $\mathbb{R}^{3+1}$ , where the issue of translation invariance poses significant difficulties, see [10], the finite propagation speed of wave maps, as well as energy conservation, imply that for general wave maps without symmetry assumptions, singularity formation can only happen via a finite number of energy concentration scenarios within a forward light cone, say (posing data at time  $t = 0$ ). One may then concentrate on what happens within a single such light cone. Note that each Lorentz boost centered at the tip of the light cone leaves the latter invariant, but moves its time axis. Then the following general blow up scenario may be plausible:

**Conjectured general bubbling off scenario:** Let  $u : \mathbb{R}^{2+1} \rightarrow S^2$  be a wave map, no longer necessarily equivariant, which becomes singular at  $(t, x) = (0, 0)$ . Then there exists  $N$  and a sequence of times  $t_i \rightarrow 0$ , and for each  $i$  Lorentz boosts  $T_i^k$ ,  $k = 1, \dots, N$ , as well as parameters  $\lambda_i^k$  with  $|t_i \lambda_i^k| \rightarrow \infty$  and harmonic maps  $Q_k(x)$ , such that the wave map, restricted to the half of the light cone on which it is defined, may be decomposed into

$$u(t_i, x) = \sum_{k=1}^N T_i^k [Q_k(\lambda_i^k x)] + \epsilon(t_i, x),$$

where we have  $E_{loc}(\epsilon(t_i, x)) \rightarrow 0$ . Here the wave map is represented in terms of the ambient coordinates  $S^2 \hookrightarrow \mathbb{R}^3$ .

## 7. SINGULARITY FORMATION IN THE CRITICAL DIMENSION

Recall from section 2 that due to the supercritical nature of wave maps  $\mathbb{R}^{n+1} \rightarrow M$ ,  $n \geq 3$ , generic targets  $M$ , one expects singularity formation for large data. This has been achieved for a rather eclectic list of targets in [47], [6], via a self-similar

ansatz, i. e. writing  $u(t, x) = v(\frac{x}{t})$ , and then solving for  $v$ , which leads to an elliptic problem for  $v$  that can be solved by variational techniques in certain cases. In particular, it is known that self-similar blow up wave maps with smooth data exist with domain  $\mathbb{R}^{3+1}$  and target  $S^3$ . In the latter case, the explicit self-similar solutions are believed to represent the generic blow up phenomenon, see e. g. [2], although this appears out of reach of present techniques. No other types of blow up solutions are known in the supercritical regime. As seen below, recent blow up results in the critical dimension are of a strikingly different character. In particular, the rates are never self-similar.

**7.1. The co-rotational case.** In the critical dimension  $n = 2$ , existence of singular wave maps with target  $S^2$  has been conjectured for a while, on account of numerical experiments, see e. g. [3]. These have been for the most part restricted to co-rotational wave maps  $u : \mathbb{R}^{2+1} \rightarrow S^2$ , see preceding section, and thus deal with the scalar equation (with  $u$  the polar angle)

$$(7.1) \quad -u_{tt} + u_{rr} + \frac{1}{r}u_r = \frac{\sin(2u)}{2r^2}$$

Recall that these admit the static solution  $u(t, r) = Q_1(r) = 2 \arctan(r)$ . It was observed in [3] that sufficiently large perturbations of these appear to lead to singularities in finite time, while small enough perturbations lead to global existence and scattering. Rigorous existence of blow up solutions with precise blow up dynamics is provided by the following

**Theorem 7.1.** [37] *Let  $\nu > \frac{1}{2}$ , and let  $t_0 > 0$  sufficiently small. Define  $\lambda(t) = t^{-1-\nu}$ , and fix a large integer  $N$ . Then there exists a function  $u^e$  satisfying*

$$u^e \in C^{\nu+\frac{1}{2+}}(t_0 > t > 0, |x| \leq t), E_{loc}(u^e)(t) \lesssim [\lambda(t)t]^{-2} |\log t|^2$$

and a blow up solution  $u(t, r)$  to (7.1) on  $[0, t_0]$  which has the form

$$u(r, t) = Q_1(\lambda(t)r) + u^e(r, t) + \epsilon(r, t), \quad 0 \leq r \leq t,$$

with  $\epsilon \in t^N H_{loc}^{1+\nu-}(\mathbb{R}^2)$ ,  $\epsilon_t \in t^{N-1} H_{loc}^{\nu-}(\mathbb{R}^2)$ ,  $E_{loc}(\epsilon) \lesssim t^N$ . The corresponding solution  $u(t, x)$  can be extended of class  $H^{1+\nu-}$  to all of  $[0, t_0] \times \mathbb{R}^2$ .

A surprising feature is that there is a continuum of blow up rates. Indeed, the fact that a nonlinear equation may admit such a family of blow up rates seems to have been observed here for the first time. Furthermore, the fact that the energy of the static solutions  $Q_k(\cdot)$  is  $4k\pi$ , together with the conjectured general bubbling off scenario from above, appear to suggest that theorem 7.1 is optimal in some sense: any initial data of energy strictly below that of  $Q_1(\cdot)$ , which is  $4\pi$ , should lead to globally regular solutions, while the theorem produces blow up solutions of energy arbitrarily close to  $4\pi$ .

We now outline the strategy for proving this theorem: one may be tempted to try a naive perturbative ansatz, namely

$$(7.2) \quad u(t, r) = Q(\lambda(t)r) + \epsilon(t, r),$$

where  $\epsilon(t, \cdot)$  is small in a suitable sense, and where  $\lambda(t)$  is chosen to blow up at  $t = 0$ . In particular, one may try to construct  $\epsilon$  by means of an iterative procedure.

This ansatz however appears to fail: note that substituting  $Q(\lambda(t)r)$  into (7.1), one obtains an error of the form

$$(7.3) \quad \ddot{\lambda}(t)rQ'(\lambda(t)r) + (\dot{\lambda})^2r^2Q''(\lambda(t)r)$$

However, note that this function is not in  $L^2(rdr)$ , unless  $\lambda(t) = t^{-1}$ , which we know is excluded from Struwe's theorem 6.3. Furthermore, this function is not going to be small, and hence we cannot expect that adding on a small function  $\epsilon$ , which is obtained by means of a direct iteration, will counteract this error. The fact that (7.3) is not  $L^2$ -integrable is related to an interesting feature of the co-rotational case  $k = \pm 1$ , namely the fact that *the spectrum of the linearization around the static solution  $Q_1(\cdot)$  has a resonance at its endpoint  $\xi = 0$* . On the other hand, for equivariance indices  $|k| \geq 2$ , the linearization around  $Q_k(\cdot)$  has an eigenvalue at  $\xi = 0$ . Indeed, the resonance and eigenfunctions are given by  $\frac{\partial}{\partial \lambda} Q_k(\lambda r)|_{\lambda=1} = rQ'_k(r)$ .

Note that one may be tempted to deal with the non-square-integrability of (7.3) by truncation at infinity. Indeed, one is really only interested in constructing solutions on the backward light cone  $r \leq t$ , as one may extend these arbitrarily outside the light cone. Hence one may try to replace  $Q(\lambda(t)r)$  by  $\chi(\frac{r}{t})Q(\lambda(t)r)$  for some smooth cutoff  $\chi(\frac{r}{t})$ , where we write  $Q_1(\cdot) =: Q$  for the remainder of this subsection. Indeed, rather than imposing a fixed form of  $\chi$ , one may try to substitute an ansatz  $\chi(\frac{r}{t})Q(\lambda(t)r) + \epsilon(t, r)$  into (7.1) and thereby obtain the correct form of  $\chi$ . Indeed, it turns out that one obtains an equation of the form

$$(7.4) \quad (a^2 - 1)\chi_{aa} + [2a(1 - \frac{\lambda_t t}{\lambda}) + a^{-1}]\chi_a = (\text{independent of } \chi) + O(\chi^2),$$

where we have introduced the self-similar variable  $a = \frac{r}{t}$ . The operator  $(a^2 - 1)\partial_a^2 + [2a(1 - \frac{\lambda_t t}{\lambda}) + a^{-1}]\partial_a$  is a singular linear differential operator, which for  $\lambda(t) = t^{-1-\nu}$  admits a fundamental system of solutions of regularity  $C^{\nu + \frac{1}{2+}}$  across  $a = 1$ , which corresponds to the light cone. Unfortunately, equation (7.4) is effectively inconsistent, as the right hand side depends on  $t$  in addition to  $a$ , and hence we really would have to choose  $\chi$  of the form  $\chi(\frac{r}{t}, t)$ , introducing additional error terms due to time differentiation. Nevertheless, these observations motivate one to look for an approximate solution (which is the  $Q(\lambda(t)r) + u^\epsilon(t, r)$  in the theorem), which is in effect a large profile modification of  $Q(\lambda(t)r)$  for fixed times  $t$  away from  $t = 0$ , and obtained by solving certain elliptic problems on fixed time slices which in some sense approximate the hyperbolic problem.

The precise procedure for proving theorem 7.1 is then as follows: to find  $u^\epsilon$ , one solves a sequence of elliptic problems on fixed time slices (thus in some sense neglecting time derivatives) which improve the accuracy of the problem near the symmetry axis  $r = 0$ , as well as the light cone  $r = t$ . The latter is achieved by working with the coordinates  $(a, t)$  instead of  $(r, t)$ , as in the preceding paragraph. Specifically, if  $u$  is an approximate solution of (7.1), an exact solution  $u + \epsilon$  is given with  $\epsilon$  satisfying

$$(-\partial_t^2 + \partial_r^2 + \frac{1}{r}\partial_r)\epsilon - \frac{\cos(2u)}{2r^2}\sin(2\epsilon) + \frac{\sin(2u)}{2r^2}(1 - \cos(2\epsilon)) = e,$$

where the error  $e$  generated by  $u$  is given by

$$e = (-\partial_t^2 + \partial_r^2 + \frac{1}{r}\partial_r)u - \frac{\sin(2u)}{2r^2}$$

We approximate  $\sin(2\epsilon)$  by  $2\epsilon$ , and moving nonlinear (in  $\epsilon$ ) terms to the right, we encounter the problem

$$(-\partial_t^2 + \partial_r^2 + \frac{1}{r}\partial_r - \frac{\cos(2u)}{r^2})\epsilon = \text{small}$$

Near the origin  $r = 0$ , one expects the time derivatives to play less of a role. Moreover, the smaller  $t$  becomes the less  $u$  should differ from  $Q(\lambda(t)r)$ , which suggests replacing  $u$  by the latter. Hence one obtains the problem

$$(\partial_r^2 + \frac{1}{r}\partial_r - \frac{\cos[Q(\lambda(t)r)]}{r^2})\epsilon = \text{small}$$

On the other hand, near  $r = t$ , one expects the time derivatives to play a role, and replaces the above by

$$(-\partial_t^2 + \partial_r^2 + \frac{1}{r}\partial_r - \frac{1}{r^2})\epsilon = \text{small}$$

In short, one constructs  $u^\epsilon = \sum_{k=1}^M v_k$ , where we put

$$(7.5) \quad (\partial_r^2 + \frac{1}{r}\partial_r - \frac{\cos[Q(\lambda(t)r)]}{r^2})v_{2k+1} = e_{2k}$$

$$(7.6) \quad (-\partial_t^2 + \partial_r^2 + \frac{1}{r}\partial_r - \frac{1}{r^2})v_{2k} = e_{2k+1}$$

where the errors  $e_k$  generated become increasingly smaller. It turns out that (7.6), when translated into the coordinates  $a = \frac{r}{t}$ ,  $t$ , leads to a problem of the form (7.4) for a suitable auxiliary function.. For details we refer to section 2 in [37]. Choosing  $M$  large enough, one may achieve  $e_M = O(\frac{1}{[\lambda(t)t]^N})$ , for arbitrary  $N$ .

Having constructed an approximate solution  $Q(\lambda(t)r) + u^\epsilon(t, r)$  with

$$(-\partial_t^2 + \partial_r^2 + \frac{1}{r}\partial_r)[Q(\lambda(t)r) + u^\epsilon(t, r)] - \frac{\sin[2(Q(\lambda(t)r) + u^\epsilon(t, r))]}{2r^2} = O(\frac{1}{[\lambda(t)t]^N})$$

one needs to correct it to a precise solution  $u = Q(\lambda(t)r) + u^\epsilon(t, r) + \epsilon(t, r)$ , where  $\epsilon(t, r)$  is to be solved for by means of iteration. We now consider the equation for  $\epsilon$ . In order to avoid a time-dependent elliptic operator, one replaces the coordinates  $(t, r)$  by the new ones  $\tau = -\int_t^1 \lambda(s)ds + \frac{1}{v}$ , and  $R = \lambda(t)r$ . Also, we replace  $\epsilon$  by  $\tilde{\epsilon}(\tau, R) := R^{\frac{1}{2}}\epsilon(t, r)$ . Then one obtains the problem

$$(7.7) \quad [-(\partial_\tau + \frac{\lambda_\tau}{\lambda}R\partial_R)^2 + \frac{1}{4}(\frac{\lambda_\tau}{\lambda})^2 + \frac{1}{2}\partial_\tau(\frac{\lambda_\tau}{\lambda})]\tilde{\epsilon} - \mathcal{L}\tilde{\epsilon} = \text{small},$$

where

$$\mathcal{L} := -\partial_R^2 + \frac{3}{4R^2} - \frac{8}{(1+R^2)^2}$$

The latter operator has the property that in spite of being defined on the half-line, it is self-adjoint even without the extra imposition of a boundary condition at  $R = 0$ . This is due to the fact that the singularity at  $R = 0$  already forces a vanishing condition at  $R = 0$  for functions in its domain. The theory of operators of the type of  $\mathcal{L}$ , so called *strongly singular operators*, is developed in [13]. We need

the fact that the spectrum of  $\mathcal{L}$  consists of  $[0, \infty)$ , and 0 is a resonance for  $\mathcal{L}$ , which is now considered as an operator on  $L^2([0, \infty))$ . Indeed, we have that

$$\mathcal{L}\left(\frac{R^{\frac{3}{2}}}{1+R^2}\right) = 0$$

It turns out that  $\mathcal{L}$  admits for each  $z \in \mathbb{C}$  a fundamental system of solutions  $\phi(R, z)$ ,  $\theta(R, z)$  for the eigenvalue problem  $(\mathcal{L} - z)f = 0$ , which obey the asymptotics

$$\phi(R, z) \sim R^{\frac{3}{2}}, \theta(R, z) \sim R^{-\frac{1}{2}}$$

One also says that  $\mathcal{L} - z$  is in the limit point case at  $R = 0$ , see e. g. [38], and  $\phi(R, z)$  is the Weyl-Titchmarsh solution of  $\mathcal{L} - z$  at  $R = 0$ .

Further, assuming now  $\text{Im } z > 0$ , the Weyl-Titchmarsh solution at  $R = \infty$  is given asymptotically by

$$\psi^+(R, z) \sim z^{-\frac{1}{4}} e^{iz^{\frac{1}{2}}R}, \text{Im } z^{\frac{1}{2}} > 0,$$

and letting  $\text{Im } z \rightarrow 0$ , one obtains an oscillatory (at  $R = \infty$ ) fundamental system of solutions  $\psi^+(R, \xi)$ ,  $\psi^-(R, \xi) = \overline{\psi^+(R, \xi)}$ ,  $\xi > 0$ . We then also have the condition

$$W(\theta, \phi) = 1, W(f, g) = \left(\frac{d}{dR}f\right)g - \left(\frac{d}{dR}g\right)f$$

With these tools, one can then introduce a distorted Fourier transform, as follows:

$$f(R) \longrightarrow \hat{f}(\xi) := \int_0^\infty \phi(R, \xi) f(R) dR$$

which is to be interpreted in a weighted  $L^2$ -sense, similarly to the ordinary Fourier transform. Further, one has the inversion relation

$$f(R) = \int_0^\infty \phi(R, \xi) \hat{f}(\xi) \rho(\xi) d\xi, \rho(\xi) = \frac{1}{\pi} \text{Im}(m(\xi + i0)),$$

where the spectral density  $\rho$  is defined in terms of

$$m(\xi) := \frac{W(\theta(\cdot, \xi), \psi^+(\cdot, \xi))}{W(\psi^+(\cdot, \xi), \phi(\cdot, \xi))}$$

The construction of  $\tilde{\epsilon}(\tau, R)$  as above now proceeds via representing

$$\tilde{\epsilon}(\tau, R) = \int_0^\infty x(\tau, \xi) \phi(R, \xi) \rho(\xi) d\xi,$$

and working with the Fourier coefficients  $x(\tau, \xi)$  instead. The difficulty one encounters here is that it is not immediate (as in the case of the free d'Alembertian  $\square$ ) to deduce a transport equation for  $x(\tau, \xi)$  from (7.7). Nevertheless, neglecting the terms  $\frac{\lambda_\mp}{\lambda} R \partial_R$  in  $(\partial_\tau + \frac{\lambda_\mp}{\lambda} R \partial_R)$  allows one to obtain an approximate transport equation for the  $x(\tau, \xi)$ , which is enough thanks to the rapid decay of the error (namely  $(\lambda(t)t)^{-N}$ ). For details, we again refer to [37].

We conclude this subsection by noting that interesting open questions remain:

- (i) What are the stability properties of the solutions constructed? One may conjecture that there is a high co-dimensional manifold of data resulting in the same blow up, which would be somewhat analogous to [5].
- (ii) Can the same construction be carried out for the higher homotopy indices  $k \geq 2$ ? It appears that the resonance for the linearization around the ground state plays a

fundamental role.

(iii) Are there  $C^\infty$ -smooth data resulting in the kind of blow up rates constructed in the theorem?

(iv) What other targets admit such blow up solutions? Can one say that the existence of non-trivial finite-energy harmonic maps originating on  $\mathbb{R}^2$  to the target suffices?

**7.2. Higher equivariance classes.** Now consider equivariant wave maps  $u : \mathbb{R}^{2+1} \rightarrow S^2$  but of index  $k \geq 4$ . Then we have the following result due to Rodnianski and Sterbenz:

**Theorem 7.2.** [45] *For  $k \geq 4$ , there exists an open (with respect to a suitable Sobolev topology) set of  $k$ -equivariant initial data (in particular  $C^\infty$ -data) arbitrarily close to  $Q_k(\cdot)$ , resulting in blow up solutions of the form*

$$u(t, r) = Q_k(\lambda(t)r) + \epsilon(t, r),$$

where we have  $\lambda(t) \sim \frac{\sqrt{|\log(T-t)|}}{T-t}$ , and  $T$  is the blow up time.

Observe that this theorem guarantees an open (within the  $k$ -equivariant category) set of data resulting in a kind of *stable blow up*. Not surprisingly, the method of proof here is quite different than the one of theorem 7.1, and deduces the blow up rate from monotonicity type arguments and an orthogonality relation, rather than imposing it. Moreover, rather than constructing  $\epsilon(t, r)$  via iteration, Rodnianski-Sterbenz control it by means of a Morawetz type estimate (for a time dependent wave operator!), hence via a priori type estimates. The restriction  $k \geq 4$  should be relaxable to  $k \geq 2$ , while the case  $k = 1$  appears unreachable, as the method heavily relies on the  $L^2$ -integrability of the zero mode  $\frac{\partial}{\partial \lambda} [Q_k(\lambda r)]|_{\lambda=1}$ . We refer to [45] for details. It remains an interesting open issue to see whether the methods of theorem 7.1, theorem 7.2 can be combined to deduce a stable blow up regime in the co-rotational case  $k = 1$ . Further, the issue of *whether any of these blow up solutions remain stable in the full category of (non-equivariant) wave maps  $u : \mathbb{R}^{2+1} \rightarrow S^2$  appears a quite difficult open problem.*

**Acknowledgment** The author thanks the referee for pointing out numerous improvements.

## REFERENCES

- [1] P. D’Ancona, V. Georgiev *On the continuity of the solution operator of the wave maps system*, preprint
- [2] P. Bizon, *Comm. Math. Phys.* 215(2000), 45
- [3] P. Bizon, T. Chmaj, T. Zbislaw, *Formation of singularities for equivariant (2+1)-dimensional wave maps into the 2-sphere.*, *Nonlinearity* 14 (2001), no. 5, 1041–1053.
- [4] J. Bourgain, *Global wellposedness of defocusing critical nonlinear Schrödinger equation in the radial case.*, *J. Amer. Math. Soc.* 12 (1999), no. 1, 145–171.
- [5] J. Bourgain, Wang, W., *Construction of blowup solutions for the nonlinear Schrödinger equation with critical nonlinearity*, *Annali Scuola Ecole Normale Cl.*(4) 25 (1997), no.1-2, 197-215
- [6] Cazenave, Thierry; Shatah, Jalal; Tahvildar-Zadeh, A. Shadi *Harmonic Maps and the development of singularities in Wave Maps and Yang-Mills fields.*, *Ann. Inst. H. Poincaré Phys. Theor.* 68 (1998), no.3, 315-349

- [7] Y.Choquet-Bruhat, V.Moncrief, *Existence theorem for solutions of Einstein's equations with 1 parameter spacelike isometry groups*, Quantization, nonlinear partial differential equations, and operator algebra(Cambridge, MA, 1994), 67-80, Proc. Sympos. Pure Math., 59, Amer. Math. Soc., Providence, RI, 1996
- [8] D.Christodoulou, A. Tahvildar-Zadeh *On the regularity of spherically symmetric wave maps*, C.P.A.M., 46(1993), 1041-1091
- [9] D. Christodoulou, A. Tahvildar-Zadeh *On the asymptotic behavior of spherically symmetric wave maps*, Duke Math. J.71 (1993), no.1, 31-69
- [10] J.Coliander, M.Keel, G.Staffilani, H.Takaoka, T.Tao, *Global well-posedness and scattering in the energy space for the critical nonlinear Schrodinger equation in  $R^3$* , preprint
- [11] A.Freire, S.Muller, M.Struwe, *Weak convergence of wave maps from  $(1+2)$ -dimensional Minkowski space to Riemannian manifolds*, Inventiones Math. 130 (1997), 589-617.
- [12] M.Gell-Mann, M.Levy, *The axial vector current in beta decay*, Nuovo Cim. 16(1960) 705
- [13] F.Gesztesy, M.Zinchenko, *On spectral theory for Schrödinger operators with strongly singular potentials*. preprint 2005.
- [14] M.Grillakis, *Regularity and asymptotic behaviour of the wave equation with a critical nonlinearity.*, Ann. of Math. (2) 132 (1990), no. 3, 485-509
- [15] M. Guenther, *Isometric embeddings of Riemannian manifolds*, Proceedings Internatl. Congress of Mathematicians, Vol. 1-2 (Kyoto 1990),p.1137-1143, Tokyo, 1991, Math. Soc. Japan.
- [16] C.-H. Gu, *On the Cauchy problem for harmonic maps defined on two-dimensional Minkowski space*, Comm. Pure Appl. Math. 33: 727-737, 1980.
- [17] F.Helein, *Regularite des applications faiblement harmoniques entre une surface et une varietee Riemannienne*, C.R.Acad.Sci.Paris Ser.1 Math 312(1991), 591-596
- [18] J.Isenberg, S.Liebling, *Singularity formation in  $2 + 1$  wave maps*. J. Math. Phys. 43 (2002), no. 1, 678-683.
- [19] M.Keel, T.Tao *Endpoint Strichartz estimates*, Amer.J. Math.120 (1998), no.5, 955-980
- [20] S.Klainerman, *The null condition and global existence to nonlinear wave equations*. Nonlinear systems of partial differential equations in applied mathematics, Part 1 (Santa Fe, N.M., 1984), 293-326, Lectures in Appl. Math., 23, Amer. Math. Soc., Providence, RI, 1986.
- [21] S.Klainerman, *Uniform decay estimates and the Lorentz invariance of the classical wave equation.*, Comm. Pure Appl. Math. 38 (1985), no. 3, 321-332.
- [22] S.Klainerman, *UCLA lectures on nonlin. wave eqns.*, preprint (2001)
- [23] S.Klainerman, D.Foschi, *Bilinear Space-Time Estimates for Homogeneous Wave Equations*, Ann. Scient. Ec. Norm. Sup., 4e serie, t.33(2000), 211-274
- [24] S.Klainerman, M.Machedon, *Smoothing estimates for null forms and applications*, Duke Math.J., 81(1995), 99-133
- [25] S.Klainerman, M.Machedon, *On the algebraic properties of the  $H^{\frac{n}{2}, \frac{1}{2}}$  spaces*, I.M.R.N. 15(1998), 765-774
- [26] S.Klainerman, M.Machedon, *On the regularity properties of a model problem related to wave maps*, Duke Math.J., 87(1997), 553-589
- [27] S.Klainerman, I.Rodnianski, *On the global regularity of wave maps in the critical Sobolev norm*, I.M.R.N. 13(2001), 655-677
- [28] S.Klainerman, S.Selberg, *Remark on the optimal regularity for equations of wave maps type*, C.P.D.E., 22(1997), 901-918
- [29] S.Klainerman, S.Selberg, *The spaces  $H^{s, \theta}$  and applications to nonlinear wave equations.*, preprint
- [30] S.Klainerman, S.Selberg, *Bilinear estimates and applications to nonlinear wave equations*, preprint
- [31] S.Klainerman, D.Tataru, *On the optimal regularity for the Yang-Mills equations in  $\mathbf{R}^{4+1}$* , Journal of the American Math. Soc., 12(1999), 93-116
- [32] J.Krieger, *Global Regularity of Wave Maps in 2 and 3 spatial dimensions*, Ph. D. Thesis, Princeton University (2003)
- [33] J. Krieger, *Global regularity of Wave Maps from  $\mathbf{R}^{3+1}$  to surfaces*, CMP 238/1-2 (2003), 333-366
- [34] J.Krieger, *Null-Form estimates and nonlinear waves*, Adv. Differential Equations 8(2003), no.10, 1193-1236
- [35] J. Krieger, *Global regularity of Wave Maps from  $\mathbf{R}^{2+1}$  to  $\mathbf{H}^2$* , CMP 250(2004), 507-580

- [36] J.Krieger, *Stability of Spherically Symmetric Wave Maps*, Memoirs of the AMS Vol. 181(2006), no. 853
- [37] J. Krieger, W.Schlag, D.Tataru, *Renormalization and blow up for charge one equivariant critical wave maps*, preprint
- [38] B.Levitan, I.Sargsjan, *Sturm-Liouville and Dirac operators*. Translated from the Russian. Mathematics and its Applications (Soviet Series), 59. Kluwer Academic Publishers Group, Dordrecht, 1991.
- [39] H.Lindblad, *A sharp counterexample to the local existence of low-regularity solutions to nonlinear wave equations*, Duke Math. J. 72(1993), 503-539
- [40] M.Machedon, J.Sterbenz, *Almost optimal local well-posedness for the  $(3 + 1)$ -dimensional Maxwell-Klein-Gordon equations.*, J. Amer. Math. Soc. 17 (2004), no. 2, 297-359
- [41] S.Muller, M.Struwe, *Global existence of wave maps in  $1+2$  dimensions with finite energy data*, Topological methods in nonlinear analysis 7 (1996), 245-259.
- [42] A.Nahmod, A.Stefanov, K.Uhlenbeck, *On the well-posedness of the wave maps problem in high dimensions*, Comm. Anal. Geom. 11(2003) no.1, 49-83
- [43] J.Nash, *John The imbedding problem for Riemannian manifolds.*, Ann. of Math. (2) 63 (1956), 20-63.
- [44] I.Rodnianski, *The wave map problem small data critical regularity*, Seminaire Bourbaki, 58eme annee, 2005-2006, no. 965
- [45] I.Rodnianski, J.Sterbenz, *On the formation of singularities in the critical  $O(3)$   $\sigma$ -model*, preprint
- [46] S.Selberg, *Multilinear space-time estimates and applications to local existence theory for nonlinear wave equations*, Ph.D. thesis, Princeton University, 1999
- [47] J.Shatah, *Weak solutions and development of singularities of the  $SU(2)$   $\sigma$ -model.*, Comm. Pure Appl. Math. 41 (1988), no. 4, 459-469
- [48] J.Shatah, A. Tahvildar-Zadeh, *On the Cauchy Problem for Equivariant Wave Maps*, Comm. Pure Appl. Math. 47(1994), 719-754
- [49] J.Shatah, M.Struwe *Regularity results for nonlinear wave equations.*, Ann. of Math. (2) 138 (1993), no. 3, 503-518.
- [50] J.Shatah, M.Struwe *The Cauchy problem for wave maps*, I.M.R.N.11(2002), 555-571
- [51] J.Shatah, M. Struwe *Geometric Wave Equations*, AMS Courant Lecture Notes 2
- [52] T.Sideris, *Global existence of harmonic maps in Minkowski space*, Comm. Pure Appl. Math. 42(1989), no.1, 1-13
- [53] C.D.Sogge, *Lectures on nonlinear wave equations*, Monographs in Analysis II, Internatl. Press, Boston, MA, 1995
- [54] E. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality and Oscillatory Integrals*, Princeton University Press, Princeton, NJ, 1993
- [55] Sterbenz, J. *Angular regularity and Strichartz estimates for the wave equation*, preprint
- [56] M.Struwe, *Globally regular solutions to the  $u^5$  Klein-Gordon equation.* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 15 (1988), no. 3, 495-513 (1989)
- [57] M.Struwe, *Equivariant Wave Maps in 2 space dimensions*, preprint
- [58] M.Struwe, *Radially Symmetric Wave Maps from  $1+2$  dimensional Minkowski space to the sphere*, Math.Z.242(2002)
- [59] M.Struwe, *Radially symmetric wave maps from  $(1 + 2)$ -dimensional Minkowski space to general targets.*, Calc. Var. Partial Differential Equations 16 (2003), no. 4, 431-437.
- [60] T.Tao, *Ill-posedness for one-dimensional Wave Maps at the critical regularity*, Am. Journal of Math.122 No.3(200), 451-463
- [61] T.Tao, *Global regularity of wave maps I*, I.M.R.N. 6(2001), 299-328
- [62] T.Tao, *Global regularity of wave maps II*, Comm.Math.Phys.224(2001), 443-544
- [63] T.Tao, *Counterexamples to the  $n=3$  endpoint Strichartz estimate for the wave equation*, preprint
- [64] T.Tao, *Geometric renormalization of large energy wave maps*, Journees Equations aux derives partielles, Forges les Eaux, 7-11 June 2004, XI 1-32
- [65] T.Tao, *Nonlinear dispersive equations: local and global analysis*, CBMS Regional Conference Series in Mathematics no. 106(2006), 373 pp
- [66] D.Tataru, *Local and global results for wave maps I*, Comm. PDE 23(1998), 1781-1793
- [67] D.Tataru, *On global existence and scattering for the wave maps equation*, Amer. Journal. Math.123(2001), no.1, 37-77

- [68] D. Tataru, *Rough solutions for the Wave Maps equation*, preprint
- [69] K. Uhlenbeck, *Connections with  $L^p$ -bounds on curvature*, CMP 83(1982), no. 1, 31-42
- [70] T. Wolff, *A sharp bilinear cone restriction estimate.*, Ann. of Math. (2) 153 (2001), no. 3, 661–698.

HARVARD UNIVERSITY, DEPT. OF MATHEMATICS, SCIENCE CENTER, 1 OXFORD STREET, CAMBRIDGE, MA 02138, USA  
*email:* jkrieger at math.harvard.edu

HARVARD UNIVERSITY, DEPT. OF MATHEMATICS, SCIENCE CENTER, 1 OXFORD STREET, CAMBRIDGE, MA 02138, U.S.A.  
*E-mail address:* jkrieger@math.harvard.edu