

## THE WAVE MAPS EQUATION

DANIEL TATARU

ABSTRACT. The wave maps equation has become a very popular topic in recent years. The aim of these expository notes is to present a non-technical survey of the ideas and methods which have proved useful in the study of wave maps, leading up to the latest results. The remaining open problems are also stated and explained.

### 1. THE EQUATIONS

Let us begin with the Laplace equation,

$$-\Delta\phi = 0, \quad \phi : \mathbb{R}^n \rightarrow \mathbb{R}.$$

The solutions of this equation are called harmonic functions, and they can be thought of as critical points for the Lagrangian

$$L^e(\phi) = \frac{1}{2} \int_{\mathbb{R}^n} |\nabla\phi(x)|^2 dx.$$

The same equation is obtained when we consider vector valued functions, i.e. if we replace the target space  $\mathbb{R}$  by  $\mathbb{R}^m$ .

The problem, however, becomes much richer and considerably more interesting if instead of  $\mathbb{R}^m$  we consider a Riemannian manifold  $(M, g)$ . Then

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

and its first order derivatives take values in the tangent space of  $M$ ,

$$\partial_\alpha\phi : \mathbb{R}^n \rightarrow T_\phi M, \quad \alpha = \overline{1, n}.$$

Thus they are sections of the pull-back bundle

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

To measure the size of  $\nabla\phi$  it is natural to use the metric  $g$ ; therefore the modified Lagrangian has the form

$$L_M^e(\phi) = \frac{1}{2} \int_{\mathbb{R}^n} |\nabla\phi(x)|_g^2 dx.$$

In order to describe the corresponding Euler-Lagrange equations we need to use covariant differentiation. To do so, associate to the metric  $g$  the natural covariant

---

Received by the editors May 10, 2003, and, in revised form, September 28, 2003.

2000 *Mathematics Subject Classification*. Primary 35L70.

Lecture presented at the AMS Special Session on Mathematical Current Events: Expository Reports in Baltimore, MD, January 17, 2003.

differentiation on  $TM$  defined by its Levi-Civita connection. This induces a connection  $\mathbf{D}$  on the pull-back bundle  $\phi^*(TM)$ . If  $V$  is a section of  $\phi^*(TM)$ , then we set

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

With this notation, the Euler-Lagrange equations have the form

$$(1) \quad -\mathbf{D}_\alpha \partial_\alpha \phi = 0,$$

with the usual summation convention.

To understand the type of this equation it is useful to write it in local coordinates on  $M$ . By doing so, the equation assumes the form

$$-\Delta \phi^i = \Gamma_{jk}^i(\phi) \partial_\alpha \phi^j \partial_\alpha \phi^k,$$

where  $\Gamma$  is the Riemann-Christoffel symbol on  $M$ . This equation is semilinear elliptic, and it is called the *harmonic maps equation*. Its solutions are called *harmonic maps*. The harmonic maps have their own story (see [12] and references therein) and at some point also have had an impact on the study of wave maps.

We now switch to the wave equation,

$$\square \phi = 0, \quad \phi : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$$

where  $\square$  is the d'Allembertian,  $\square = \partial_t^2 - \Delta_x$ . On the Minkowski space  $\mathbb{R} \times \mathbb{R}^n$  we have the pseudo-Riemannian metric

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

By lifting indices with respect to this metric, we can rewrite the wave equation in the form

$$-\partial^\alpha \partial_\alpha \phi = 0$$

where  $\partial^t = -\partial_t$  and  $\partial^x = \partial_x$ .

Notice that we could also interpret the wave equation as the Euler-Lagrange equation for the Lagrangian

$$L^h(u) = \frac{1}{2} \int_{\mathbb{R} \times \mathbb{R}^n} -|\partial_t \phi|^2 + |\nabla_x \phi|^2 dt dx,$$

but this is far less useful than in the elliptic case because it lacks coercivity. If, instead, we consider functions

$$\phi : \mathbb{R} \times \mathbb{R}^n \rightarrow M,$$

then the corresponding Lagrangian has the form

$$L_M^h(\phi) = \int_{\mathbb{R} \times \mathbb{R}^n} -|\partial_t \phi|_g^2 + |\nabla_x \phi|_g^2 dt dx.$$

Its Euler-Lagrange equation has the form

$$(2) \quad \mathbf{D}^\alpha \partial_\alpha \phi = 0,$$

and it is called the *wave maps equation*. In local coordinates it has the form

$$(3) \quad \square \phi^i = \Gamma_{jk}^i(\phi) \partial^\alpha \phi^j \partial_\alpha \phi^k$$

where  $\Gamma_{jk}^i$  are the usual Christoffel symbols. This expression can be used only as long as  $\phi$  is continuous. Still, it shows that the wave maps equation is essentially a semilinear wave equation.

There are many reasons to study the wave maps. On one hand, this is the simplest geometric nonlinear wave equation, so it provides an excellent setup where

one can begin the study of nonlinear wave interactions. On the other hand, wave maps arise in the analysis of the more difficult hyperbolic Yang-Mills equations either as special cases or as equations for certain families of gauge transformations. They arise as well in general relativity for spacetimes with two Killing vector fields.

In the analysis of nonlinear problems one often considers the linearized equations, which in our case can be easily computed:

$$(4) \quad \mathbf{D}^\alpha \mathbf{D}_\alpha \psi + R(\psi, \partial^\alpha \phi) \partial_\alpha \phi = 0$$

where  $R$  is the Riemann curvature tensor. In local coordinates this expression has a representation which is similar to the equation (3) for  $\phi$ , only somewhat longer.

A special case of the wave maps equation is when  $M$  is a submanifold of  $\mathbb{R}^m$  with the euclidean metric. In this setting the second fundamental form plays a role. In our case it is convenient to interpret it as a symmetric quadratic form on the tangent space with values into the normal space,

$$S_p : TM \times TM \rightarrow NM, \quad p \in M \\ \langle S(X, Y), N \rangle = \langle \partial_X Y, N \rangle.$$

Here the wave maps equation has the form

$$\square \phi = -S_\phi(\partial^\alpha \phi, \partial_\alpha \phi).$$

In particular if  $M$  is a sphere  $M = S^{n-1}$ , then the second fundamental form is

$$S_\phi(X, Y) = -\phi \langle X, Y \rangle$$

and the wave maps equation becomes

$$\square \phi = \phi \langle \partial^\alpha \phi, \partial_\alpha \phi \rangle.$$

Another interesting special case is when  $M$  is the hyperbolic space  $\mathbb{H}^m$ . We can think of  $\mathbb{H}^m$  as the space-like hyperboloid

$$\phi_0^2 = 1 + \phi_1^2 + \cdots + \phi_m^2$$

in the Minkowski space  $\mathbb{R} \times \mathbb{R}^m$  with the Lorentzian metric

$$(ds)^2 = -(dt)^2 + (dx_1)^2 + \cdots + (dx_n)^2.$$

On the tangent space  $T\mathbb{H}^m$  we use the induced inner product

$$\langle X, Y \rangle_L = -X_0 Y_0 + X_1 Y_1 + \cdots + X_m Y_m.$$

The second fundamental form is

$$S_\phi(X, Y) = \phi \langle X, Y \rangle_L,$$

and the wave maps equation becomes

$$\square \phi = -\phi \langle \partial^\alpha \phi, \partial_\alpha \phi \rangle_L.$$

Finally, let us turn our attention to a special class of solutions. Let  $M_1$  be a submanifold of  $M$  and ask the following question. Are the wave maps into  $M_1$  also wave maps into  $M$ ? It is not difficult to see that the answer is affirmative iff  $M_1$  is a totally geodesic submanifold of  $M$ ; i.e. the  $M_1$  geodesics are also  $M$  geodesics.

A special case when this happens is if  $M_1$  is a geodesic  $\gamma$  of  $M$ . Then  $\gamma$  has one dimension and no curvature. Hence, with respect to the arclength parametrization, the wave maps equation into  $\gamma$  is nothing but the linear wave equation. Thus for any target manifold  $M$  we have at our disposal a large supply of wave maps associated to the geodesics of  $M$ .

## 2. THE PROBLEMS

Consider the Cauchy problem for the wave maps equation

$$(5) \quad \begin{cases} \mathbf{D}^\alpha \partial_\alpha \phi = 0 & \text{in } \mathbb{R} \times \mathbb{R}^n \\ \phi(0, x) = \phi_0(x), \quad \partial_t \phi(0, x) = \phi_1(x) & \text{in } \mathbb{R}^n \end{cases}$$

The initial data  $(\phi_0, \phi_1)$  must be chosen so that

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

We begin with two useful observations:

**1. Energy.** The wave maps equation has a conserved energy, namely the quadratic functional

$$E(\phi) = \frac{1}{2} \int_{\mathbb{R}^n} |\partial_t \phi|_g^2 + |\partial_x \phi|_g^2 dx.$$

**2. Scaling.** The wave maps equation is invariant with respect to the dimensionless scaling

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

Note however that the energy is scale invariant only in dimension  $n = 2$ .

**2.1. Local well-posedness.** The natural problem is to consider initial data in Sobolev spaces

$$(\phi_0, \phi_1) \in H^s(\mathbb{R}^n) \times H^{s-1}(\mathbb{R}^n)$$

and to seek a solution

$$\phi \in C([-T, T]; H^s(\mathbb{R}^n)), \quad \partial_t \phi \in C([-T, T]; H^{s-1}(\mathbb{R}^n))$$

with a lifespan  $T$  which depends on the initial data or, perhaps, only on the size of the initial data. This is fairly easy to do if  $s$  is large enough, but it becomes increasingly difficult to accomplish as  $s$  decreases in value.

For low  $s$  one has difficulties even with the definition of the Sobolev spaces. What is the meaning of  $H^s(\mathbb{R}^n)$  for functions which take values into a manifold? The reason this question is easier to answer for  $s > \frac{n}{2}$  is that the  $H^s$  functions are continuous; therefore locally the image of  $\phi$  is contained in the domain of a local map for  $M$ . Thus one can measure the regularity of  $\phi$  using the corresponding local coordinates on  $M$ .

If  $s \leq \frac{n}{2}$ , the problem becomes nonlocal and the answer may depend on global properties of the manifold  $M$ . If, for instance,  $M$  embeds isometrically into  $\mathbb{R}^m$ , then one might use this to define the space of  $H^s$  functions with values in  $M$  for all  $s \geq 0$ . To do so, however, one needs to ask whether these spaces depend on the isometric embedding or not.

Indeed, the problem is that topological information is lost whenever  $s < \frac{n}{2}$ . Consider for instance the case when  $M = S^n$  and ask how many times does an  $S^n$  valued  $H^s$  function wrap around the sphere. As shown by Brezis-Nirenberg [2], this rotation number is well defined for  $s = \frac{n}{2}$ , but not for  $s < \frac{n}{2}$ .

Another piece of information comes from scaling. By rescaling the equation one can balance the size of the initial data and the lifespan of the solution. The initial data space<sup>1</sup> is scale invariant if  $s = \frac{n}{2}$ . Then one has the following relations

<sup>1</sup>In order to take advantage of scaling one should really use homogeneous Sobolev spaces, but this issue turns out not to be important unless  $s = n/2$ .

concerning local well-posedness:

$$\begin{array}{lll}
 s > \frac{n}{2} & \begin{array}{l} \text{small data} \\ \text{large time} \end{array} & \Leftrightarrow \begin{array}{l} \text{large data} \\ \text{small time} \end{array} \\
 s = \frac{n}{2} & \begin{array}{l} \text{small data} \\ \text{small time} \end{array} & \Leftrightarrow \begin{array}{l} \text{small data} \\ \text{large time} \end{array} \\
 s < \frac{n}{2} & \begin{array}{l} \text{small data} \\ \text{small time} \end{array} & \Leftrightarrow \begin{array}{l} \text{large data} \\ \text{large time} \end{array}
 \end{array}$$

Given the above considerations, the first problem one is led to is

**Problem 1.** Prove local well-posedness for  $s > \frac{n}{2}$ .

As  $s$  decreases toward  $\frac{n}{2}$  one gains better information concerning the lifespan of solutions. For  $s = \frac{n}{2}$  a local result yields a global result, but one needs to distinguish between small and large data. Hence it makes sense to consider

**Problem 2.** Under reasonable assumptions on  $M$  prove global well-posedness for small data and  $s = \frac{n}{2}$ .

The need for assumptions on  $M$  stems from the fact that for  $s = \frac{n}{2}$  the problem becomes nonlocal. For instance, by using the special solutions contained on geodesics, one easily sees that  $M$  needs to be geodesically complete. It is likely that other assumptions may also be needed.

An alternate venue which was pursued was to obtain weak global energy solutions using various penalization techniques; see for instance [25], [22], [6]. It would be interesting to understand whether in two dimensions these must coincide with the unique solutions obtained later in better classes of functions.

Finally, for  $s < \frac{n}{2}$  a local result seems unreasonable except perhaps for very special cases. Thus one is led to consider the following:

**Problem 3.** Prove ill-posedness for  $s < \frac{n}{2}$ .

This was solved in [5] using the special solutions which are contained on geodesics and will not be discussed further.

**2.2. Global solutions vs. blowup.** A local well-posedness result at scaling yields a global result for small data. However, the question of what happens for large data is of a different nature. Experience indicates that in general large data global well-posedness results can be obtained only by exploiting conserved or decreasing “energy” functionals.

For the wave maps equation the only conserved quantity that we know is the energy (and variations of it, e.g. momentum). Unfortunately in dimension  $n \geq 3$  the energy is below scaling, and this essentially renders it useless. We call such problems *supercritical*:

**Problem 4.** Prove that large data solutions for the wave maps equation can blow up in finite time if  $n \geq 3$ .

Indeed, for a reasonably large class of manifolds  $M$  self-similar blow-up solutions were constructed in [25], [28], [3]. It is, however, interesting to note that for positively curved target manifolds  $M$  (e.g. the sphere) such counterexamples have been obtained for all  $n \geq 3$ , while for negatively curved manifolds the current counterexamples apply only to dimensions  $n \geq 7$ .

The other easier case is in dimension  $n = 1$ , where the energy is above scaling. This problem is called *subcritical*:

**Problem 5.** Prove that large data solutions for the wave maps equation are global if  $n = 1$ .

This is true and was proved in [9], [7], [20]; see also [14] for related work.

It remains to discuss the most difficult case,  $n = 2$ . Here the energy is precisely at scaling, which makes it very difficult to exploit. We call this problem *critical*.

Suppose that  $M$  is a manifold for which the wave maps equation is known to be well-posed for small data in the energy space  $H^1 \times L^2$ . Combining this with the finite speed of propagation, the large data problem reduces to a nonconcentration argument, which asserts that energy cannot focus at the tip of a light cone. Such nonconcentration arguments are known for other critical semilinear wave equations (see [26] and references therein), but none are available yet for wave maps.

Furthermore, numerical evidence in [13] and [1] indicates that the outcome seems to depend on the geometry of  $M$ . For the sphere blow-up appears to occur above a certain energy threshold, while for the hyperbolic space no such phenomena occurs.

**Open Problem 6.** Consider the wave maps equation in  $2 + 1$  dimensions. Prove that

- a) blow-up of large energy solutions occurs for certain target manifolds (e.g. the sphere),
- b) large energy solutions are global for other target manifolds (e.g. the hyperbolic space).

Results of this type have been previously obtained only for special classes of solutions; see e.g. [4].

**2.3. The main result.** The rest of these notes is devoted to the following result:

**Theorem 7.** [32] *Let  $n \geq 2$ . For all “reasonable” target manifolds  $M$  the wave maps equation is globally well-posed for initial data which is small in  $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$ .*

This is most interesting, but also most difficult, in  $2 + 1$  dimensions. On one hand, it shows that if the initial data is smooth and has small energy, then the solution cannot develop singularities in finite time; therefore it remains smooth globally in time. On the other hand, it also shows that for any small energy data there is a meaningful notion of solution, which can be defined as the unique limit of smooth solutions. This happens despite the fact that such rough solutions cannot in general be continuous.

This result was proved in a recent article [32] by the author, but this proof would not have been possible without earlier work in [16], [18], [34] and [31].

Here a “reasonable” target manifold is a manifold which admits a uniform isometric embedding into a Euclidean space  $\mathbb{R}^m$ . By Nash’s theorem [24] (see also the more recent work in [8], [10]) such an embedding exists for all smooth compact manifolds and also for a large class of smooth unbounded manifolds.

In the next two sections we outline the evolution of ideas which led to the above result, and we attempt to provide the reader with a road map for the proof. The actual proof is quite technical, and it would be hopeless to try to present it here. Hopefully, significant simplifications can be achieved given sufficient time.

## 3. WAVE MAPS AS A SEMILINEAR EQUATION

Since in local coordinates the wave maps equation looks like a semilinear wave equation, it is natural to try to treat it as such. The usual approach is to consider the nonlinear term as a small perturbation of the linear operator. Suppose, more generally, that we are trying to solve a semilinear wave equation

$$(6) \quad \square\phi = N(\phi), \quad \phi(0) = \phi_0, \quad \partial_t\phi(0) = \phi_1$$

with initial data  $(\phi_0, \phi_1)$  in some Sobolev space  $X_0$ . Then the idea is to consider the nonlinear term as a small perturbation with respect to the linear equation and to absorb the nonlinearity in a fixed point argument.

We introduce the operators  $H$  and  $\square^{-1}$  which give the solution  $\phi$  for the linear equation

$$(7) \quad \square\phi = f, \quad \phi(0) = \phi_0, \quad \partial_t\phi(0) = \phi_1$$

in the form

$$\phi = H(\phi_0, \phi_1) + \square^{-1}f.$$

There are two cases to consider:

**(a)  $X_0$  is above scaling.** The approach here is to seek some Banach spaces  $X$  for the solution  $\phi$  and  $Y$  for the nonlinearity with the following properties:

(a1) (linear mapping property) The solution  $\phi$  to (7) satisfies

$$\|\chi(t)\phi\|_X \lesssim \|(\phi_0, \phi_1)\|_{X_0} + \|f\|_Y$$

for any smooth compactly supported  $\chi$ .

(a2) The truncated nonlinear term  $\chi(t)N$  has the mapping property

$$\chi(t)N : X \rightarrow Y, \quad \text{Lipschitz continuous.}$$

If this is done, we can use the contraction principle in  $X$  to solve the nonlinear problem

$$\phi = H(\phi_0, \phi_1) + \square^{-1}(\chi(t)N(\phi)).$$

If  $\chi = 1$  in some time interval  $[-T, T]$ , then the function  $\phi$  must also solve (6) in  $[-T, T]$ . Thus we obtain

(a3) For each initial data  $(\phi_0, \phi_1) \in X_0$  there is some time  $T$  which depends only on the size of the initial data and a function  $\phi$  in  $X$  which solves the equation in  $[-T, T]$  and has a Lipschitz dependence on the initial data.

**(b)  $X_0$  is scale invariant.** Then we seek some Banach spaces  $X$  for the solution  $\phi$  and  $Y$  for the nonlinearity which are compatible with the scaling and have the following properties:

(b1) (linear mapping property) The solution  $\phi$  to (7) satisfies

$$\|\phi\|_X \lesssim \|(\phi_0, \phi_1)\|_{X_0} + \|f\|_Y$$

(b2) The nonlinear term  $N$  maps

$$N : X \rightarrow Y \quad \text{Lipschitz continuous.}$$

If this is done, then we recast the equation (6) in the form

$$\phi = H(\phi_0, \phi_1) + \square^{-1}N(\phi)$$

which we solve in  $X$  using the contraction principle. This yields

(b3) For each small initial data  $(\phi_0, \phi_1) \in X_0$  there is a unique global solution  $\phi$  in  $X$  which solves the equation and depends smoothly on the initial data.

The small initial data is needed in order to gain the small Lipschitz constant in the fixed point argument. This is different from problems above scaling, where the small Lipschitz constant is usually obtained by shortening the time interval  $[-T, T]$ .

Next we consider possible strategies for implementing this approach in the case of the wave maps equation.

**3.1. Energy estimates.** In this case the choice of spaces is made based on the energy estimates for the linear equation (7), namely

$$\|\nabla\phi\|_{L^\infty(H^{s-1})} \lesssim \|\phi_0\|_{H^s} + \|\phi_1\|_{H^{s-1}} + \|f\|_{L^1H^{s-1}}.$$

This corresponds to choosing

$$X_0 = H^s \times H^{s-1}, \quad X = C(H^s) \cap C^1H^{s-1}, \quad Y = L^1H^{s-1}.$$

The nonlinear mapping property (a2) yields the constraint

$$s > \frac{n}{2} + 1$$

which is quite far from scaling.

**3.2. Strichartz estimates.** The nonlinear estimates can be improved if one takes advantage of the dispersive effect of the wave equation. The idea is that solutions to the wave equation cannot stay spatially concentrated for a long time; instead they must spread out and decay. A quantitative way of measuring this is provided by the Strichartz estimates. For solutions to (7) they yield

$$\|\nabla\phi\|_{L^\infty(H^{s-1})} + \|\nabla\phi\|_{L^4L^\infty} \lesssim \|\phi_0\|_{H^s} + \|\phi_1\|_{H^{s-1}} + \|f\|_{L^1H^{s-1}}$$

for  $n = 2$ ,  $s = \frac{5}{4}$ , respectively

$$\|\nabla\phi\|_{L^\infty(H^{s-1})} + \|\nabla\phi\|_{L^2L^\infty} \lesssim \|\phi_0\|_{H^s} + \|\phi_1\|_{H^{s-1}} + \|f\|_{L^1H^{s-1}}$$

for<sup>2</sup>  $n \geq 3$ ,  $s = \frac{n+1}{2}$ .

Then we modify the choice of  $X$  to

$$X = \{\phi \in C(H^s), \nabla\phi \in C(H^{s-1}) \cap L^2L^\infty\},$$

and we gain the improved range

$$s \geq \frac{5}{4} \quad (n = 2), \quad s > 2 \quad (n = 3), \quad s \geq \frac{n+1}{2} \quad (n > 3).$$

**3.3. The null condition.** So far we have treated the wave maps equation as a generic semilinear wave equation of the form

$$\square\phi = (\nabla\phi)^2.$$

Unfortunately, for such equations a counterexample of Lindblad [21] shows that the Strichartz estimates actually give the sharp result at least in low dimension ( $n = 2, 3$ ). In this counterexample, blow-up occurs due to concentration along a light ray or, in other words, due to nonlinear interaction of waves which travel in the same direction.

The first key insight into wave maps came from Klainerman, who observed that the quadratic form which occurs in the wave maps equation is not generic; instead it has a special structure which he called the null condition. Exploiting this observation Klainerman was able to prove in [15] that for  $n \geq 3$  solutions with small,

<sup>2</sup>While false for  $n = 3$  it is almost true.

smooth and compactly supported data are global. In two dimensions, the proof was more difficult, but the same result was proved later by Sideris [29].

To explain the meaning of the null condition it is useful to switch to the Fourier space. We denote by  $\tau$  the time Fourier variable and by  $\xi$  the space Fourier variable. The symbol of the wave operator is

$$p(\tau, \xi) = \tau^2 - \xi^2;$$

therefore solutions to the wave equation are concentrated in frequency near the characteristic cone

$$K = \{\tau^2 - \xi^2 = 0\}.$$

A solution to the wave equation which is concentrated in frequency near some  $(\tau, \xi) \in K$  travels roughly in the normal direction to the cone,

$$\nabla_{\tau, \xi} p(\tau, \xi) = (2\tau, -2\xi).$$

Consider now two waves  $\phi^{(1)}$  and  $\phi^{(2)}$  which are concentrated in frequency near  $(\tau^{(1)}, \xi^{(1)}) \in K$ , respectively  $(\tau^{(2)}, \xi^{(2)}) \in K$ . Their product  $\phi^{(1)}\phi^{(2)}$  has Fourier transform

$$\widehat{\phi^{(1)}\phi^{(2)}} = \int_{\mathbb{R}^{n+1}} \widehat{\phi^{(1)}}(s, \eta) \widehat{\phi^{(2)}}(\tau - s, \xi - \eta) ds d\eta$$

which is concentrated in frequency near

$$(\tau, \xi) = (\tau^{(1)} + \tau^{(2)}, \xi^{(1)} + \xi^{(2)}).$$

If we consider the wave equation

$$\square\phi = \phi^{(1)}\phi^{(2)},$$

the largest input comes from  $(\tau, \xi)$  near the cone  $K$ . But the only way we can have

$$(\tau, \xi) \in K, \quad (\tau^{(1)}, \xi^{(1)}) \in K, \quad (\tau^{(2)}, \xi^{(2)}) \in K$$

is if all three vectors are parallel. This computation indicates that, for the wave equation with generic quadratic nonlinearities, the worst nonlinear interaction is between waves which travel in the same direction. A partial physical space interpretation of the same phenomena is that waves which travel in different directions intersect on a smaller set and therefore have less interaction.

Now replace the product  $\phi^{(1)}\phi^{(2)}$  by the quadratic form arising in the wave maps equation,

$$Q_0(\phi^{(1)}, \phi^{(2)}) = \partial^\alpha \phi^{(1)} \partial_\alpha \phi^{(2)}.$$

Its Fourier transform is given by

$$\mathcal{F}Q_0(\phi^{(1)}\phi^{(2)}) = \int_{\mathbb{R}^{n+1}} q_0(s, \eta, \tau - s, \xi - \eta) \widehat{\phi^{(1)}}(s, \eta) \widehat{\phi^{(2)}}(\tau - s, \xi - \eta) ds d\eta$$

where the symbol  $q_0$  of the quadratic form  $Q_0$  is given by

$$q_0(\tau^{(1)}, \xi^{(1)}, \tau^{(2)}, \xi^{(2)}) = \tau^{(1)}\tau^{(2)} - \xi^{(1)}\xi^{(2)}.$$

The interesting observation is that  $q_0$  vanishes if both  $(\tau^{(1)}, \xi^{(1)})$  and  $(\tau^{(2)}, \xi^{(2)})$  are on the cone and collinear. This is a consequence of the relation

$$2q_0(\tau^{(1)}, \xi^{(1)}, \tau^{(2)}, \xi^{(2)}) = p(\tau^{(1)} + \tau^{(2)}, \xi^{(1)} + \xi^{(2)}) - p(\tau^{(1)}, \xi^{(1)}) - p(\tau^{(2)}, \xi^{(2)}).$$

Thus the worst kind of bilinear interaction cannot occur in the wave maps equation. Then we say that this nonlinearity satisfies the null condition.

**3.4. The  $X^{s,b}$  spaces.** How can one choose the spaces  $X$  and  $Y$  in order to best take advantage of the null condition? One answer to this is provided by the  $X^{s,b}$  spaces introduced by Klainerman and Machedon. These spaces are associated to the wave equation essentially in the same way the  $H^s$  spaces are associated to the Laplacian. More precisely we define

$$\|u\|_{X^{s,b}} = \|(1 + |\tau| + |\xi|)^s (1 + ||\tau| - |\xi||)^b \hat{u}(\tau, \xi)\|_{L^2}.$$

The index  $s$  accounts for regular derivatives, while  $b$  corresponds to “wave” derivatives, i.e. powers of the wave operator.

It is easy to see that

$$\square : X^{s,b} \rightarrow X^{s-1,b-1}$$

and that the wave equation parametrix  $\square^{-1}$  has the inverse mapping properties

$$\square^{-1} : X_{comp}^{s-1,b-1} \rightarrow X_{loc}^{s,b}.$$

Furthermore,  $H^s$  solutions to the homogeneous wave equation are in  $X^{s,b}$  for all  $b$ . Hence we seek spaces  $X, Y$  of the form

$$X = X^{s,b}, \quad Y = X^{s-1,b-1}.$$

The choice of  $b$  is related to scaling, and the optimal value is  $b = \frac{1}{2}$ .

Given the way the  $X^{s,b}$  spaces are defined, the estimates (a2) for the nonlinearity reduce to weighted convolution estimates in the Fourier space. This venue was pursued by Klainerman-Machedon [16], who proved local well-posedness for the wave maps equation in  $H^s \times H^{s-1}$  for all  $s > \frac{n}{2}$ ,  $n \geq 3$ . The similar result in the more difficult case  $n = 2$  was obtained shortly afterward by Klainerman-Selberg [18].

**3.5. Scale invariant results.** Can the same method be used in the scale invariant case  $s = \frac{n}{2}$ ? Then the spaces  $X, Y$  should be compatible with scaling, which would appear to imply that we should use the homogeneous version of the  $X^{s,b}$  spaces, namely

$$\begin{aligned} X &= \dot{X}^{\frac{n}{2}, \frac{1}{2}} = \{u; (|\tau| + |\xi|)^{\frac{n}{2}} ||\tau| - |\xi||^{\frac{1}{2}} \hat{u} \in L^2\}, \\ Y &= \dot{X}^{\frac{n}{2}-1, -\frac{1}{2}} = \{u; (|\tau| + |\xi|)^{\frac{n}{2}-1} ||\tau| - |\xi||^{-\frac{1}{2}} \hat{u} \in L^2\}. \end{aligned}$$

However,  $X$  is not well defined as a space of distributions, while  $Y$  does not even contain all the smooth compactly supported functions. Any naive attempt to fix one problem worsens the other.

Consider the nonlinear estimate one needs to prove:

$$f(X) \partial^\alpha X \partial_\alpha X \rightarrow Y.$$

We can split this into bilinear pieces. Clearly  $X$  should be an algebra; in addition, two more multiplicative properties are required:

$$(8) \quad \begin{aligned} X \cdot X &\rightarrow X, \\ X \cdot Y &\rightarrow Y, \\ \partial^\alpha X \cdot \partial_\alpha X &\rightarrow Y. \end{aligned}$$

A standard technique in harmonic analysis is to use a Littlewood-Paley decomposition to break up functions into dyadic pieces, i.e. in parts which are frequency localized in dyadic regions. For  $j \in \mathbb{Z}$  let us loosely denote by  $X_j$ , respectively

$Y_j$ , the spaces of functions in  $X$ , respectively  $Y$ , which have Fourier transform supported in the region

$$\{|\xi| + |\tau| \in [2^{j-1}, 2^{j+1}]\}.$$

The structure of  $X_j, Y_j$  is determined by that of  $X, Y$ , but the converse is also true. Indeed, on one hand, the rescaling must map one  $X^j$  into another. On the other hand, the square summability with respect to  $j$  is inherited from the initial data. Thus, we have

$$(9) \quad \|\phi\|_X^2 \approx \sum_{j \in \mathbb{Z}} \|P_j \phi\|_{X_j}^2$$

where  $P_j = P_j(D)$  are multipliers which select the frequency  $\approx 2^j$  part of  $\phi$ .

A preliminary step in obtaining (8) is to prove their dyadic counterparts. When we multiply two frequency localized functions, there are two possible types of interaction.

*a) High-high interactions.* If we multiply two functions of comparable frequency, the product is at the same frequency or lower. Thus we have dyadic estimates of the form

$$(10) \quad \begin{aligned} P_j(X_k \cdot X_k) &\rightarrow X_j, \\ P_j(X_k \cdot Y_k) &\rightarrow Y_j, \quad j \leq k \\ P_j(\partial^\alpha X_k \cdot \partial_\alpha X_k) &\rightarrow Y_j. \end{aligned}$$

*b) High-low interactions.* If we multiply a high frequency function by a low frequency one, the product is at high frequency. Hence the dyadic estimates are

$$(11) \quad \begin{aligned} X_j \cdot X_k &\rightarrow X_k, \\ X_j \cdot Y_k &\rightarrow Y_k, \quad X_k \cdot Y_j \rightarrow Y_k \quad j < k \\ \partial^\alpha X_j \cdot \partial_\alpha X_k &\rightarrow Y_k. \end{aligned}$$

Aside from the bilinear dyadic estimates, we also want to have the linear mapping property for the parametrix of the wave equation,

$$(12) \quad \square^{-1} : Y_j \rightarrow X_j.$$

If one attempts to produce function spaces  $X, Y$  which have all the desired properties, it soon becomes clear that there are two major hurdles to overcome.

*i) The division problem,* which is to find dyadic spaces  $X_j$  satisfying all the dyadic estimates in (10), (11) and (12). The name comes from the fact that in the Fourier space the parametrix  $\square^{-1}$  for the wave equation is essentially the division by  $\tau^2 - \xi^2$ . The function  $(\tau^2 - \xi^2)^{-1}$  is not locally integrable, and without considerable care this generates logarithmic divergences in the estimates.

*ii) The summation problem,* which is to somehow be able to sum up the dyadic estimates (10), (11) in order to obtain (8). A priori, in both high-high and high-low interactions there is a logarithmic divergence in the summation with respect to the index  $k$ .

The division problem was solved in two consecutive articles by the author. The first one, [33], applies to high dimension ( $n \geq 4$ ) and uses function spaces  $X, Y$  which are essentially modified  $X^{s,b}$  spaces combined with Strichartz type norms. The second one [34] applies in low dimension and is much more involved. This time

an essential part of the  $X$  and  $Y$  spaces consists of functions tied to rotating characteristic frames which are frequency localized in certain neighborhoods of sectors on the characteristic cone.

In that setup there is an additional gain of  $2^{-\epsilon|j-k|}$  which insures the  $k$  summation for the high-high interactions. However, the  $k$  summation in the high-low interactions remains a problem. In [33], [34] this is avoided by replacing the  $l^2$  dyadic summation in (9) with an  $l^1$  summation. The main result has the form:

**Theorem 8.** [33], [34] *The wave maps equation is well-posed for initial data which is small in the homogeneous Besov space  $\dot{B}_{2,1}^{\frac{n}{2}} \times \dot{B}_{2,1}^{\frac{n}{2}-1}$ .*

On the positive side, the above Besov space is scale invariant; therefore the solutions which are obtained are global and depend smoothly on the initial data.

On the negative side, the Besov space is smaller than the corresponding Sobolev space. Due to the embedding  $\dot{B}_{2,1}^{\frac{n}{2}} \subset L^\infty$  the solutions obtained are small in  $L^\infty$ . Hence the problem becomes local with respect to the target manifold  $M$ ; this means that the geometry of  $M$  does not come into the picture.

One might ask whether it is possible to also solve the summation problem by making a better choice of the function spaces  $X, Y$ . At the time the answer was not clear. However, we now know that this is not the case; indeed, the results in [5] show that there is no uniformly continuous dependence of the solutions on the initial data in  $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$ . Thus the semilinear approach to wave maps can go no further than this.

#### 4. WAVE MAPS AS A NONLINEAR EQUATION

The main idea of this section is that at scaling the wave maps equation behaves like a genuine nonlinear wave equation. As in the semilinear case, we begin with a presentation of the common approach for nonlinear waves.

**4.1. Local solutions for the nonlinear wave equation.** Suppose that we have a nonlinear wave equation

$$(13) \quad P(\phi, \partial\phi, \partial^2\phi) = 0, \quad \phi(0) = \phi_0, \quad \partial_t\phi(0) = \phi_1.$$

**Definition 9.** The problem (13) is well-posed in  $H^{s_0} \times H^{s_0-1}$  if:

For each  $M > 0$  there is some  $T > 0$  so that:

(i) (*a priori bound for smooth solutions*) For each smooth initial data  $(\phi_0, \phi_1)$  with

$$(14) \quad \|(\phi_0, \phi_1)\|_{H^{s_0} \times H^{s_0-1}} \leq M$$

there is a unique smooth solution  $\phi$  in  $[-T, T]$  which satisfies uniform bounds

$$\|\phi\|_{C(-T, T; H^s) \cap C^1(-T, T; H^{s-1})} \leq C \|(\phi_0, \phi_1)\|_{H^s \times H^{s-1}}, \quad s \geq s_0$$

with  $C$  depending only on  $M, s$ .

(ii) (*weak stability estimates*) There is some  $s < s_0$  so that for any two smooth solutions  $\phi, \psi$  whose initial data is subject to (14) we have

$$\|\phi - \psi\|_{C(-T, T; H^s) \cap C^1(-T, T; H^{s-1})} \leq C \|(\phi_0 - \psi_0, \phi_1 - \psi_1)\|_{H^s \times H^{s-1}}$$

for some  $s < s_0$ , with  $C$  depending only on  $M, s$ .

(iii) (*rough solutions as limits of smooth solutions*) For any initial data  $(\phi_0, \phi_1)$  which satisfies (14) there is a solution  $\psi \in C(-T, T; H^{s_0}) \cap C^1(-T, T; H^{s_0-1})$ ,

depending continuously on the initial data, which can be obtained as the unique limit of smooth solutions.

Most notably one loses the Lipschitz dependence on the initial data, which implies that the solution can no longer be obtained via a fixed point argument.

For the wave maps equation, because we are at scaling, there is also another twist:  $M$  must be small while the lifespan  $T$  must be infinite (arbitrarily large). Note that if  $s_0$  is above scaling, then the smallness of  $M$  can be gained by rescaling. Consequently, in what follows we restrict ourselves to small data solutions.

To outline a standard proof of well-posedness we first introduce some notations. Begin with the initial data space, where we set

$$X_0 = H^{s_0} \times H^{s_0-1}, \quad X_0^s = H^s \times H^{s-1}.$$

Next we need to find some Banach space  $X$  which measures the regularity of  $H^{s_0} \times H^{s_0-1}$  solutions. In effect it is better to have a family of spaces  $X_s$  which measure the regularity of  $H^s \times H^{s-1}$  solutions. Usually these spaces are related by multipliers just like the Sobolev spaces,

$$X^s = |D|^{s_0-s} X.$$

One might be tempted to set

$$X^s = C(-T, T; H^s) \cap C^1(-T, T; H^{s-1}),$$

but this is rarely sufficient. As it turns out, one needs spaces which provide a better description of the solutions.

A main difficulty in obtaining estimates for nonlinear equations is that on one hand in order to obtain bounds for the solution one needs bounds for the coefficients, while on the other hand the bounds for the coefficients follow from bounds for the solution.

A way out of this circular argument is provided by a bootstrap lemma of the form

**Lemma 10.** *a) For all smooth solutions  $\phi$  to (13) in  $[-T, T]$  we have*

$$\left. \begin{array}{l} \|\phi[0]\|_{X_0} \ll 1 \\ \|\phi\|_X \leq 2 \end{array} \right\} \implies \|\phi\|_X \leq 1.$$

*b) In addition for  $s > s_0$  and  $\alpha > 0$*

$$\left. \begin{array}{l} \|\phi[0]\|_{X_0^s} \ll \alpha \\ \|\phi\|_X \leq 2 \\ \|\phi\|_{X^s} \leq 2\alpha \end{array} \right\} \implies \|\phi\|_{X^s} \leq \alpha.$$

This is used in a continuity argument as follows. Let  $s$  be large and  $\alpha > 0$  arbitrary. Given some smooth initial data  $(\phi_0, \phi_1)$  satisfying

$$\|\phi[0]\|_{X_0} \ll 1, \quad \|\phi[0]\|_{X_0^s} \ll \alpha,$$

we consider a smooth one parameter family of initial data sets

$$(\phi_0^h, \phi_1^h) \quad h \in [0, 1],$$

uniformly satisfying similar bounds with

$$(\phi_0^0, \phi_1^0) = (0, 0), \quad (\phi_0^1, \phi_1^1) = (\phi_0, \phi_1).$$

Here we assume for simplicity that 0 is a solution and that the linearization at 0 is the linear wave equation. Then for small  $h$  there is a smooth solution  $\phi^h$  in  $[-T, T]$  depending smoothly on  $h$  which satisfies

$$(15) \quad \|\phi^h\|_X \leq 2, \quad \|\phi^h\|_{X_s} \leq 2\alpha.$$

From the lemma we obtain

$$\|\phi^h\|_X \leq 1, \quad \|\phi^h\|_{X_s} \leq \alpha.$$

By continuity it follows that a smooth solution exists and satisfies (15) for a larger set of values of  $h$ . By the same token one argues that the set  $A \subset [0, 1]$  of those  $h$  for which a solution satisfying (15) is both open and closed in  $[0, 1]$ . Hence  $A = [0, 1]$ ; therefore for  $h = 1$  we obtain a solution with initial data  $(\phi_0, \phi_1)$ .

For the weak stability part of the estimates we need to consider the linearized equations

$$(16) \quad P^{lin}(\phi)\psi = 0, \quad \psi(0) = \psi_0, \quad \partial_t\psi(0) = \psi_1$$

around some solution  $\phi$ . These equations have variable coefficients, and for them we need the next lemma:

**Lemma 11.** *There is  $s < s_0$  so that the linearized equations (16) are uniformly well-posed in  $H^{s-1} \times H^s$  for all smooth solutions  $\phi$  to (13) in  $[-T, T]$  with  $\|\phi[0]\|_{X_0} \ll 1$ .*

Note that from the previous lemma we obtain that the solutions  $\psi$  must be bounded in  $X$ , which provides us with some regularity for the coefficients of  $P^{lin}$ . A robust way to prove the lemma is to obtain an estimate of the form

$$(17) \quad \|\psi\|_{X^s} \lesssim \|(\psi_0, \psi_1)\|_{H^s \times H^{s-1}}.$$

**4.2. The paradifferential calculus.** Since it is difficult to directly prove nonlinear estimates, one usually attempts to reduce them to linear estimates. A convenient tool which can be used for that is the paradifferential calculus. Given a Littlewood-Paley decomposition of a solution  $\phi$  to (13), the basic principle is that we can transform the equation into an infinite system

$$(18) \quad P^{lin}(\phi_{<k})\phi_k = \text{error}$$

where  $\phi_{<k}$ , respectively  $\phi_k$ , loosely denotes the part of  $\phi$  which is at frequency less than  $2^{k-10}$ , respectively  $\approx 2^k$ .

The term “error” is intended to mean an acceptable error, i.e. which is small in an appropriate sense. To measure this we need Banach spaces  $Y$ , respectively  $Y_s$ , so that the forward parametrix for  $P^{lin}(\phi_{<k})$  has the linear mapping properties

$$[P^{lin}(\phi_{<k})]^{-1} : Y_k \rightarrow X_k, \quad [P^{lin}(\phi_{<k})]^{-1} : Y_{s,k} \rightarrow X_{s,k}.$$

One advantage in doing this is that the equations (18) are frequency localized. Thus the statements in Lemma 10(a), Lemma 10(b) and Lemma 11 become more or less equivalent after such a transformation.

Another advantage is that we peel off some less important parts of the nonlinearity and retain the main contributions only. This amounts to saying that the low frequency contributions of the high-high frequency interactions is negligible.

**4.3. Wave maps.** Can such a strategy be implemented for the critical wave maps equation? As it turns out, the answer is affirmative. This shows that although the wave maps equation has the form of a semilinear wave equation, its behavior at scaling is genuinely nonlinear.

The first step in this direction was carried out by Tao, initially in high dimension in [30] and then in low dimension [31]. He considered the case of a target which is a sphere and proved only part (i) in Definition 9 of well-posedness (i.e. the wave maps counterpart of the bootstrap Lemma 10).

However, as it turns out, related ideas can be used for more general target manifolds. Klainerman and Rodnianski [17] did this for  $n \geq 5$ , Shatah-Struwe [27] and Nahmod-Stefanov-Uhlenbeck [23] for  $n \geq 4$  and Krieger [19] for the hyperbolic space and  $n = 3$ .

Finally, the result in low dimension ( $n \geq 2$ ) for general target manifolds was proved in a very recent article by the author [32]. Parts (ii), (iii) in Definition 9 of nonlinear well-posedness are also obtained there.

Independently Krieger was also able to extend his results to include general manifolds in dimension  $n = 3$ , as well as the hyperbolic space in dimension  $n = 2$ . At the same time Nahmod-Stefanov-Uhlenbeck made some progress toward extending their results in dimension  $n = 3$ .

To keep the exposition simple we begin the discussion below using Tao's setup of the spherical target. Then we move on to general target manifolds and use the spherical case for comparison.

**4.4. The spaces.** The first step in the proof of Theorem 7 is to construct Banach spaces  $X, Y$  for the solutions, respectively the inhomogeneous terms in the wave equation. Just as in Section 3.5, we omit the description of these spaces because it is too technical.

The spaces used by Tao in [31] and also later in [32] are variations of the spaces introduced in [34]. The  $Y$  space is essentially the same, but the  $X$  space is slightly enlarged in order to gain the key algebra property for  $X \cap L^\infty$ . This already removes the logarithmic divergence in the first relation in (8). However, the logarithmic divergence in the last two relations in (8) is genuinely nonlinear and cannot be removed. Instead, the best one can do is to use the paradifferential calculus.

We note that there seems to be considerable room for improvement in this functional setup. Hopefully, given enough time it is possible that a simpler framework will emerge.

**4.5. Paradifferential calculus and the trilinear estimate.** The paradifferential form of the wave maps equation into the sphere is

$$\square \phi_k^i = 2\phi_{<k}^i \partial^\alpha \phi_{<k}^j \partial_\alpha \phi_k^j + \text{error}.$$

However, this is not entirely satisfactory. Instead one uses the equation for the sphere to get  $\phi^2 = 1$  which yields  $\phi^j \partial_\alpha \phi^j = 0$ . In paradifferential form this gives good control over the product  $\phi_{<k}^j \partial_\alpha \phi_k^j$ . Thus we replace the above equation with

$$(19) \quad \square \phi_k^i = 2(\phi_{<k}^i \partial^\alpha \phi_{<k}^j - \phi_{<k}^j \partial^\alpha \phi_{<k}^i) \partial_\alpha \phi_k^j + \text{error}.$$

In doing this, one obtains antisymmetric matrices

$$(A_{<k}^\alpha)^{ij} = (\phi_{<k}^i \partial^\alpha \phi_{<k}^j - \phi_{<k}^j \partial^\alpha \phi_{<k}^i)$$

as the coefficients of the first order term, which is crucial later on.

Passing from the original equation to its paradifferential form turns out to be quite nontrivial. In particular, it cannot be done using bilinear estimates; as it turns out, in addition one needs a trilinear estimate, which is central in Tao's work. To explain it let us begin with a consequence of the dyadic estimates (10), (11), namely

$$P_k(X_{k_1} \partial^\alpha X_{k_2} \partial_\alpha X_{k_3}) \rightarrow Y_k.$$

The  $2^{\varepsilon|j-k|}$  gain in (10) already shows that there is a gain above as well unless we have a high-low-low type interaction, i.e.  $k = \max\{k_1, k_2, k_3\}$ . Tao goes one step further and proves that there is an additional gain unless  $k_1 < \min\{k_2, k_3\}$ . This shows that the only nonremovable logarithmic divergence occurs when the undifferentiated term is the smallest in frequency.

The need for a trilinear estimate seems to be connected to the choice of the  $X$  and  $Y$  spaces. Hopefully, a more careful choice<sup>3</sup> could lead to better bilinear estimates which would imply the trilinear one.

**4.6. The gauge transformation.** The right hand side of (19) cannot be treated as an error term. The idea used by Tao, inspired from similar work on harmonic maps (see Hélein's book [11] and references therein), is to do a gauge transformation,

$$\phi_k \rightarrow U_{<k} \phi_k$$

for some  $U_{<k}$  in  $L^\infty \cap X$ . Ideally  $U_{<k}$  should be orthogonal, and, in order to cancel the  $A_{<k}^\alpha$ , it should satisfy

$$U_{<k} A_{<k}^\alpha + \partial^\alpha U_{<k} = 0.$$

This would imply that  $A^\alpha$ 's are antisymmetric and justifies in part the earlier choice of the paradifferential formulation of the wave maps equation.

Unfortunately, the above system can be solved exactly only if the following compatibility conditions hold:

$$\partial^\alpha A_{<k}^\beta - \partial^\beta A_{<k}^\alpha = [A_{<k}^\alpha, A_{<k}^\beta]$$

which are not true in general. However, we do have good control over the curl of  $A$ , so there is hope to find at least an approximate gauge transformation. This is obtained inductively using a paradifferential type setup,

$$U_k = -U_{<k}(\phi_{<k}^i \phi_k^j - \phi_{<k}^j u_k^i).$$

The main estimate in this context is

$$(20) \quad \square(U_{<k} \phi_k) = \text{error}.$$

This allows one to use linear estimates for  $U_{<k} \phi_k$  and closes the loop.

**4.7. Embedded manifolds and Moser type estimates.** Assume now that the target manifold  $M$  is isometrically embedded in  $\mathbb{R}^n$ , which is the setup considered in [32]. Then the wave maps equation has the form

$$(21) \quad \square \phi^i = -S_{jl}^i(\phi)(\partial^\alpha \phi^j, \partial_\alpha \phi^l).$$

The image of the second fundamental form is contained in the normal space of  $M$ ; therefore we have the compatibility condition

$$S_{jl}^i(\phi) \partial^\alpha \phi^i = 0.$$

---

<sup>3</sup>Krieger's recent work seems to go in the right direction.

Taking the two relations above into account, we arrive at a paradifferential formulation of (21) of the form

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

where  $A_{<k}^\alpha$  are the antisymmetric matrices

$$(22) \quad (A_{<k}^\alpha)_{ij} = ([S_{jl}^i(\phi)]_{<k} - [S_{il}^j(\phi)]_{<k})\partial_\alpha\phi_{<k}^l.$$

The gauge transformation is defined inductively as

$$U_k = U_{<k}([S_{jl}^i(\phi)]_{<k} - [S_{il}^j(\phi)]_{<k})\phi_k^l,$$

and the analysis is not very different from the case of the sphere. However, here are two new problems which we need to consider.

a) If the manifold is unbounded, then the restriction  $\phi \in L^\infty$  becomes unreasonable. The solution is to simply drop the boundedness assumption (and not use it in estimates).

b) We need to bound the expression  $S(\phi)$  in  $X$  in terms of  $\phi$  in  $X$ .  $X \cap L^\infty$  is an algebra, so this is not a problem if  $S$  is analytic. However, this restriction is too severe. Instead, one should seek a Moser-type estimate. The classical Moser estimates have the form

$$\|f(u)\|_{H^s} \leq c(\|u\|_{L^\infty})(1 + \|u\|_{H^s})$$

and are proved using the paradifferential calculus.

Its counterparts in our case, proved in [32], have the form

$$\begin{aligned} \|f(\phi)\|_X &\leq c\|\phi\|_X(1 + \|\phi\|_X^N), \\ \|f(\phi)\|_{X^s} &\leq c\|\phi\|_{X^s}(1 + \|\phi\|_X^N). \end{aligned}$$

Here  $f$  is a bounded function with sufficiently many bounded derivatives. In exchange, no boundedness condition is imposed on the functions in  $X$ . Also  $N$  is a large enough integer (probably 2 or 3 should do, but this is not easy to verify).

To prove these Moser-type estimates the usual paradifferential calculus is no longer good enough. Instead one needs to consider a family of multilinear paradifferential expansions for the nonlinear expression  $f(\phi)$ .

**4.8. The linearized equation.** Let  $\phi$  be a wave map with values in  $M$ . Then a solution  $\psi$  to the linearized equations around  $\phi$  takes values in the tangent bundle of  $M$ ,

$$(23) \quad \psi(x) \in T_\phi(x)M.$$

Since the wave maps equation is translation invariant, it follows that in particular the functions  $\partial_\alpha\phi$  are solutions for the linearized equations. This is exploited in the higher dimensional approach in [17], [27] and [23].

One possibility at this point, followed for instance in [17], [27], is to express the linearized equations in a local basis in  $TM$ . Instead, we simply linearize the equation (21),

$$(24) \quad \square\psi^i = -2S_{jl}^i(\phi)(\partial^\alpha\phi^j, \partial_\alpha\psi^l) - (\partial_m S)_{jl}^i(\phi)(\partial^\alpha\phi^j, \partial_\alpha\phi^l)\psi^m.$$

Next we write the appropriate paradifferential formulation. Recall that we measure  $\psi$  in a lower regularity space  $X^s$  with  $s < n/2$ . Consequently, every time a low frequency in  $\psi$  contributes to a higher frequency output in the nonlinearity, there is an additional gain compared to the case when  $s = \frac{n}{2}$ . Then in the error term we

can include the second right hand side term entirely and also the part of the first term where  $\psi$  does not have the highest frequency. Hence we arrive at

$$\square\psi_k^i = -2S_{jl}^i(\phi)_{<k}(\partial^\alpha\phi_{<k}^j, \partial_\alpha\psi_k^l) + \text{error}.$$

On the other hand, the compatibility condition (23) yields

$$S_{jl}^i(\phi)\psi^i = 0.$$

Differentiating this we conclude that the expression

$$S_{jl}^i(\phi)_{<k}\partial_\alpha\psi_k^i$$

should also be better behaved. This leads us to the paradifferential formulation of the linearized equations

$$\square\psi = -2A_{<k}^\alpha\partial_\alpha\psi + \text{error}$$

where, unsurprisingly,  $A^\alpha$  are the same as in (22). But this can be dealt with as before by making a gauge transformation.

## 5. CONCLUSIONS

Nonlinear dispersive equations in general, and nonlinear wave equations in particular, represent an important part of the current research in partial differential equations.

The initial work in this field was mostly devoted to the local study of problems with reasonably smooth initial data. Often this reduces to analyzing linear problems. The well-posedness questions become harder to answer when one considers either problems with less regular data or solutions which exist for a larger time. Then one needs to gain a better and better understanding of the nonlinear wave interactions. The work on the wave maps that we described in these notes is representative of what happens for other nonlinear wave equations.

Up to a certain point one can use linear tools combined with the contraction principle, although, perhaps, with greater and greater difficulty. The solutions obtained have a Lipschitz continuous dependence on the initial data. One may call this a *semilinear* behavior pattern. In the case of the wave maps equation it corresponds to  $s > \frac{n}{2}$ .

Beyond this point, as  $s$  decreases, the low frequency waves become large enough so that they significantly alter the background geometry on which the high frequency waves live. At the same time, the interaction between high frequencies remains weaker; i.e. it retains a semilinear character. Because of these factors, only a continuous dependence of the solutions on the initial data survives. We could call this a *quasilinear* behavior. For the wave maps equation it corresponds to the critical energy level  $s = \frac{n}{2}$ .

Finally, for an even smaller  $s$ , one could conceivably enter a regime where strong interactions occur between high frequency waves. Examples of this type are scarce and less understood, possibly because such interactions are more likely to lead to an immediate breakdown. This could be called a *fully nonlinear* behavior.

We hope that this brief presentation on the study of the wave maps equation may serve as a guide for other problems of a similar nature.

## REFERENCES

1. Piotr Bizoń, Tadeusz Chmaj, and Zbislav Tabor, *Formation of singularities for equivariant  $(2 + 1)$ -dimensional wave maps into the 2-sphere*, Nonlinearity **14** (2001), no. 5, 1041–1053. MR **2003b**:58043
2. H. Brezis and L. Nirenberg, *Degree theory and BMO. I. Compact manifolds without boundaries*, Selecta Math. (N.S.) **1** (1995), no. 2, 197–263. MR **96g**:58023
3. Thierry Cazenave, Jalal Shatah, and A. Shadi Tahvildar-Zadeh, *Harmonic maps of the hyperbolic space and development of singularities in wave maps and Yang-Mills fields*, Ann. Inst. H. Poincaré Phys. Théor. **68** (1998), no. 3, 315–349. MR **2000g**:58042
4. Demetrios Christodoulou and A. Shadi Tahvildar-Zadeh, *On the regularity of spherically symmetric wave maps*, Comm. Pure Appl. Math. **46** (1993), no. 7, 1041–1091. MR **94e**:58030
5. Piero D’Ancona and V. Georgiev, *Ill posedness results for the two dimensional wave maps equation*, preprint.
6. Alexandre Freire, Stefan Müller, and Michael Struwe, *Weak compactness of wave maps and harmonic maps*, Ann. Inst. H. Poincaré Anal. Non Linéaire **15** (1998), no. 6, 725–754. MR **2000a**:58045
7. J. Ginibre and G. Velo, *The Cauchy problem for the  $O(N)$ ,  $CP(N-1)$ , and  $G_C(N, p)$  models*, Ann. Physics **142** (1982), no. 2, 393–415. MR **84i**:58027
8. M. L. Gromov, *Isometric imbeddings and immersions*, Dokl. Akad. Nauk SSSR **192** (1970), 1206–1209. MR **43**:1212
9. Chao Hao Gu, *On the Cauchy problem for harmonic maps defined on two-dimensional Minkowski space*, Comm. Pure Appl. Math. **33** (1980), no. 6, 727–737. MR **82g**:58027
10. Matthias Günther, *Isometric embeddings of Riemannian manifolds*, Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990) (Tokyo), Math. Soc. Japan, 1991, pp. 1137–1143. MR **93b**:53049
11. Frédéric Hélein, *Régularité des applications faiblement harmoniques entre une surface et une sphère*, C. R. Acad. Sci. Paris Sér. I Math. **311** (1990), no. 9, 519–524. MR **92a**:58034
12. ———, *Harmonic maps, conservation laws and moving frames*, second ed., Cambridge Tracts in Mathematics, vol. 150, Cambridge University Press, Cambridge, 2002. MR **2003g**:58024
13. James Isenberg and Steven L. Liebling, *Singularity formation in  $2 + 1$  wave maps*, J. Math. Phys. **43** (2002), no. 1, 678–683. MR **2002i**:58034
14. Markus Keel and Terence Tao, *Local and global well-posedness of wave maps on  $\mathbf{R}^{1+1}$  for rough data*, Internat. Math. Res. Notices (1998), no. 21, 1117–1156. MR **99k**:58180
15. 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), Lectures in Appl. Math., vol. 23, Amer. Math. Soc., Providence, RI, 1986, pp. 293–326. MR **87h**:35217
16. S. Klainerman and M. Machedon, *Space-time estimates for null forms and the local existence theorem*, Comm. Pure Appl. Math. **46** (1993), no. 9, 1221–1268. MR **94h**:35137
17. Sergiu Klainerman and Igor Rodnianski, *On the global regularity of wave maps in the critical Sobolev norm*, Internat. Math. Res. Notices (2001), no. 13, 655–677. MR **2002h**:58051
18. Sergiu Klainerman and Sigmund Selberg, *Remark on the optimal regularity for equations of wave maps type*, Comm. Partial Differential Equations **22** (1997), no. 5-6, 901–918. MR **99c**:35163
19. Joachim Krieger, *Global regularity of wave maps from  $R^{3+1}$  to surfaces*, Comm. Math. Phys. **238** (2003), no. 1-2, 333–366.
20. O. A. Ladyzhenskaya and V. I. Shubov, *On the unique solvability of the Cauchy problem for equations of two-dimensional relativistic chiral fields with values in complete Riemannian manifolds*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **110** (1981), 81–94, 242–243, Boundary value problems of mathematical physics and related questions in the theory of functions, 13. MR **83h**:81051
21. Hans Lindblad, *A sharp counterexample to the local existence of low-regularity solutions to nonlinear wave equations*, Duke Math. J. **72** (1993), no. 2, 503–539. MR **94h**:35165
22. Stefan Müller and Michael Struwe, *Global existence of wave maps in  $1 + 2$  dimensions with finite energy data*, Topol. Methods Nonlinear Anal. **7** (1996), no. 2, 245–259. MR **99b**:58063
23. Andrea Nahmod, Atanas Stefanov, and Karen Uhlenbeck, *On the well-posedness of the wave map problem in high dimensions*, preprint.

24. John Nash, *The imbedding problem for Riemannian manifolds*, Ann. of Math. (2) **63** (1956), 20–63. MR **17f**:782b
25. Jalal Shatah, *Weak solutions and development of singularities of the SU(2)  $\sigma$ -model*, Comm. Pure Appl. Math. **41** (1988), no. 4, 459–469. MR **89f**:58044
26. Jalal Shatah and Michael Struwe, *Geometric wave equations*, Courant Lecture Notes in Mathematics, vol. 2, New York University Courant Institute of Mathematical Sciences, New York, 1998. MR **2000i**:35135
27. ———, *The Cauchy problem for wave maps*, Int. Math. Res. Not. (2002), no. 11, 555–571. MR **2003m**:58042
28. Jalal Shatah and A. Shadi Tahvildar-Zadeh, *On the Cauchy problem for equivariant wave maps*, Comm. Pure Appl. Math. **47** (1994), no. 5, 719–754. MR **96c**:58049
29. Thomas C. Sideris, *Global existence of harmonic maps in Minkowski space*, Comm. Pure Appl. Math. **42** (1989), no. 1, 1–13. MR **89k**:58069
30. Terence Tao, *Global regularity of wave maps. I. Small critical Sobolev norm in high dimension*, Internat. Math. Res. Notices (2001), no. 6, 299–328. MR **2001m**:35200
31. ———, *Global regularity of wave maps. II. Small energy in two dimensions*, Comm. Math. Phys. **224** (2001), no. 2, 443–544. MR **2002h**:58052
32. Daniel Tataru, *Small solutions for the wave maps equation in critical Sobolev spaces*, preprint.
33. ———, *Local and global results for wave maps. I*, Comm. Partial Differential Equations **23** (1998), no. 9-10, 1781–1793. MR **99j**:58209
34. ———, *On global existence and scattering for the wave maps equation*, Amer. J. Math. **123** (2001), no. 1, 37–77. MR **2002c**:58045

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720  
E-mail address: [tataru@math.berkeley.edu](mailto:tataru@math.berkeley.edu)