

*Journées*

# **ÉQUATIONS AUX DÉRIVÉES PARTIELLES**

Forges-les-Eaux, 7 juin–11 juin 2004

Terence Tao

**Geometric renormalization of large energy wave maps**

*J. É. D. P.* (2004), Exposé n° XI, 32 p.

<[http://jedp.cedram.org/item?id=JEDP\\_2004\\_\\_\\_\\_A11\\_0](http://jedp.cedram.org/item?id=JEDP_2004____A11_0)>

**cedram**

*Article mis en ligne dans le cadre du*

*Centre de diffusion des revues académiques de mathématiques*

<http://www.cedram.org/>

# Geometric renormalization of large energy wave maps

Terence Tao

## Abstract

There has been much progress in recent years in understanding the existence problem for wave maps with small critical Sobolev norm (in particular for two-dimensional wave maps with small energy); a key aspect in that theory has been a renormalization procedure (either a geometric Coulomb gauge, or a microlocal gauge) which converts the nonlinear term into one closer to that of a semilinear wave equation. However, both of these renormalization procedures encounter difficulty if the energy of the solution is large. In this report we present a different renormalization, based on the harmonic map heat flow, which works for large energy wave maps from two dimensions to hyperbolic spaces. We also observe an intriguing estimate of “non-concentration” type, which asserts roughly speaking that if the energy of a wave map concentrates at a point, then it becomes asymptotically self-similar.

## 1. Introduction

Let  $n, m \geq 1$  be integers, and let  $\mathbf{R}^{1+n}$  be Minkowski space  $\{(t, x) : t \in \mathbf{R}, x \in \mathbf{R}^n\}$  with the usual metric  $g_{\alpha\beta}x^\alpha x^\beta = -dt^2 + dx^2$ . We parameterize this space by Greek indices  $\alpha, \beta = 0, 1, \dots, n$ , raised and lowered in the usual manner.

Let  $(N, h)$  be a smooth connected complete  $m$ -dimensional Riemannian manifold without boundary, and let  $I \subseteq \mathbf{R}$  be an interval. A *Schwartz wave map* from a slab  $I \times \mathbf{R}^n$  in Minkowski space to  $N$  is defined to be any smooth map  $\phi : I \times \mathbf{R}^n \rightarrow N$  which decays rapidly to some constant  $\phi(\infty) \in N$  at spatial infinity, and whose

derivatives all decay rapidly to 0.<sup>1</sup>, and is a (formal) critical point of the Lagrangian

$$(1) \quad \int_{\mathbf{R}^{1+n}} \langle \partial^\alpha \phi(t, x), \partial_\alpha \phi(t, x) \rangle_{h(\phi(t, x))} dt dx.$$

The Euler-Lagrange equation for Schwartz wave maps can then be written as

$$(2) \quad (\phi^* \nabla)^\alpha \partial_\alpha \phi = 0,$$

where  $(\phi^* \nabla)^\alpha$  is covariant differentiation on the vector bundle  $\phi^*(TN)$  with respect to the pull-back  $\phi^* \nabla$  via  $\phi$  of the Levi-Civita connection  $\nabla$  on  $N$ . Wave maps are of course the Minkowski space analogue of harmonic maps.

In this paper we shall restrict primarily to the case when  $n = 2$ ,  $m \geq 2$  (the  $m = 1$  case being reducible to the free (linear) wave equation) and when  $N$  has constant negative curvature  $\kappa < 0$ , thus

$$(3) \quad [\nabla_X, \nabla_Y]Z - \nabla_{[X, Y]}Z =: R(X, Y)Z = \kappa(\langle Y, Z \rangle_h X - \langle X, Z \rangle_h Y)$$

for all smooth vector fields  $X, Y, Z$  on  $N$ . Note that  $\kappa$  can be normalized to  $-1$  by scaling the metric  $h$ . A typical example of such a manifold  $N$  (with  $\kappa = -1$ ) is the hyperbolic space  $\mathbf{H}^m := O(m, 1)/O(m) = SO(m, 1)/SO(m)$ , where  $O(m)$  is the orthogonal group on  $\mathbf{R}^m$  and  $O(m, 1)$  is the Lorentz group on  $\mathbf{R}^{1+m}$ , with  $O(m, 1)$  and  $SO(m, 1)$  being the respective orientation-preserving components of these groups. More generally, one can quotient  $\mathbf{H}^m$  by a discrete subgroup of  $O(m, 1)$  (acting on the left in the usual manner), and as is well known one can recover all constant negative curvature  $m$ -dimensional manifolds  $N$  this way, up to isometry and normalization of  $\kappa$ . The case  $N = \mathbf{H}^2$  is of particular interest as it arises as a model problem for the Einstein equations under some additional symmetry assumptions (see e.g. [1] for a discussion).

*Remark 1.1.* With the constant negative curvature assumption, Schwartz maps  $\phi$  become topologically trivial (contractable to a constant map); by compactifying infinity, this remark is equivalent to the assertion that any map from a sphere  $S^n$  to  $N$  is topologically trivial. When  $N$  is hyperbolic space  $\mathbf{H}^m$  this claim is clear since  $N$  is contractible; more generally, if  $N$  is  $\mathbf{H}^m$  quotiented out by a discrete group, then the claim follows by lifting the map up to hyperbolic space (the discreteness of the fibers keeps the monodromy trivial). Later on we will see that these maps can in fact be explicitly deformed into a constant map by means of the harmonic map heat flow. Note that while  $N$  itself need not be parallelizable or even orientable, this will not particularly concern us since maps into  $N$  can be lifted up into  $\mathbf{H}^m$ , which is both parallelizable and orientable. Indeed one could reduce to the case  $N = \mathbf{H}^m$  by lifting the wave map up to hyperbolic space, and observing that the wave map equation (2) is preserved by this lift.

---

<sup>1</sup>In other words, for each fixed  $t$ ,  $\phi(t, x)$  converges to  $\phi(\infty)$  faster than any negative power of  $|x|$ , and all derivatives of  $\phi$  converge to zero faster than any negative power of  $|x|$ . We need to consider Schwartz wave maps rather than classical wave maps (which are constant outside of in a finite union of light cones) because we will need to take (nonlinear) Littlewood-Paley projections of these wave maps, obtained via the heat flow for harmonic maps, in our argument.

In local co-ordinates  $\phi = \phi^a$  on  $N$ , the wave map equation (2) can be written as

$$(4) \quad \partial^\alpha \partial_\alpha \phi^a = -\Gamma_{bc}^a(\phi) \partial_\alpha \phi^b \partial^\alpha \phi^c$$

where the  $\Gamma_{bc}^a$  are Christoffel symbols; in particular we see that  $\phi$  evolves by a nonlinear wave equation. It is then natural to consider the Cauchy problem, specifying initial data  $\phi[t_0] := (\phi(t_0), \partial_t \phi(t_0))$  at some initial time  $t_0 \in \mathbf{R}$ . Of course, this initial data  $\phi[t_0]$  must take values in the tangent bundle  $TN$  of  $N$ , and decay rapidly to  $(\phi(\infty), 0)$  as  $x \rightarrow \infty$ ; we refer to such data as *Schwartz initial data*.

The Cauchy problem for wave maps has been extensively studied (see references); we refer the interested reader to the surveys in [23], [30], [50], [59], [69]. For arbitrary dimension  $n$  and arbitrary targets  $N$ , one can easily show by general energy methods (which indeed apply to any nonlinear wave equation, see e.g. [56]) that for any Schwartz<sup>2</sup> initial data  $\phi[t_0]$ , one has a unique *local-in-time* Schwartz solution to (2) in some spacetime slab  $I \times \mathbf{R}^n$ , where  $I$  is a neighborhood of  $t_0$ , attaining the value of  $\phi(\infty)$  at infinity (thus the value of  $\phi$  at spatial infinity is fixed by the wave map evolution). Indeed, energy methods can be used to show such smooth solutions can be continued whenever the norm  $\|\nabla_{t,x} \phi(t)\|_{H^{n/2+\varepsilon}(\mathbf{R}^n)}$  remains bounded for some  $\varepsilon > 0$ . In particular, by use of the conserved energy

$$(5) \quad E(\phi) = E(\phi[t]) := \int_{\mathbf{R}^n} \frac{1}{2} |\nabla_{t,x} \phi(t, x)|_h^2 dx$$

we see that one has global existence of Schwartz solutions in one dimension; see [16], [33], [12], [46], [21]. The one dimensional wave map equation is actually rather special, being completely integrable and thus supporting special solutions such as homoclinic breather solutions; see [48].

Now we turn to higher dimensions  $n \geq 2$ . Since the wave map equation (2) is invariant under the scaling

$$(6) \quad \phi(t, x) \leftarrow \phi\left(\frac{t}{\lambda}, \frac{x}{\lambda}\right)$$

for any  $\lambda > 0$ , we see that the two-dimensional case  $n = 2$  is critical with respect to the energy (5), as this energy is now scale-invariant (dimensionless) in this case. For higher dimensions  $n > 2$  it is known that singularities can form<sup>3</sup> in finite time when the manifold  $N$  is positively curved, [46], [53], [5], [50] (and there is strong numerical evidence to suggest that there is also blowup when  $n = 2$  [18], [2] in this case); when  $n \geq 7$  one can also obtain singularities for negatively curved manifolds [5], despite such manifolds being well-behaved for other equations (such as the harmonic map

---

<sup>2</sup>The smoothness condition has been relaxed substantially, see for instance [26], [29], [21], [66], [67], [68], [69] and the discussion below; however we shall restrict our attention here to the Schwartz category. The rapid decay assumptions can also be removed by finite speed of propagation, though possibly at the cost of creating a domain of existence which is not a spacetime slab  $I \times \mathbf{R}^n$ .

<sup>3</sup>Despite the presence of these singularities, it is known that global *weak* solutions exist globally in time, see [36], [11], although it is quite possible that uniqueness and energy conservation will fail for these solutions.

heat flow). However, one still has global solutions for small data. This was first observed for initial data sufficiently close to a point [6] or a geodesic [55], and for data which is small in the critical Besov norm  $\|\nabla_{t,x}\phi\|_{\dot{B}_1^{n/2-1}(\mathbf{R}^n)}$  [66], [67], assuming that the manifold  $N$  is uniformly smooth. This was achieved by viewing the wave map equation as a nonlinear wave equation in  $\phi$  (as in (4)) and iterating in a certain sophisticated function space. These Besov space results appear to be essentially the limit of the method; for instance, it is known that such iteration methods cannot work if the data is merely assumed to be small in  $\|\nabla_{t,x}\phi\|_{\dot{H}^{n/2-1}(\mathbf{R}^n)}$  (see [21], [39], [9]). However, this small Sobolev norm global regularity result turned out to be achievable upon either a microlocal renormalization of the wave map equation, or by passing to a derivative formulation and selecting a special gauge (such as the Coulomb gauge). In the case of the sphere  $N = S^{m-1}$ , this was achieved for  $n \geq 2$  in [64], [65] via the microlocal renormalization method; for the more general class of boundedly parallelizable manifolds (which includes the hyperbolic spaces  $\mathbf{H}^m$ ), this was achieved for  $n \geq 5$  in [27] (by a hybrid of the microlocal renormalization and Coulomb gauge methods) and then for  $n \geq 4$  in [51], [38] and  $n \geq 3$  in [31] (by use of the Coulomb gauge). For the hyperbolic plane  $N = \mathbf{H}^2$  the  $n = 2$  case was recently treated in [32] (again using the Coulomb gauge), while for manifolds which are uniformly isometrically embeddable in Euclidean space the result was obtained for all  $n \geq 2$  in [68] (together with more precise well-posedness results). While the result in [68] does not directly cover the hyperbolic spaces  $\mathbf{H}^m$ , which cannot be uniformly isometrically embedded, it may be possible that the argument can be modified to treat this case by first quotienting  $\mathbf{H}^m$  by a discrete group in order to compactify the target, and lifting back to  $\mathbf{H}^m$  at the end of the argument.

Thus we now have a fairly satisfactory regularity theory when the critical Sobolev norm is small; in particular in two dimensions we have a small energy regularity theory for a reasonably large class of target manifolds. Now we turn to the question of large energy data at  $n = 2$ . In light of the numerical work in [18], [2] we do not expect global regularity in the positive curvature case, but it was still conjectured (see e.g. [24]) that one has global regularity for large energy data in negative curvature manifolds such as  $\mathbf{H}^m$ ; note that this would in fact imply an unconditional global regularity result for smooth data thanks to finite speed of propagation. In the case where the wave map has some  $U(1)$  symmetry (either rotation-invariant, or rotation equivariant), such results have already been obtained in [7] (for spherically symmetric maps) and [52], [49], [50] (for equivariant maps). The spherically symmetric regularity result in fact extends to arbitrary targets [60], [61] but the equivariant one seems restricted to sufficiently “negatively curved” targets; see [62] (and the numerics in [2], [18]) for further discussion of this issue, which is in particular related to the existence of non-constant *harmonic* maps from  $S^2$  to  $N$ , which among other things can be used to generate non-trivial finite-energy stationary solutions to (2)). numerics that strongly suggest blowup occurs for instance when the target is a sphere, even for equivariant data. In [62] it is shown that for equivariant data, blowup can only occur if the wave map relaxes to a (rescaled) harmonic map from  $\mathbf{R}^2$  to  $S^2$  for a sequence of times converging to the blowup time.

The conjecture of large energy global regularity for wave maps into  $\mathbf{H}^m$ , if true, would be analogous to other critical large energy global regularity results in the literature, notably that of the energy-critical nonlinear wave equation  $\square u = u^5$  in three dimensions (see e.g. [50]), the critical Yang-Mills-Higgs equation in three dimensions [19], and the energy-critical nonlinear Schrödinger equation  $iu_t + \Delta u = |u|^4 u$  in three dimensions (see [14], [4] for the radial case, and [8] for the general case). These large energy results are substantially more difficult than the small energy theory, and require at least two new ingredients:

- A perturbation, local existence, and theory suitable for large energy solutions, assuming that a suitable spacetime norm (e.g. the  $L_{t,x}^{10}$  norm for the energy-critical three-dimensional nonlinear Schrödinger equation, or  $L_t^4 L_x^{12}$  norm for the energy-critical three-dimensional nonlinear wave equation or Yang-Mills-Higgs equation) is bounded;
- Some sort of energy non-concentration argument which establishes that such a spacetime norm is bounded.

The energy non-concentration argument generally relies on some sort of monotonicity formula, often arising from a Morawetz inequality, although in some cases (notably that of the energy-critical Schrödinger equation) such an inequality is not sufficient by itself to control the desired spacetime quantity; one must supplement that inequality with additional arguments, such as the induction-on-energy strategy pioneered by Bourgain [4] and which also plays a fundamental role in [8]. On the other hand, in these types of arguments, the large energy perturbation analysis does not require such a Morawetz inequality, and instead proceeds by a variant of the standard existence theory, using Strichartz estimates as the primary tool. In order to do this, one relies crucially on the *fungibility in time* of spacetime norms such as the  $L_{t,x}^{10}$ ; what we mean by this is that if the  $L_{t,x}^{10}$  norm is large but bounded, then one can decompose the time interval into a bounded number of sub-intervals, such that the  $L_{t,x}^{10}$  norm is small on each interval. This allows one (at least in principle) to obtain a perturbation theory for large  $L_{t,x}^{10}$  norm solutions from that for small  $L_{t,x}^{10}$  norm solutions. This is in contrast with norms such as  $L_t^\infty L_x^6$  which are clearly bounded by the energy for (say) the energy-critical NLS, but are not fungible in time and so cannot be used directly to establish a global regularity result for this equation.

We do not fully resolve either of these two issues here, however we can announce some progress on both. More precisely, we are able to establish a large energy perturbation theory for wave maps, in the case when the target is a space of constant negative curvature, although the actual details of this theory are rather technical and will be only partially sketched here. Secondly, we have a simple monotonicity formula argument which shows that if the energy of a wave map (with arbitrary target) concentrates at a point, then it must be asymptotically self-similar in an averaged sense (see Proposition 3.1).

We now discuss the gauge renormalization (leaving the discussion of the non-concentration result to Section 3). For the model problem of constant negative curvature, a direct attempt to mimic either the microlocal renormalization procedure in [64], [65], [27], [68] or the Coulomb gauge renormalization procedure in [51], [38], [31], [32] runs into difficulty (although the Coulomb gauge renormalization will work in the abelian case, in which the target is a two-dimensional space such as  $\mathbf{H}^2$ , so that the gauge group  $SO(2)$  is abelian). More specifically, one begins to encounter difficulty in large energy in keeping the microlocal gauge change approximately unitary, and in the Coulomb gauge one has problems establishing uniqueness, regularity, and ellipticity of the gauge in the non-abelian case where the target has dimension  $m > 2$  and so the gauge group  $SO(m)$  is non-abelian. Also, the microlocal gauge has not yet been successfully applied to manifolds such as hyperbolic space, which are not easy to embed inside Euclidean spaces; meanwhile, the Coulomb gauge introduces some additional (and somewhat artificial) singularities at the spatial frequency origin  $\xi = 0$  (arising from the Hodge decomposition) which cause additional technical complications, especially in two dimensions (see [31], [32] for a demonstration of this phenomenon). Finally, while both these gauges are well suited for establishing regularity of a *single* solution, the task of establishing a perturbation theory - i.e. comparing two nearby solutions - is also not trivial, as one must then establish that the gauge transforms themselves have suitable continuity properties, which is an extremely delicate issue in the critical setting.

In this report we give a new renormalization procedure, which is based on the harmonic map heat flow, which is intrinsic to the geometry and seems a extremely natural choice of gauge; we shall dub it the “canonical heat-temporal gauge” or “caloric gauge”. This caloric gauge has essentially the same effect as the microlocal gauge change (which can be viewed as a discretized, partially linearized, version of the harmonic map heat flow renormalization), in that it replaces the nonlinearity by a “paraproduct” variant in which at least one derivative falls on a low frequency term. However, unlike the microlocal gauge change, the caloric gauge works well in large energies and negative curvature targets, essentially thanks to the famous result of Eells and Sampson [10] which shows that the harmonic map heat flow for those targets is globally smooth and converges to a constant. The caloric gauge also is well adapted to perturbation theory, since the harmonic map heat flow also enjoys a good perturbation theory. When compared against the Coulomb gauge, the caloric gauge has the advantage of renormalizing the nonlinearity into a slightly better form (with no singularity near the spatial frequency origin); we shall give a heuristic comparison of the two renormalized nonlinearities using Littlewood-Paley analysis later on in this report. It is also insensitive to non-abelian behaviour in the gauge group, and does not develop difficulties with uniqueness or regularity, as long as the target is negatively curved.

The discussion here will consist of rigorous geometric computations, and non-rigorous analytical heuristics; a rigorous treatment of the perturbation theory which is informally discussed here will appear elsewhere (and in a much lengthier paper).

*Acknowledgements.* The author is indebted to Ben Andrews and Andrew Hassell for a crash course in Riemannian geometry and manifold embedding, and in particular to Ben Andrews for explaining the harmonic map heat flow. The author also thanks Mark Keel for background material on wave maps, and Daniel Tataru for sharing some valuable insights on multilinear estimates and function spaces. The author thanks Andrew Hassell and the Australian National University for their hospitality when much of this work was conducted. The author is a Clay Prize Fellow and is supported by a grant from the Packard Foundation.

## 2. The derivative formulation

Here and in the rest of the paper we fix  $N$  to be a smooth complete Riemannian manifold of constant negative curvature  $\kappa < 0$ . We now describe the (standard) derivative formulation of the wave map equation (2), writing  $\nabla_{t,x}\phi$  in terms of an orthonormal frame  $e(t, x)$ ; this formulation was used for instance in [27], [51], [38] (or in [17] for harmonic maps). Our discussions here shall be primarily geometric rather than analytic. As we shall need this derivative formulation not just for the wave map equation, but also for the harmonic map heat flow (and in fact for a concatenation of the two equations), we shall first describe the derivative formulation of a general smooth map  $\phi : \Omega \rightarrow N$ , where  $\Omega$  is an open contractible subset of either a Minkowski space  $\mathbf{R}^{1+n}$  or a Euclidean space  $\mathbf{R}^n$ . In either case we will use the co-ordinates  $x^\alpha$  to parameterize the domain  $\Omega$ , with the associated partial derivative operators  $\partial_\alpha$ . For this general discussion we will not need the metric structure of  $\Omega$  (i.e. we will not raise and lower indices), although we will use the fact that  $\Omega$  is flat in the sense that the partial derivatives  $\partial_\alpha$  commute with each other, thus  $[\partial_\alpha, \partial_\beta] = 0$ .

We recall the *orthonormal frame bundle*  $\text{Frame}(N)$  of  $N$ , defined as the collection of all pairs  $(\phi, e)$  where  $\phi \in N$  and  $e : \mathbf{R}^m \rightarrow T_\phi N$  is an orthogonal transformation (or equivalently, an orthonormal basis of  $T_\phi N$ ). It is geometrically obvious that,  $\text{Frame}(N)$  is a smooth principal  $O(m)$ -bundle of  $N$ , where  $O(m)$  is the orthogonal group on  $\mathbf{R}^m$  (not necessarily orientation preserving). Since  $\Omega$  is contractible, we can lift any map  $\phi : \Omega \rightarrow N$  to a map  $(\phi, e) : \Omega \rightarrow \text{Frame}(N)$  into the orthonormal frame bundle. In fact there are multiple such lifts; given any smooth function  $U : \Omega \rightarrow O(m)$ , we have the *gauge transform*  $(\phi, e) \mapsto (\phi, eU^{-1})$  which transforms any lift of  $\phi$  to any other lift of  $\phi$ ; indeed, since  $\text{Frame}(N)$  is a principal  $O(m)$ -bundle, all such lifts can be obtained from a single specified lift by a (unique) gauge transform.

Let  $(\phi, e) : \Omega \rightarrow \text{Frame}(N)$  be a smooth map. The bundle  $\phi^*(TN)$  (the pullback of the tangent bundle of  $N$  under  $\phi$ ) is an  $m$ -dimensional vector bundle on  $\Omega$ , and the inverse  $e^{-1}$  of the orthonormal frame  $e$  can be used to identify this with the trivial bundle  $\Omega \times \mathbf{R}^m$ . In particular, since the derivatives  $\partial_\alpha \phi$  are sections of  $\phi^*(TN)$ , we



can define the  $\mathbf{R}^m$ -valued one-form  $\psi_\alpha : U \rightarrow \mathbf{R}^m$  by this procedure, thus

$$\psi_\alpha := e^{-1} \partial_\alpha \phi$$

or equivalently

$$(7) \quad \partial_\alpha \phi = e \psi_\alpha$$

The Levi-Civita connection  $\nabla$  on  $TN$  induces a pullback connection  $\phi^*(\nabla)$  on  $\phi^*(TN)$ , which in turn induces a connection on the bundle  $\Omega \times \mathbf{R}^m$  by the above identification. This connection is compatible with the gauge group  $O(m)$  (since the original Levi-Civita connection is also), and is given by the covariant derivatives

$$(8) \quad \mathbf{D}_\alpha := \partial_\alpha + A_\alpha.$$

where the  $A_\alpha : \Omega \rightarrow o(m)$  take values in  $o(m)$ , the skew-symmetric linear transformations of  $\mathbf{R}^m$  (and the Lie algebra of  $O(m)$ ) and are defined by

$$(9) \quad (\phi^* \nabla)_\alpha e = e A_\alpha.$$

From the Leibnitz rule and (9) we have

$$(10) \quad (\phi^* \nabla)_\alpha (ef) = e(\mathbf{D}_\alpha f)$$

for any smooth scalar function  $f : \Omega \rightarrow \mathbf{R}^m$ . Since the Levi-Civita connection  $\nabla$  is torsion free, we have

$$\nabla_{\partial_\alpha \phi} \partial_\beta \phi = \nabla_{\partial_\beta \phi} \partial_\alpha \phi$$

which upon pulling back by  $\phi$  becomes

$$(\phi^* \nabla)_\alpha \partial_\beta \phi = (\phi^* \nabla)_\beta \partial_\alpha \phi.$$

From (7), (10) we thus have the fundamental *zero torsion identity*

$$(11) \quad \mathbf{D}_\alpha \psi_\beta - \mathbf{D}_\beta \psi_\alpha = 0.$$

This identity allows us (at least in principle) to write any of the fields  $\psi_\alpha$  in terms of a single field  $\psi_\beta$ , thus converting the derivative wave map system into something resembling a scalar equation.

*Remark 2.1.* The tangent bundle  $T\text{Frame}(N)$  of  $\text{Frame}(N)$  can be canonically identified with  $\text{Frame}(N) \times \mathbf{R}^m \times o(m)$ , since for any  $(\phi, e) \in \text{Frame}(N)$ , the orthonormal basis  $e$  can be used to parameterize the horizontal space  $T_\phi N$  (lifted up to  $T_{(\phi, e)} \text{Frame}(N)$  via the Levi-Civita connection) as  $\mathbf{R}^m$ , while the vertical space of the bundle  $\text{Frame}(N)$  at  $(\phi, e)$  can be parameterized as  $o(m)$ . The fields  $(\psi_\alpha, A_\alpha)$  can thus be thought of as the derivatives  $\partial_\alpha(\phi, e)$  of the original fields  $(\phi, e)$  with respect to this identification. Furthermore, one can show that the Lie algebra structure of this  $\mathbf{R}^m \times o(m)$ -bundle induces an action of the special Lorentz group  $O(m, 1)$ , which then acts transitively on  $\text{Frame}(N)$ . This in turn can be used to verify the assertion made earlier that  $N$  is the quotient of  $\mathbf{H}^m$  by a discrete group. We were however unable to exploit this group structure to any particular advantage, and shall mostly treat the fields  $\phi$ ,  $e$  and  $\psi_\alpha, A_\alpha$  separately.

The connection  $A_\alpha$  has a *curvature tensor*  $F_{\alpha\beta} : \Omega \rightarrow o(m)$  defined by

$$(12) \quad F_{\alpha\beta} := [\mathbf{D}_\alpha, \mathbf{D}_\beta] = \partial_\alpha A_\beta - \partial_\beta A_\alpha + [A_\alpha, A_\beta].$$

We can use the constant negative curvature hypothesis (3) to compute  $F_{\alpha\beta}$ . Indeed, from (3) we have

$$[(\phi^* \nabla)_\alpha, (\phi^* \nabla)_\beta]Z = \kappa(\langle \partial_\beta \phi, Z \rangle_h \partial_\alpha \phi - \langle \partial_\alpha \phi, Z \rangle_h \partial_\beta \phi)$$

for all sections  $Z$  of  $\phi^*(TN)$ ; applying (7), (10), (12) we obtain the *curvature identity*

$$(13) \quad F_{\alpha\beta} = \kappa \psi_\alpha \wedge \psi_\beta$$

where we use  $u \wedge v \in o(m)$  to denote the skew-symmetric linear operator on  $\mathbf{R}^m$  defined by<sup>4</sup>

$$(u \wedge v)w := u\langle v, w \rangle - v\langle u, w \rangle.$$

This identity will allow us (given a suitable fixing of gauge, see below) to recover  $A$  as some sort of quadratic integral of  $\psi$ .

*Remark 2.2.* In the Riemann surface case  $m = 2$ , the gauge group  $O(m)$  is abelian and so the commutator term  $[A_\alpha, A_\beta]$  vanishes in (12). This is in particular useful for placing  $A$  in the Coulomb gauge. However in the canonical heat-temporal gauge, we will not be unduly bothered by this commutator term, indeed in that gauge this commutator will have a number of vanishing components and will thus essentially be ignored.

We refer to  $(\phi, e)$  as the *undifferentiated fields* and  $(\psi_\alpha, A_\alpha)$  as the *differentiated fields*. The gauge freedom  $e \mapsto eU^{-1}$  mentioned earlier affects these fields by the formulae

$$(14) \quad (\phi, e) \mapsto (\phi, eU^{-1}); \quad (\psi_\alpha, A_\alpha) \mapsto (U\psi_\alpha, UA_\alpha U^{-1} + (\partial_\alpha U)U^{-1}); \quad F_{\alpha\beta} \mapsto UF_{\alpha\beta}U^{-1}.$$

In particular we observe that the magnitudes  $|\psi_\alpha|$  and  $|F_{\alpha\beta}|$  are independent of the choice of gauge. Here we use  $|F_{\alpha\beta}|$  to denote the Hilbert-Schmidt norm, thus  $|F| := \text{tr}(FF^*)^{1/2}$ .

*Remark 2.3.* Clearly from (7) and (9) we can obtain the differentiated fields from the undifferentiated fields. Now we consider the reverse question, whether one can obtain the undifferentiated fields from the differentiated fields. To fix boundary conditions, let us now assume that  $\Omega$  is convex, unbounded, and connected at infinity, and that  $(\phi, e)$  converges rapidly (in the Schwartz sense, i.e. all derivatives  $\nabla_{x,t}^N(\phi, e)$  decay faster than any polynomial) to some point  $(\phi(\infty), e(\infty)) \in \text{Frame}(N)$  at infinity<sup>5</sup> This then implies that  $(\psi_\alpha, A_\alpha)$  are Schwartz. To reconstruct  $(\phi, e)$  from

<sup>4</sup>The cancellation inherent in this wedge product operation will be largely unexploited, in contrast to [32], [31], where this type of “ $Q_{ij}$  null structure” is needed to compensate for the singularity of  $\nabla^{-1}$  in the Coulomb gauge. We will, however, rely crucially on the negativity of  $\kappa$  to ensure that the heat flow converges properly.

<sup>5</sup>We remark that if  $\phi$  decays rapidly to  $\phi(\infty)$  at infinity and  $e(\infty)$  is any orthonormal frame of  $T_{\phi(\infty)}N$ , then one can easily find a lift  $(\phi, e) : \Omega \rightarrow \text{Frame}(N)$  of  $\phi : \Omega \rightarrow N$  which converges to  $(\phi(\infty), e(\infty))$  at infinity. If  $N$  is hyperbolic space  $\mathbf{H}^m$  this is clear since  $\mathbf{H}^m$  is parallelizable and one can simply pull back an orthonormal frame on  $\mathbf{H}^m$  which equals  $e(\infty)$  at  $\phi(\infty)$ ; more generally one can lift the map  $\phi : \Omega \rightarrow N$  to a map  $\tilde{\phi} : \Omega \rightarrow \mathbf{H}^m$  (retaining the Schwartz property at infinity, obtain a frame  $(\tilde{\phi}, \tilde{e})$  for the lift, and then push the frame back down to  $N$ ).

$(\psi_\alpha, A_\alpha)$  one simply solves the ODE (7), (9) from the given data at infinity. A simple calculation using (11), (13), and Frobenius's theorem<sup>6</sup> shows that this system is indeed solvable; we omit the details.

To summarize, we now have two equivalent ways to view a map  $(\phi, e) : \Omega \rightarrow \text{Frame}(N)$ ; either in an undifferentiated formulation, in which  $(\phi, e)$  can move freely (but do not take values in a linear space), or in a differentiated formulation, in which the fields  $(\psi_\alpha, A_\alpha)$  now live in linear spaces but are constrained to obey the compatibility conditions (11), (13). In the Schwartz category these two viewpoints are equivalent as long as we fix  $(\phi(\infty), e(\infty))$  in advance. As we shall be needing both the differentiated and undifferentiated fields at various stages our analysis, we shall however not take advantage of this redundancy, and instead consider all the fields simultaneously.

## 2.4. The derivative formulation of the wave map equation

We now specialize to Schwartz wave maps  $\phi : I \times \mathbf{R}^n \rightarrow N$  for some compact interval  $I$ , which equal some fixed point  $\phi(\infty) \in N$  at infinity. Given any fixed frame  $e(\infty)$  in  $T_{\phi(\infty)}N$ , we can lift  $\phi$  to a map  $(\phi, e) : I \times \mathbf{R}^n \rightarrow \text{Frame}(N)$  which converge rapidly to  $(\phi(\infty), e(\infty))$  at infinity.

The wave map equation (2) can be rewritten using (7), (10) as

$$\mathbf{D}^\alpha \psi_\alpha = 0$$

where we of course now use the Minkowski metric to raise and lower indices. Combining this equation with the compatibility conditions (11), (13) and the definitions (7), (9) we can thus write the wave map equation as the system (cf. [51], [38], [31], [32])

$$(15) \quad \begin{aligned} \partial_\alpha \phi &= e \psi_\alpha \\ (\phi^* \nabla)_\alpha e &= e A_\alpha \\ \mathbf{D}^\alpha \psi_\alpha &= 0 \\ \mathbf{D}_\alpha \psi_\beta - \mathbf{D}_\beta \psi_\alpha &= 0 \\ F_{\alpha\beta} &= \kappa \psi_\alpha \wedge \psi_\beta. \end{aligned}$$

Note that these equations are both over-determined (the compatibility conditions (11), (13) enforcing a number of constraints between the fields) and under-determined (since we still have the gauge freedom (14)). In light of Remark 2.3 it is possible to omit the first two equations, and just consider the last three equations as a system for the differentiated fields  $\psi_\alpha, A_\alpha$ ; since can then recover the undifferentiated fields

---

<sup>6</sup>Alternatively, if  $0 \in \Omega$ , one can use a radial gauge  $x^\alpha (\phi^* \nabla)_\alpha e = 0$  and propagate (7), (9) along radial directions; one can then use (11), (13) and a Gronwall inequality argument to show that  $(\psi_\alpha, A_\alpha)$  are indeed the derivative fields of  $(\phi, e)$ . Again, we omit the details.

$\phi$ ,  $e$  using Frobenius's theorem (and the boundary data  $\phi(\infty), e(\infty)$ ) as discussed earlier; note that we have relied crucially on the constant curvature in order to forget about  $\phi$  and  $e$ . However, we will retain both the differentiated and undifferentiated fields in our system, as the presence of the undifferentiated fields makes the compatibility conditions automatic.

We refer to the quantity  $u := \mathbf{D}^\alpha \psi_\alpha$  as the *wave-tension field* (as opposed to the *heat-tension field*  $\psi_s = \mathbf{D}_k \psi_k$  which will appear later); thus wave maps are precisely those maps whose wave-tension field vanishes.

The system (15) is not yet well-posed in time, even locally, because of the gauge freedom (14). There are a number of options available to fix the gauge freedom and recover well-posedness; examples of such gauges include the *temporal gauge*  $A_0 = 0$ , the *Lorenz gauge*  $\partial_\alpha A^\alpha = 0$ , and the *Coulomb gauge*  $\partial_j A_j = 0$ , where Latin indices such as  $j$  will run over the spatial indices  $1, \dots, n$  with the usual summation conventions. The latter gauge has been exploited for instance in [38], [51], [31], [32]; one advantage of this gauge is that it makes  $A$  a quadratic function of  $\psi$ . Indeed, from Hodge theory one can easily verify in the Coulomb gauge that

$$\begin{aligned}
(16) \quad A_\alpha &= \Delta^{-1}(\partial_j \partial_j A_\alpha) \\
&= \Delta^{-1}(\partial_j (F_{j\alpha} + [A_\alpha, A_j]) + \partial_\alpha \partial_j A_j) \\
&= \Delta^{-1} \partial_j (\kappa \psi_j \wedge \psi_\alpha + [A_\alpha, A_j]),
\end{aligned}$$

where  $\Delta$  is the spatial Laplacian. Schematically, we thus have  $A = \nabla_x^{-1}(\psi\psi + AA)$ . This identity allows the wave map equation to be quite tractable in high dimensions (see e.g. [38], [51] for the  $n \geq 4$  cases), however the inverse derivative  $\nabla_x^{-1}$  causes some difficulties in low dimensions, although they can eventually be surmounted (see [31], [32]). Also in the large energy setting the  $\nabla^{-1}(AA)$  term also becomes problematic, indeed it can cause breakdown of uniqueness in the Coulomb gauge (though this issue does not arise in the Riemann surface case  $m = 2$ ). Because of these difficulties, we shall choose a different gauge, which we shall call the *canonical heat-temporal gauge* or *caloric gauge*, which will be described later in this section and is constructed using the heat flow for the harmonic map equation; it is the covariant formulation of the microlocal gauge used in [64], [65], [27], [68], and places  $A$  in the schematic form  $A = \pi(\psi, \nabla_x^{-1}\psi)$ , where  $\pi$  is a (nonlinear) paraproduct, defined using the heat flow for the harmonic map equation, which effectively constrains the  $\nabla_x^{-1}\psi$  factor to have higher frequency than the  $\psi$  factor. This is similar in strength to the Coulomb gauge  $\nabla_x^{-1}(\psi\psi)$ , but mostly eliminates the ‘‘high-high interactions’’ where two high frequency components of  $\psi$  interact to create a low frequency component of  $A$ ; this is precisely the term for which the  $\nabla_x^{-1}$  operator causes difficulties. Also we do not encounter the  $\nabla^{-1}(AA)$  term at all.

As observed in [38], [51], [31], [32], the derivative field  $\psi$  evolves by a nonlinear wave equation. Indeed, from (15) we have

$$\partial_\alpha \psi_\beta - \partial_\beta \psi_\alpha = -A_\alpha \psi_\beta + A_\beta \psi_\alpha$$

and

$$\partial^\alpha \psi_\alpha = -A^\alpha \psi_\alpha$$

and hence

$$(17) \quad \begin{aligned} \partial^\alpha \partial_\alpha \psi_\beta &= \partial^\alpha (-A_\alpha \psi_\beta + A_\beta \psi_\alpha) + \partial^\alpha \partial_\beta \psi_\alpha \\ &= \partial^\alpha (A_\beta \psi_\alpha - A_\alpha \psi_\beta) - \partial_\beta (A^\alpha \psi_\alpha). \end{aligned}$$

Given that  $A$  should behave like a quadratic expression of  $\psi$ , containing an inverse derivative (cf. (16)), we thus see that (17) is something like a cubic wave equation for  $\psi$ .

As is well known (see e.g. [50]), the wave map equation enjoys a *stress energy tensor*

$$\mathbf{T}_{\alpha\beta} := \langle \partial_\alpha \phi, \partial_\beta \phi \rangle_h - \frac{1}{2} g_{\alpha\beta} \langle \partial^\gamma \phi, \partial_\gamma \phi \rangle_h,$$

or in terms of the differentiated fields

$$(18) \quad \mathbf{T}_{\alpha\beta} := \langle \psi_\alpha, \psi_\beta \rangle_{\mathbf{R}^m} - \frac{1}{2} g_{\alpha\beta} \langle \psi^\gamma, \psi_\gamma \rangle_{\mathbf{R}^m},$$

where  $g$  is the Minkowski metric; observe that this quantity is independent of the choice of gauge  $e$ . In particular we can define (in standard co-ordinates  $x^0 = t, x^1, x^2$ ) the *energy density*  $\mathbf{T}_{00}$  by

$$(19) \quad \mathbf{T}_{00}(t, x) := \frac{1}{2} |\psi_{t,x}(t, x)|^2.$$

Because the domain  $\mathbf{R}^{1+2}$  is invariant under translations in spacetime, Noether's theorem yields the conservation law

$$(20) \quad \partial^\alpha \mathbf{T}_{\alpha\beta} = 0$$

which can of course also be verified directly from (15). The conservation of the energy

$$(21) \quad E(\phi[t]) = \frac{1}{2} \int_{\mathbf{R}^2} |\psi_{t,x}|^2 dx = \int_{\mathbf{R}^2} \mathbf{T}_{00}(t, x) dx$$

now follows directly from (20).

*Remark 2.5.* In the case of a Euclidean target  $N = \mathbf{R}^m$  (so  $\kappa = 0$ ) and the standard orthonormal gauge  $e = \partial_{x_1}, \dots, \partial_{x_m}$ , we have  $\psi_\alpha = \partial_\alpha \phi$ ,  $A_\alpha = 0$ , and the wave map equation (2) then becomes  $\partial^\alpha \partial_\alpha \phi = 0$ , while (17) becomes  $\partial^\alpha \partial_\alpha \psi_\beta = 0$ . Thus in this case the wave maps equation collapses to the free (linear) wave equation.

## 2.6. The heat flow equation

We now leave wave maps for the moment and turn to the heat flow equation for harmonic maps (or *heat flow* for short). This equation can be written covariantly as

$$(22) \quad \partial_s \phi = (\phi^* \nabla)_k \partial_k \phi$$

where  $\phi(s, x) \in N$  is a smooth map defined for some interval  $[0, S) \times \mathbf{R}^2$  for some  $0 < S \leq +\infty$ ; formally, this is gradient flow for the energy functional  $\int_{\mathbf{R}^2} \frac{1}{2} |\nabla_x \phi|_h^2 dx$ . Actually, for our applications we will not consider just a single map  $\phi(s, x)$  evolving by the heat flow, but an entire one-parameter family  $\phi(s, t, x)$  of such maps, where  $t$  ranges in a compact time interval  $I$ . While the  $s$  variable plays the role of time in this equation, it is quite distinct from the time variable  $t$  appearing in the wave map equation, and so we shall refer to  $s$  as the *heat-temporal* variable and  $t$  as the *wave-temporal* variable to minimize confusion. We use Latin indices  $j, k$  to parameterize the spatial variable  $x$  (and write  $x$  to denote both indices, thus for instance  $A_x$  denotes the vector  $(A_1, A_2)$ ), and Greek indices  $\alpha, \beta$  to parameterize the spacetime variables  $t, x$ , but shall always treat the heat-temporal variable  $s$  separately (using neither Greek nor Latin indices to parameterize these). In our discussion of the heat flow the wave-temporal parameter  $t$  plays almost no role, although we will need to study the evolution of time derivatives such as  $\partial_t \phi$  in the  $s$  direction.

In local co-ordinates  $\phi^a$ , the equation (22) becomes a non-linear heat equation

$$(23) \quad \partial_s \phi^a = \Delta \phi^a + \Gamma_{bc}^a(\phi) \partial_\alpha \phi^b \partial^\alpha \phi^c$$

as such it is clear (e.g. from energy methods) that this equation is locally well posed in the Schwartz class, and the flow fixes the value  $\phi(\infty)$  of the map at infinity. Indeed, we have the global well-posedness result

**Theorem 2.7 (Global existence of heat flow).** *If  $\phi : I \times \mathbf{R}^2 \rightarrow N$  is Schwartz and equals  $\phi(\infty) \in N$  at infinity, then there is a unique Schwartz solution  $\phi : \mathbf{R}^+ \times I \times \mathbf{R}^2 \rightarrow N$  to (22) with initial data  $\phi(0, t, x) = \phi(t, x)$ . Furthermore,  $\phi(s)$  converges to  $\phi(\infty)$  in the  $C^m(I \times \mathbf{R}^2; N)$  topology as  $s \rightarrow +\infty$  for any  $m \geq 0$ .*

This result is essentially a special case of the results of Eells-Sampson [10] (see also [44], [35], [34], [58]), and follows primarily from the negative curvature of  $N$ . We will not prove this standard result here, but we mention that a key tool in this convergence result is the *Bocher identity*

$$(\partial_s - \Delta)|\psi_\alpha|^2 = \kappa |\psi_\alpha \wedge \psi_x|^2 - 2|\mathbf{D}_x \psi_\alpha|^2,$$

which in conjunction with the *diamagnetic inequality*  $|\nabla_x |\psi_\alpha|| \leq |\mathbf{D}_x \psi_\alpha|$  and the negative curvature hypothesis  $\kappa \leq 0$  implies that the evolution of  $|\psi_\alpha|$  is dominated by the heat flow:

$$(24) \quad (\partial_s - \Delta)|\psi_\alpha| \leq 0.$$

We can also write the heat flow equation in differentiated form by lifting  $\phi$  to  $(\phi, e) : \mathbf{R}^+ \times I \times \mathbf{R}^2 \rightarrow \text{Frame}(N)$  as before. We will however limit the gauge freedom here by imposing the *heat-temporal* gauge condition

$$(25) \quad (\phi^* \nabla)_s e = 0;$$

in other words, we require that  $e$  propagates in the heat-temporal direction  $\partial_s$  by parallel transport. It is clear geometrically (or from the Picard existence theorem) that for fixed  $\phi$ , the frame  $e(s, t, x)$  at  $(s, t, x)$  is now determined completely by the corresponding frame  $e(0, t, x)$  at  $s = 0$ . We still have the gauge freedom (14), but only if the gauge  $U = U(t, x)$  is independent of  $s$ , i.e. it depends only on  $t, x$ . We will remove this gauge freedom shortly, however, by selecting a canonical heat-temporal gauge.

Using this gauge  $e$  we can introduce the derivative fields  $\psi_\alpha, \psi_s, A_\alpha, A_s$  as before, although the heat-temporal gauge condition forces  $A_s = 0$  (or equivalently  $\mathbf{D}_s = \partial_s$ ). From (10) and (7) we see that (22) becomes  $\psi_s = \mathbf{D}_k \psi_k$ . In analogy to (15), we can now write the heat flow as the system

$$\begin{aligned}
(26) \quad & \partial_\alpha \phi = e \psi_\alpha \\
& \partial_s \phi = e \psi_s \\
& (\phi^* \nabla)_\alpha e = e A_\alpha \\
& (\phi^* \nabla)_s e = A_s = 0 \\
& \psi_s = \mathbf{D}_k \psi_k \\
& \mathbf{D}_\alpha \psi_\beta - \mathbf{D}_\beta \psi_\alpha = 0 \\
& \partial_s \psi_\alpha = \mathbf{D}_s \psi_\alpha = \mathbf{D}_\alpha \psi_s \\
& F_{\alpha\beta} = \kappa \psi_\alpha \wedge \psi_\beta \\
& \partial_s A_\alpha = F_{s\alpha} = \kappa \psi_s \wedge \psi_\alpha.
\end{aligned}$$

We refer to the quantity  $\psi_s = \mathbf{D}_k \psi_k$  as the *heat-tension field*; it is the direction that the heat flow evolution in the  $s$  variable deforms  $\phi$  in (with respect to the orthonormal frame  $e$ ), and should be thought of as a non-linear version of the Laplacian applied to the map  $\phi$ .

Similarly to (17), the derivative fields  $\psi_\alpha$  obey a (covariant) heat equation which is linear (but with potential and magnetic components). Indeed from (26) and (12) we see that

$$\begin{aligned}
(27) \quad & \partial_s \psi_\alpha = \mathbf{D}_\alpha \psi_s \\
& = \mathbf{D}_\alpha \mathbf{D}_k \psi_k \\
& = \mathbf{D}_k \mathbf{D}_\alpha \psi_k + F_{\alpha k} \psi_k \\
& = \mathbf{D}_k \mathbf{D}_k \psi_\alpha + \kappa (\psi_\alpha \wedge \psi_k) \psi_k;
\end{aligned}$$

this equation can be viewed as a linearization of the heat flow equation around  $\phi$ . Observe that the heat-tension field  $\psi_s$  also obeys the same equation:

$$(28) \quad \partial_s \psi_s = \mathbf{D}_k \mathbf{D}_k \psi_s + \kappa (\psi_s \wedge \psi_k) \psi_k.$$

As a first approximation, one can thus think of  $\psi_\alpha(s, t)$  as being a non-linear analogue of  $e^{s\Delta} \psi_\alpha(0, t)$ , where  $e^{s\Delta}$  is the propagator for the free (linear) heat equation. Similarly for the heat-tension field  $\psi_s$ . As is well known, heat operators can be used as a substitute for Littlewood-Paley operators; thus one can think of  $\psi_\alpha(s, t)$

heuristically as a nonlinear restriction of  $\psi_\alpha(0, t)$  to frequencies  $|\xi| \lesssim s^{-1/2}$ . The heat-tension field  $\psi_s$  is similar, but the presence of the derivative in the formula  $\psi_s = \mathbf{D}_k \psi_k$  suggests that this quantity is more localized to the annulus  $|\xi| \sim s^{-1/2}$  than the ball  $|\xi| \lesssim s^{-1/2}$ .

*Remark 2.8.* The assumption  $\kappa < 0$  will ensure that the equations (27), (28) will not blow up as  $s \rightarrow +\infty$ ; this is basically because of (24).

As mentioned above, we still have a gauge freedom to rotate  $e$  by an arbitrary gauge  $U(t, x)$  independent of  $s$ . However, we can remove this freedom by fixing a boundary condition at  $s = +\infty$ .

**Theorem 2.9 (Existence of canonical heat-temporal gauge).** *Suppose  $\phi : I \times \mathbf{R}^2 \rightarrow N$  is Schwartz and equals  $\phi(\infty) \in N$  at infinity, and let  $\phi : \mathbf{R}^+ \times I \times \mathbf{R}^2 \rightarrow N$  be the heat flow extension given by Theorem 2.7. Let  $e(\infty)$  be an orthonormal frame in  $T_{\phi(\infty)}N$ . Then there is a unique lift  $(\phi, e) : \mathbf{R}^+ \times I \times \mathbf{R}^2 \rightarrow \text{Frame}(N)$  which is smooth, obeys the heat-temporal condition (25), and which converges uniformly to  $(\phi(\infty), e(\infty))$  as either  $s \rightarrow +\infty$  or  $x \rightarrow \infty$ . Furthermore,  $A_\alpha(s)$  and  $\psi_\alpha(s)$  converge uniformly to zero as  $s \rightarrow +\infty$ .*

We will not prove this theorem in detail here, but follows from decay estimates on  $\psi$  as  $s \rightarrow +\infty$  which ultimately stem (24), which ensure that one can specify boundary data at  $s = +\infty$  in a well-posed manner.

We refer to the gauge  $e$  constructed in this manner as the *canonical heat-temporal gauge* or *caloric gauge* associated to  $\phi$  (and to  $e(\infty)$ ). Geometrically, it is constructed by dragging the constant frame  $(\phi(\infty), e(\infty))$  back via parallel transport from  $s = +\infty$  to  $s = 0$  by reversing the heat flow; it can also be viewed as the unique solution to (26) with boundary conditions

$$(29) \quad \begin{aligned} (\phi, e) &= (\phi(\infty), e(\infty)) \text{ when } s = +\infty \\ (\psi_\alpha, A_\alpha) &= 0 \text{ when } s = +\infty. \end{aligned}$$

Note that we are specifying data at both  $s = 0$  and  $s = +\infty$ . As we shall explain shortly, this gauge is the analogue of the microlocal gauge in [64], [65], [27], [68] in the differentiated setting, and will be used here in place of the Coulomb gauge that was used in [38], [51], [31], [32], for reasons which will also be discussed shortly.

Note that we can now compute  $A$  from  $\psi$  via the fundamental theorem of calculus and (26):

$$(30) \quad \begin{aligned} A_\alpha(s, t, x) &= - \int_s^{+\infty} \partial_s A_\alpha(s', t) ds' \\ &= -\kappa \int_s^{+\infty} \psi_s(s', t) \wedge \psi_\alpha(s', t) ds' \\ &= -\kappa \int_s^{+\infty} (\mathbf{D}_k \psi_k(s', t)) \wedge \psi_\alpha(s', t) ds', \end{aligned}$$



and in particular we can compute the connection  $A_\alpha(t, x)$  on the boundary  $s = 0$  by the formula

$$(31) \quad A_\alpha(0, t, x) = -\kappa \int_0^{+\infty} (\mathbf{D}_k \psi_k(s, t, x)) \wedge \psi_\alpha(s, t, x) ds.$$

If we continue the heuristic that  $\psi(s, t, x)$  is a non-linear version of  $e^{s\Delta}\psi(0, t, x)$ , then the above expression can then be viewed as a non-linear analogue of a paraproduct  $\pi(\psi_x(0), \nabla_x^{-1}\psi_x(0))$ , where the paraproduct  $\pi$  is a “low-high” paraproduct that restricts the frequency of the  $\nabla_x^{-1}\psi_x(0)$  factor to be larger or comparable to that of the  $\psi_x(0)$  factor. To see this informally, we restrict  $s$  to the dyadic region  $2^{-2\mathbf{k}} \leq s \leq 2^{-2\mathbf{k}+1}$ , and then heuristically view  $\psi_x(s, t)$  as  $P_{\leq \mathbf{k}}\psi_x(0, t)$ , a Littlewood-Paley type projection to frequencies  $\leq 2^{\mathbf{k}}$ , while  $\psi_s(s, t)\mathbf{D}_k\psi_k(s, t)$  is roughly like  $2^{\mathbf{k}}P_{\mathbf{k}}\psi_x(0, t)$ , a projection to frequencies  $\sim 2^{\mathbf{k}}$ . The expression (31) is then heuristically like  $\sum_{\mathbf{k}}(P_{\leq \mathbf{k}}\psi_x(0, t))(2^{-\mathbf{k}}P_{\mathbf{k}}\psi_x(0))$ , whence the claim. This paraproduct  $A = \pi(\psi_x(0), \nabla_x^{-1}\psi_x(0))$  is favorable as the inverse derivative  $\nabla_x^{-1}$  is always on the highest frequency factor; this should be compared with the Coulomb gauge, which gives a formula for  $A$  of the form  $A = \nabla_x^{-1}(\psi_x\psi_x)$  (see (16)), which is of similar strength to the paraproduct in (31) except for high-high interactions of  $\psi$ , for which the inverse derivative is in an unfavorable location; this is not particularly harmful in high dimensions (as the arguments in [38], [51] show) but begins to cause significant problems in two and three dimensions. In [31], [32] the difficulties arising from this unfavorable placement of the inverse derivative are overcome by using the overdetermined nature of the wave map system to extract additional cancellations from the expression  $\nabla_x^{-1}(\psi_x\psi_x)$ , and by quite delicate multilinear estimates. We will face similar difficulties, but they will be milder because of the more favorable location of the inverse derivative.

*Remark 2.10.* We now perform some heuristic manipulations that connect this gauge to the microlocal renormalization used for the sphere  $S^{m-1} \subset \mathbf{R}^m$  in [64], [65] (and can also be used to connect to the similar procedures used in [27], [68]). Here we think of the wave map  $\phi$  as a column vector with  $m$  entries. We choose an orthonormal frame  $e = (e_1, \dots, e_k)$  taking values not on the tangent space of the sphere, but rather in the ambient Euclidean space  $\mathbf{R}^m$ . In the standard orthonormal frame (which we call  $e^0$ ), the analogues of  $\psi_\alpha$  and  $A_\alpha$  are

$$\psi_\alpha := \partial_\alpha \phi; \quad A_\alpha := \partial_\alpha \phi \wedge \phi.$$

In a more general frame  $e = Ue^0$ , where  $U$  takes values in the rotation group  $O(m)$ , we have

$$\psi_\alpha := U^{-1}\partial_\alpha \phi; \quad A_\alpha := U^{-1}(\partial_\alpha \phi \wedge \phi)U - U^{-1}\partial_\alpha U.$$

The heat-temporal gauge condition  $A_s = 0$  thus becomes

$$\partial_s U = (\partial_s \phi(s) \wedge \phi(s))U(s)$$

where  $\partial_s \phi$  can be defined using the heat flow equation

$$\partial_s \phi = \Delta \phi + \phi |\nabla \phi|^2.$$

While this equation can blow up for large energies, for small energies it converges to a constant. In the region  $s \sim 2^{-2\mathbf{k}}$ , we can then heuristically treat  $\phi(s)$  as the low frequency projection  $P_{\leq \mathbf{k}}\phi(0)$ , and  $\partial_s\phi \approx \Delta\phi$  as the medium frequency projection  $s^{-1}P_{\mathbf{k}}\phi(0)$ . Thus the equation for  $U$  is then approximately

$$s\partial_s U \approx [P_{\mathbf{k}}\phi(0) \wedge P_{\leq \mathbf{k}}\phi(0)]U(s).$$

Comparing this with the scheme used in [64] (see also [65], [27], [68]) we see that the canonical heat-temporal gauge construction is nothing more than a continuous version of the discrete microlocal gauge used in those papers, with the Littlewood-Paley operators in the ambient Euclidean space being replaced by the more intrinsic operators generated by the harmonic map heat flow. Note that in the heat flow formulation, the matrices  $U$  are automatically orthogonal, whereas in the discrete procedure in [64], [65], [27], [68] one only obtains approximate orthogonality, and only when the energy is assumed small. Indeed for large energies it seems unlikely that one can replicate this procedure, due to topological obstructions (Schwartz maps from  $\mathbf{R}^2$  to  $S^{m-1}$  are not necessarily topologically trivial).

*Remark 2.11.* The equation (30) can be used to recover  $A$  from  $\psi_\alpha$ . One can similarly use (26), (30) and the canonical heat gauge condition  $\psi_\alpha(\infty, t, x) = 0$  to recover  $\psi_\alpha$  from the heat-tension field  $\psi_s$ :

$$\begin{aligned} \psi_\alpha(s_0, t) &= - \int_{s_0}^{+\infty} \partial_s \psi_\alpha(s, t) ds \\ (32) \qquad &= - \int_{s_0}^{+\infty} \mathbf{D}_\alpha \psi_s(s, t) ds \\ &= -\partial_\alpha \int_{s_0}^{\infty} \psi_s(s, t) ds + \Psi_\alpha(s_0, t) \end{aligned}$$

where  $\Psi$  is the cubic correction term

$$\begin{aligned} \Psi_\alpha(s_0, t) &:= \int_{s_0}^{+\infty} A_\alpha(s) \psi_s(s, t) ds \\ (33) \qquad &= -\partial_\alpha \int_{s_0}^{+\infty} \psi_s(s) ds + \kappa \int_{s_0 < s < s'} (\psi_s(s') \wedge \psi_\alpha(s')) \psi_s(s) ds' ds. \end{aligned}$$

Thus one can (at least in principle) recover  $\psi_\alpha$  from  $\psi_s$  by an iteration process, and then by (30) recover  $A_\alpha$  from  $\psi_s$  and  $\psi_\alpha$ . Thus one can view (26) as a scalar non-linear heat equation (28) for the heat-tension field  $\psi_s$ . We will thus often seek to use the above identities to write the other components of  $\psi$  and  $A$  in terms of this “dynamic variable”  $\psi_s$  in order to make the system (34) resemble a scalar equation (and thus make more amenable to solution by iterative methods); this corresponds to the “dynamic separation” strategy used in [31], [32]. Indeed one can view the heat-tension field  $\psi_s$  as a sort of<sup>7</sup> continuous Littlewood-Paley resolution of  $\phi$ , which is the scalar variable in the undifferentiated wave map equation (2). One consequence

---

<sup>7</sup>A more precise heuristic is that if  $s \sim 2^{-2\mathbf{k}}$ , then  $\psi_s \sim \frac{1}{s} P_{\mathbf{k}}\phi$ .

in particular from (32) that the functions  $\psi_\alpha$  are (modulo a cubic lower order term  $\Psi_\alpha$ ) the gradient of a scalar function  $-\int_{s_0}^\infty \psi_s(s) ds$ . This fact will be crucial for us, as it means that all of the highest order non-linear expressions involving the  $\psi_\alpha$  contain null form structure (the lower order terms will turn out to be manageable via Strichartz estimates and will not require null structure). In particular we will seek to recover null forms such as the trilinear expression  $\phi_1 \partial^\alpha \phi_2 \partial_\alpha \phi_3$ , preferably with  $\phi_1$  having “higher frequency” than  $\phi_2$  or  $\phi_3$ , in the spirit of [65]; in our heat-flow formulation the  $s$  parameter will be a proxy for this concept of frequency<sup>8</sup>. A similar strategy was carried out in [31], [32], starting not from the heat flow but instead from the equation  $\mathbf{D}_\alpha \psi_\beta - \mathbf{D}_\beta \psi_\alpha = 0$  to conclude that  $\psi$  has small curl and hence (by Hodge theory) is approximately a gradient. However the use of Hodge theory once again introduces inverse derivatives  $\nabla_x^{-1}$  into the equation, which makes the non-linearities more difficult to control than would otherwise be necessary, and so we shall rely on (32) instead of Hodge theory to write  $\psi$  as an approximate gradient.

## 2.12. Concatenating the wave map and heat flow equations

From Theorems 2.7 and 2.9 we see that the system (26) can be solved, and placed in the canonical heat-temporal gauge (29), given any Schwartz map  $\phi : I \times \mathbf{R}^2 \rightarrow N$  (as well as an orthonormal frame  $e(\infty)$  in  $T_{\phi(\infty)}N$ ). In particular we can solve this equation for any Schwartz solution to the wave map equation (15). This gives rise to the fields  $\phi, e, A_\alpha, \psi_\alpha, A_s, \psi_s$  on  $\mathbf{R}^+ \times I \times \mathbf{R}^2$  solving the system of equations and boundary conditions

$$\begin{aligned}
& \partial_\alpha \phi = e \psi_\alpha \\
& \partial_s \phi = e \psi_s \\
& (\phi^* \nabla)_\alpha e = e A_\alpha \\
& (\phi^* \nabla)_s e = A_s = 0 \\
& \mathbf{D}^\alpha \psi_\alpha = 0 \text{ when } s = 0 \\
& \psi_s = \mathbf{D}_k \psi_k \\
(34) \quad & \mathbf{D}_\alpha \psi_\beta - \mathbf{D}_\beta \psi_\alpha = 0 \\
& \partial_s \psi_\alpha = \mathbf{D}_s \psi_\alpha = \mathbf{D}_\alpha \psi_s \\
& F_{\alpha\beta} = \kappa \psi_\alpha \wedge \psi_\beta \\
& \partial_s A_\alpha = F_{s\alpha} = \kappa \psi_s \wedge \psi_\alpha \\
& (\phi, e) = (\phi(\infty), e(\infty)) \text{ when } s = +\infty \\
& (\psi_\alpha, A_\alpha) = 0 \text{ when } s = +\infty.
\end{aligned}$$

Of course, many of these equations are redundant, and this system is highly overdetermined, nevertheless it will be convenient to retain all of these equations as we will need the full structure of this system at various points in the argument.

---

<sup>8</sup>More precisely, the quantity  $s^{-1/2}$  corresponds to the Fourier notion of frequency magnitude, as can be seen by inspecting the Fourier symbol of the heat operator  $e^{s\Delta}$ ; thus larger values of  $s$  correspond to lower frequencies and vice versa.

*Remark 2.13.* If the target  $N$  were Euclidean space  $N = \mathbf{R}^m$  with the standard frame, then we have  $\psi_\alpha = \partial_\alpha \phi$ ,  $\psi_s = \Delta \phi$ ,  $A_\alpha = A_s = 0$ , and the heat flow equation (22) then becomes  $\partial_s \phi = \Delta \phi$ , while (27) becomes  $\partial_s \psi_\alpha = \Delta \psi_\alpha$ . Thus in this case the heat flow collapses to the free (linear) heat equation. It is clear that (34) similarly collapses to an evolution which is the wave equation in the  $x, t$  variables and the heat equation in the  $x, s$  variables, with the two flows commuting with each other. For non-Euclidean targets, the same statement is true to top order; thus all fields evolve in the  $s$  direction by a nonlinear heat equation and in the  $t$  direction by a nonlinear wave equation.

*Remark 2.14.* One can use the transitive group action of  $O(1, m)$  of  $\text{Frame}(N)$  to place  $\phi(\infty)$  and  $e(\infty)$  wherever one pleases. However there will be no need for us to perform such a normalization.

*Remark 2.15.* We caution the reader that the heat-temporal variable  $s$  scales like twice the dimension of space, in contrast to the wave-temporal variable  $t$ , which scales like one dimension of space. Indeed, the scale invariance of (34) is given by

$$\begin{aligned}
(\phi, e)(s, t, x) &\leftarrow (\phi, e)\left(\frac{s}{\lambda^2}, \frac{t}{\lambda}, \frac{x}{\lambda}\right) \\
(\psi_\alpha, A_\alpha)(s, t, x) &\leftarrow \frac{1}{\lambda}(\psi_\alpha, A_\alpha)\left(\frac{s}{\lambda^2}, \frac{t}{\lambda}, \frac{x}{\lambda}\right) \\
\psi_s(s, t, x) &\leftarrow \frac{1}{\lambda^2}\psi_s\left(\frac{s}{\lambda^2}, \frac{t}{\lambda}, \frac{x}{\lambda}\right) \\
F_{\alpha\beta}(s, t, x) &\leftarrow \frac{1}{\lambda^2}F_{\alpha\beta}\left(\frac{s}{\lambda^2}, \frac{t}{\lambda}, \frac{x}{\lambda}\right) \\
F_{s\alpha}(s, t, x) &\leftarrow \frac{1}{\lambda^3}F_{s\alpha}\left(\frac{s}{\lambda^2}, \frac{t}{\lambda}, \frac{x}{\lambda}\right).
\end{aligned}
\tag{35}$$

Note that the wave map equation  $\mathbf{D}^\alpha \psi_\alpha = 0$  only holds at the boundary  $s = 0$  of the slab  $\mathbf{R}^+ \times I \times \mathbf{R}^2$ . Unfortunately, this equation does not commute with the heat flow equation (even in the completely integrable one-dimensional case  $n = 1$ ), so we do not obtain this equation in the interior of the slab. However, we can assert that the wave map equation holds “approximately” in the interior region. Indeed, if we define the *wave-tension field*  $u$  by  $u := \mathbf{D}^\alpha \psi_\alpha$ , then we have (from various equations

in (34) and covariant versions of the Leibnitz rule)

$$\begin{aligned}
(36) \quad \partial_s u &= \mathbf{D}_s \mathbf{D}^\alpha \psi_\alpha \\
&= F_{s\alpha} \psi^\alpha + \mathbf{D}^\alpha \mathbf{D}_s \psi_\alpha \\
&= \kappa(\psi_s \wedge \psi_\alpha) \psi^\alpha + \mathbf{D}^\alpha \mathbf{D}_\alpha \psi_s \\
&= \kappa(\mathbf{D}_k \psi_k \wedge \psi_\alpha) \psi^\alpha + \mathbf{D}^\alpha \mathbf{D}_\alpha \mathbf{D}_k \psi_k \\
&= \kappa(\mathbf{D}_k \psi_k \wedge \psi_\alpha) \psi^\alpha + \mathbf{D}^\alpha (F_{\alpha k} \psi_k) + \mathbf{D}^\alpha \mathbf{D}_k \mathbf{D}_\alpha \psi_k \\
&= \kappa[(\mathbf{D}_k \psi_k \wedge \psi_\alpha) \psi^\alpha + \mathbf{D}^\alpha((\psi_\alpha \wedge \psi_k) \psi_k)] + \mathbf{D}^\alpha \mathbf{D}_k \mathbf{D}_k \psi_\alpha \\
&= \kappa[(\mathbf{D}_k \psi_k \wedge \psi_\alpha) \psi^\alpha + (\mathbf{D}^\alpha \psi_\alpha \wedge \psi_k) \psi_k + (\psi_\alpha \wedge \mathbf{D}^\alpha \psi_k) \psi_k + (\psi_\alpha \wedge \psi_k) \mathbf{D}^\alpha \psi_k] \\
&\quad + F_{\alpha k} \mathbf{D}_k \psi^\alpha + \mathbf{D}_k (F_{\alpha k} \psi^\alpha) + \mathbf{D}_k \mathbf{D}_k \mathbf{D}^\alpha \psi_\alpha \\
&= \kappa[(\mathbf{D}_k \psi_k \wedge \psi_\alpha) \psi^\alpha + (u \wedge \psi_k) \psi_k + (\psi_\alpha \wedge \mathbf{D}_k \psi^\alpha) \psi_k + (\psi_\alpha \wedge \psi_k) \mathbf{D}^\alpha \psi_k] \\
&\quad + 2(\psi_\alpha \wedge \psi_k) \mathbf{D}_k \psi^\alpha + (\mathbf{D}_k \psi_\alpha \wedge \psi_k) \psi^\alpha + (\psi_\alpha \wedge \mathbf{D}_k \psi_k) \psi^\alpha + \mathbf{D}_k \mathbf{D}_k u \\
&= \mathbf{D}_k \mathbf{D}_k u + \kappa(u \wedge \psi_k) \psi_k + \kappa[(\psi_\alpha \wedge \mathbf{D}_k \psi^\alpha) \psi_k + 3(\psi_\alpha \wedge \psi_k) \mathbf{D}_k \psi^\alpha + (\mathbf{D}_k \psi_\alpha \wedge \psi_k) \psi^\alpha] \\
&= \mathbf{D}_k \mathbf{D}_k u + \kappa(u \wedge \psi_k) \psi_k + 4\kappa(\psi_\alpha \wedge \psi_k) \mathbf{D}_k \psi^\alpha.
\end{aligned}$$

Thus the wave-tension field  $u$  propagates in the  $s$  direction by a covariant, inhomogeneous heat equation. Also, from (34) we have  $u = 0$  on the boundary  $s = 0$ . Thus we expect  $u$  to be somewhat “small” for  $s > 0$  as well.

Just as the wave-tension field obeys a heat equation in  $s$ , the heat-tension field  $\psi_s = \mathbf{D}_k \psi_k$  obeys a wave equation in  $t$ :

$$\begin{aligned}
(37) \quad \partial^\alpha \partial_\alpha \psi_s &= \partial^\alpha (\mathbf{D}_\alpha \psi_s - A_\alpha \psi_s) \\
&= \partial^\alpha \partial_s \psi_\alpha - \partial^\alpha (A_\alpha \psi_s) \\
&= \partial_s (\mathbf{D}^\alpha - A^\alpha) \psi_\alpha - \partial^\alpha (A_\alpha \psi_s) \\
&= \partial_s u - \partial_s (A^\alpha \psi_\alpha) - \partial^\alpha (A_\alpha \psi_s).
\end{aligned}$$

One can use (36) and (37) to design a (local-in-time) iteration scheme for solving (34) in the large energy setting. The rigorous details of this scheme will appear elsewhere, but the informal ideas are as follows. We first use (36) propagate the wave-tension field  $u$  in the  $s$  direction (using values of  $\psi_{t,x}$  and  $A_{t,x}$  obtained from previous iterates), using the boundary condition  $u = 0$  at  $s = 0$ . Then one uses (37) to propagate the heat-tension field  $\psi_s$  forward in the  $t$  direction (using the value of  $u$  just obtained, and values of  $A_\alpha$  and  $\psi_s$  obtained from previous iterates), using the initial data of  $\psi_s$  at the initial time  $t = t_0$ . Finally, one uses (32) and (30) to compute  $\psi_{t,x}$  and  $A_{t,x}$  from  $\psi_s$  (and using values of  $\psi_{t,x}$  obtained from previous iterates), using the boundary condition  $\psi_{t,x} = A_{t,x} = 0$  at  $s = +\infty$ . This scheme may appear convoluted, but it is necessary to do this in the large energy setting in order that the iterative algorithms to compute each of the fields  $\psi_{t,x}$ ,  $A_{t,x}$ ,  $\psi_s$  eventually involve some integration in time (so that we can take advantage of time localization). In the small energy case there are simpler schemes available<sup>9</sup>, but the

---

<sup>9</sup>For instance, one can propagate  $\psi_{t,x}$  forward in  $t$  on the boundary  $s = 0$  using (17), then

convergence of such schemes in the large energy case becomes problematic because of the possibility of an infinite number of iterations backwards and forwards in the  $s$  variable (which cannot be made to converge by localizing  $t$ ).

*Remark 2.16.* An overly simplified model of the system (34) can be obtained by assuming the approximation  $\psi_\alpha \approx \partial_\alpha \Phi$  for some function  $\Phi$  (cf. (32), (11)), and assuming that the nonlinear heat flow behaves like the linear heat equation (and in particular behaves somewhat like a family of Littlewood-Paley operators); then the equation for  $\Phi$  on the boundary  $s = 0$  is roughly of the form

$$\square \Phi \approx \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3: \mathbf{k}_1 \geq \min(\mathbf{k}_2, \mathbf{k}_3)} \mathbf{O}((P_{\mathbf{k}_1} \Phi)(\partial_\alpha P_{\mathbf{k}_2} \Phi)(\partial^\alpha P_{\mathbf{k}_3} \Phi));$$

thus  $\Phi$  evolves by a nonlinear wave equation with non-linearity of the form  $\Phi \partial_\alpha \Phi \partial^\alpha \Phi$ , but with at least one of the derivatives falling on a low frequency term. This is essentially the type of renormalized equation obtained for the wave map flow when the target is a sphere [65] or an isometrically embedded manifold in Euclidean space [68], and it is a good heuristic to keep in mind for the local theory (including the large data perturbation theory). For comparison, the corresponding heuristic equation for the Coulomb gauge formulation of the wave map equation would look something like

$$\square \Phi \approx \nabla_x^{-1} [\nabla_x \Phi (\partial_\alpha \Phi)] \partial^\alpha \Phi$$

(see [51], [38], [31] and especially [32]), which is a similar type of equation but has an additional divergence arising from “high-high” interactions inside the  $\nabla_x^{-1}$  term that must be controlled<sup>10</sup>.

### 3. Asymptotic local self-similarity

We now discuss the asymptotic self-similarity of wave maps that concentrate at a point. Our arguments here are based entirely on an analysis of the stress energy tensor  $\mathbf{T}_{\alpha\beta}$ , and in particular are independent of the choice of gauge or on the curvature properties of the manifold.

---

propagate in  $s$  using (27), and then recover  $A$  using (30). One can pursue these ideas to obtain a new proof of the small energy regularity results, at least in the constant negative curvature case, which is close in spirit to the arguments in [65] and particularly in [68], but we will not do so here. One could also abandon the caloric gauge, and use the Coulomb gauge instead; for instance, the arguments in [32] can (in principle) be modified to obtain an iteration scheme when the target is  $\mathbf{H}^2$ . However the  $\nabla^{-1}$  factors arising from this gauge would require one to prove far more delicate multilinear estimates, and in particular it becomes even more difficult to take advantage of the time localization than in our current argument. Also the Coulomb gauge has some uniqueness problems in the large energy non-abelian ( $m > 2$ ) case, being in some sense even more non-local than the caloric gauge.

<sup>10</sup>There is an additional null structure on the above trilinear form of “ $Q_{ij}$  type” which can be used to handle this divergence, see [32]. However, in the caloric gauge we neither have the divergence, nor the null structure required to deal with that divergence, and so the analytical treatment is somewhat simpler in this gauge.

Based in prior experience with large data critical non-linear wave equation regularity results (see e.g. [50] for some examples), one might hope (in the negative curvature case) to establish an energy non-concentration result, for instance showing that

$$(38) \quad \lim_{t \rightarrow 0^-} \int_{|x| < |t|} \mathbf{T}_{00}(t, x) dx = 0$$

whenever  $\psi$  is a Schwartz derivative wave map on  $[-1, 0) \times \mathbf{R}^2$ ; this for instance would be sufficient to establish global regularity for large energy wave maps by combining such an energy non-concentration result with a small energy regularity result (e.g. the one in [32]) and exploiting a finite speed of propagation argument<sup>11</sup>.

As is well known, one can hope to obtain (38), at least for certain components of the energy density  $\mathbf{T}_{00}$ , by contracting the stress energy tensor  $\mathbf{T}_{\alpha\beta}$  against some well-chosen vector field<sup>12</sup>  $X^\beta$  in spacetime<sup>13</sup> and then applying Stokes' theorem on a truncated backwards light cone such as  $\{|x| < |t|; -T < t < -\varepsilon\}$ ; the point is that (20) allows us to compute the divergence of  $\mathbf{T}_{\alpha\beta}X^\beta$  as<sup>14</sup>

$$(39) \quad \partial^\alpha(\mathbf{T}_{\alpha\beta}X^\beta) = \mathbf{T}_{\alpha\beta}\partial^\alpha X^\beta.$$

The boundary terms on the cone can be handled by a standard flux decay argument and can be considered negligible as a first approximation. The strategy is then to choose  $X^\beta$  so that the expression in (39) (as well as the boundary terms  $\mathbf{T}_{0\beta}X^\beta$ ) consists of terms which are either non-negative or small.

In the case of spherical symmetry or equivariant symmetry, one can demonstrate energy decay away from the time axis  $\{x = 0\}$ , not directly by the above strategy, but rather by taking advantage of the symmetry assumptions to eliminate the angular components of the stress-energy tensor, thus reducing (20) to what is essentially a set of transport equations for the remaining components of the stress-energy tensor in the null directions  $\partial_t \pm \partial_r$ . In particular it is relatively easy (via a Gronwall inequality argument, see e.g. [50]) to show in these special cases that

$$(40) \quad \lim_{t \rightarrow 0^-} \int_{\lambda|t| < |x| < |t|} \mathbf{T}_{00}(t, x) dx = 0 \text{ for any } 0 < \lambda < 1.$$

If one then applies (39) to vector fields such as  $X = r\partial_r = x \cdot \nabla_x$  and applies the above estimate, one can easily obtain (time-averaged) decay of the time component

---

<sup>11</sup>Actually, there is a slight difficulty in truncating the wave map properly to take advantage of finite speed of propagation, as one has to take a little care to ensure that the truncated wave map still has small energy and obeys the required compatibility conditions; we will not discuss these (minor) technicalities here however. See [68], [69] for some closely related issues.

<sup>12</sup>We shall abuse notation and identify vector fields  $X = X^\beta$  with their corresponding first order differential operators  $X = X^\beta\partial_\beta$  without further comment.

<sup>13</sup>In related equations such as the semilinear wave equation, one often adds lower order correction terms to  $\mathbf{T}_{\alpha\beta}X^\beta$ ; however there are no such lower-order terms available for wave maps.

<sup>14</sup>One can of course exploit the symmetry of  $\mathbf{T}_{\alpha\beta}$  to replace  $\partial^\alpha X^\beta$  by the more symmetric *deformation tensor*  $\pi^{\alpha\beta} := \frac{1}{2}(\partial^\alpha X^\beta + \partial^\beta X^\alpha)$ , but we shall not need to do so here.

of the energy:

$$(41) \quad \lim_{T \rightarrow 0^-} \frac{1}{T} \int_{-T}^0 \int_{|x| < |t|} |\psi_0(t, x)|^2 dx dt = 0.$$

From this and the negative curvature of  $N$  we can then obtain decay of the spatial component of the energy also.

In the absence of symmetry, one cannot hope to apply this type of argument directly. First of all, the stress-energy tensor no longer propagates solely in the null directions  $\partial_t \pm \partial_r$ , but can now propagate in any timelike or lightlike direction, which seems to frustrate any attempt to use Gronwall's inequality to obtain much decay, even very close to the light cone. A more serious objection, however, arises from the Lorentz invariance of the wave map equation once the symmetry assumptions are removed. Indeed, if one could prove the decay estimate (40) away from the time axis for arbitrary Schwartz wave maps, then by Lorentz invariance one could also prove a similar decay estimate away from any other timelike ray emanating backwards in time from the spacetime origin  $(0, 0)$ . Using two such disjoint timelike rays, it is then an easy matter to obtain (38). Thus there is no advantage in excluding the time axis in proving (40) for general wave maps; this estimate is as difficult as the original estimate (38). In particular, one cannot hope to obtain an estimate such as (40) without<sup>15</sup> relying more heavily on the negative curvature of  $N$ ; this in particular rules out the possibility of proving (40) purely from an analysis of the stress-energy tensor. Similar considerations hold for (41), basically because the vector field  $\partial_t$  is not Lorentz-invariant.

However, a closer inspection of (41) shows that, in light of (40), the content of (41) is only new when  $(t, x)$  is close to the time axis; thus we only need the relevant vector field to point in the direction of  $\partial_t$  near the time axis. This now allows for the possibility of a Lorentz-invariant estimate, for instance using the scaling vector field  $S := x^\alpha \partial_\alpha = t \partial_t + x \cdot \nabla_x$ , or more generally  $S/\rho^k$ , where  $\rho := \sqrt{-x^\alpha x_\alpha} = (|t|^2 - |x|^2)^{1/2}$  and  $k \in \mathbf{R}$  is some parameter. For *harmonic* maps in three dimensions  $\mathbf{R}^3$  it is well known (see e.g. [43]) that the choice  $k = 1$  (which would give the radial vector field  $\partial_r$  in three dimensions) yields a useful monotonicity formula of Pohozaev type; in light of the algebraic analogy between Euclidean space  $\mathbf{R}^3$  and Minkowski space  $\mathbf{R}^{1+2}$  it is then natural to try the same thing for wave maps on  $\mathbf{R}^{1+2}$ . This indeed works, and gives the following decay estimate:

**Proposition 3.1 (Asymptotic local self-similarity).** *Let  $\psi$  be a Schwartz derivative wave map on  $[-1, 0) \times \mathbf{R}^2$  (not assumed to have any symmetry assumptions). Then*

$$(42) \quad \lim_{T \rightarrow 0^-} \frac{1}{|\log |T||} \int_{-1}^{-T} \int_{|x| < |t|} \left| \frac{1}{t} \psi_S(t, x) \right|^2 dx \frac{dt}{t} = 0,$$

---

<sup>15</sup>For instance, the numerical work in [2] strongly suggests that when the target is  $S^2$  and the initial data is equivariant, then the energy can concentrate along the time axis leading to blowup. Applying a Lorentz transformation, one can then obtain non-equivariant initial data which concentrates on another timelike ray, and in particular (40) fails for this choice of data.



where  $\psi_S$  is the component of  $\psi$  in the scaling direction  $S$ , thus

$$\frac{1}{t}\psi_S := \frac{1}{t}S^\alpha\psi_\alpha = \psi_0 + \frac{x_j}{t}\psi_j.$$

This estimate should be compared with (41). On one hand, it replaces the vector field  $\partial_t$  by the variant  $\frac{1}{t}S = \partial_t + \frac{x}{t} \cdot \nabla_x$ , which is (conformally) invariant under mild Lorentz transformations (i.e. Lorentz transformations with bounded coefficients<sup>16</sup>) and applies for general wave maps. On the other hand, it requires somewhat more averaging in time than is present in (41). Note that conservation of energy only allows us to say that the expression in the left-hand side of (42) is bounded, but not that it decays to zero.

We now give a proof of the above proposition.

**Proof** [Proof of Proposition 3.1] Let  $\psi$  be a smooth derivative wave map on  $[-1, 0) \times \mathbf{R}^2$ . We now elaborate the stress energy analysis begun in (39). For any two times  $-1 \leq t_2 < t_1 < 0$ , define the truncated solid cone

$$K_{[t_2, t_1]} := \{(t, x) : t_2 \leq t \leq t_1; |x| \leq |t|\}.$$

This truncated cone has the boundary  $\partial K_{t_2} \cup \partial K_{(t_2, t_1)} \cup \partial K_{t_1}$ , where  $\partial K_t$  is the disk

$$\partial K_t := \{(t, x) : |x| \leq |t|\} \subset \{t\} \times \mathbf{R}^2$$

and  $\partial K_{(t_2, t_1)}$  is the truncated light cone

$$\partial K_{(t_2, t_1)} := \{(t, x) : t_2 < t < t_1; |x| = |t|\}.$$

The surfaces  $\partial K_t$  can be given Lebesgue measure  $dx$  (which co-incides with the measure induced from the ambient Minkowski metric). The surfaces  $\partial K_{(t_2, t_1)}$  are null and thus have no canonical measure, however we will give them the artificial measure  $d\sigma$ , defined by

$$\int_{\partial K_{(t_2, t_1)}} u \, d\sigma := \int_{t_2}^{t_1} \int_{S^1} u(t, |t|\omega) |t| \, d\omega \, dt$$

for all test functions  $u$ , where  $d\omega$  is arclength measure on the circle  $S_1$ . From Stokes' theorem we thus have

$$-\int_{K_{[t_2, t_1]}} \partial^\alpha P_\alpha(t, x) \, dt dx = \int_{\partial K_{t_1}} P_0(t_2, x) \, dx + \int_{\partial K_{(t_2, t_1)}} P_L \, d\sigma - \int_{\partial K_{t_2}} P_0(t_1, x) \, dx,$$

for any one-form  $P_\alpha$  smooth on  $K_{[t_2, t_1]}$ , where

$$P_L(t, x) := P_0(t, x) - \frac{x_j}{|x|} P_j(t, x)$$

---

<sup>16</sup>We do not expect our estimates to be invariant under extreme Lorentz transformations since such transformations tend to increase the energy substantially.

is the  $L := \partial_t - \partial_r$  component of  $P$ . In particular, from (39) we have

$$(43) \quad - \int_{K_{[t_2, t_1]}} \mathbf{T}_{\alpha\beta} \partial^\alpha X^\beta dx dt = \int_{\partial K_{t_1}} \mathbf{T}_{0\beta} X^\beta dx + \int_{\partial K_{(t_2, t_1)}} \mathbf{T}_{L\beta} X^\beta d\sigma - \int_{\partial K_{t_2}} \mathbf{T}_{0\beta} X^\beta dx$$

for any vector field  $X^\beta$  smooth on  $K_{[t_2, t_1]}$ .

We can now derive the standard facts of energy monotonicity and flux decay by choosing the vector field  $X = \partial_t$  in (43), obtaining the energy identity

$$(44) \quad 0 = \int_{\partial K_{t_1}} \mathbf{T}_{00} dx + \int_{\partial K_{(t_2, t_1)}} \mathbf{T}_{L0} d\sigma - \int_{\partial K_{t_2}} \mathbf{T}_{00} dx.$$

From (18) one easily verifies that

$$T_{L0} = \frac{1}{2} |\psi_L|_N^2 + \frac{1}{2} |\psi_{\nabla}|_N^2 \geq 0$$

on  $\partial K_{(t_2, t_1)}$ , where  $\nabla := \frac{x_1}{|x|} \partial_{x_2} - \frac{x_2}{|x|} \partial_{x_1}$  is the spatial angular derivative, thus  $\psi_{\nabla} = \frac{x_1}{|x|} \psi_2 - \frac{x_2}{|x|} \psi_1$ . In particular, we have *energy monotonicity*

$$\int_{\partial K_{t_1}} \mathbf{T}_{00} dx \leq \int_{\partial K_{t_2}} \mathbf{T}_{00} dx \text{ for all } -1 \leq t_2 \leq t_1 < 0;$$

if we then set  $E_0 := \int_{\partial K_{-1}} \mathbf{T}_{00} dx$ , then we have

$$(45) \quad \int_{\partial K_t} \mathbf{T}_{00} dx \leq E_0 \text{ for all } -1 \leq t < 0.$$

From this, (44), and monotone convergence we obtain the bounded flux property

$$\int_{\partial K_{(-1, 0)}} \mathbf{T}_{L0} d\sigma \leq E_0 < \infty.$$

In particular we have *flux decay*

$$(46) \quad \lim_{t \rightarrow 0^-} \int_{\partial K_{(t, 0)}} \mathbf{T}_{L0} d\sigma = 0.$$

Now let  $\lambda > 4$  be a large number, and let  $-1/\lambda < t_1 < 0$ . We set  $t_2 := \lambda t_1$ , and let  $\eta(t, x) = \eta(t)$  be a time cutoff which equals 1 when  $t_2/2 < t < 2t_1$ , vanishes when  $t < t_2$  or  $t > t_1$ , and is smooth otherwise (so in particular  $\eta'(t) = O(1/t_2)$  when  $t_2 \leq t \leq t_2/2$  and  $\eta'(t) = O(1/t_1)$  when  $2t_1 \leq t \leq t_1$ ). We would now like to apply (43) to the vector field

$$X^\beta := \frac{\eta x^\beta}{\rho} + (\partial^\beta \eta) \rho,$$

where  $\rho := \sqrt{-x^\alpha x_\alpha} = \sqrt{|t|^2 - |x|^2}$ , as the left-hand side of (43) will then give something very similar to the expression in (42). Unfortunately this vector field

is singular on the light cone and so (43) does not directly apply. To resolve this problem we shall mollify the above vector field slightly. Let  $0 < \varepsilon \leq 1$  be a small parameter, and define the modified co-ordinates  $\tilde{x}^\alpha$  and  $\tilde{\rho}$  by

$$\tilde{x}^0 := x^0 + \varepsilon t_1; \quad \tilde{x}^j := x^j; \quad \tilde{\rho} = \sqrt{-\tilde{x}^\alpha \tilde{x}_\alpha}$$

for  $j = 1, 2$ ; thus the  $\tilde{x}^\alpha$  co-ordinates are just the  $x^\alpha$  co-ordinates but with the origin shifted upwards to  $(-\varepsilon t_1, 0)$ . In particular  $\tilde{\rho}$  is now smooth on all of  $K_{[t_2, t_1]}$ . We now apply (43) with the vector field

$$\tilde{X}^\beta := \frac{\eta \tilde{x}^\beta}{\tilde{\rho}} + (\partial^\beta \eta) \tilde{\rho}.$$

Using the identities

$$\tilde{\rho}^2 = -\tilde{x}^\gamma \tilde{x}_\gamma; \quad \partial^\alpha \tilde{\rho} = -\frac{\tilde{x}^\alpha}{\tilde{\rho}}; \quad \partial^\alpha \tilde{x}^\beta = g^{\alpha\beta},$$

where  $g$  is the Minkowski metric, we see that

$$\partial^\alpha \tilde{X}^\beta = \frac{\eta}{\tilde{\rho}^3} (\tilde{x}^\alpha \tilde{x}^\beta - g^{\alpha\beta} \tilde{x}^\gamma \tilde{x}_\gamma) + \rho \partial^\alpha \partial^\beta \eta.$$

Contracting this against  $\mathbf{T}_{\alpha\beta}$  using (18) (and the fact that  $g^{\alpha\beta} g_{\alpha\beta} = \dim(\mathbf{R}^{1+2}) = 3$ ), we obtain

$$\mathbf{T}_{\alpha\beta} \partial^\alpha \tilde{X}^\beta = \frac{\eta}{\tilde{\rho}^3} |\tilde{x}^\alpha \psi_\alpha|^2 + \rho \mathbf{T}_{\alpha\beta} \partial^\alpha \partial^\beta \eta.$$

Inserting this into (43), and observing that the cutoff  $\eta$  eliminates the boundary terms at  $\partial K_{t_1}$ ,  $\partial K_{t_2}$ , we obtain

$$(47) \quad - \int_{K_{[t_2, t_1]}} \frac{\eta}{\tilde{\rho}^3} |\tilde{x}^\alpha \psi_\alpha|^2 + \rho \mathbf{T}_{\alpha\beta} \partial^\alpha \partial^\beta \eta \, dx dt = \int_{\partial K_{(t_2, t_1)}} \mathbf{T}_{L\beta} \left( \frac{\eta \tilde{x}^\beta}{\tilde{\rho}} + (\partial^\beta \eta) \tilde{\rho} \right) d\sigma.$$

Consider the second term on the left-hand side, which simplifies to  $\rho \mathbf{T}_{00} \eta''$ . This is supported on the regions  $t_2 < t < t_2/2$  and  $2t_1 < t < t_1$ , and we have the crude bound  $\rho \eta'' = O(1/t)$  on those regions. Thus by the energy bound (45), the contribution of this second term is  $O(E_0)$ . Now consider the right-hand side of (47). From (18) it is easy to show<sup>17</sup> that  $\mathbf{T}_{L\beta} = O(\mathbf{T}_{L0})$  for all  $\beta$ . Also, we have the pointwise bound

$$\left| \frac{\eta \tilde{x}^\beta}{\tilde{\rho}} + (\partial^\beta \eta) \tilde{\rho} \right| \leq C(\lambda, \varepsilon) < \infty$$

on  $\partial K_{(t_2, t_1)}$ , for some quantity  $C(\lambda, \varepsilon)$  depending on  $\lambda$  and  $\varepsilon$  (the value of  $C$  may vary from line to line). Thus we have

$$\int_{K_{[t_2, t_1]}} \frac{\eta}{\tilde{\rho}^3} |\tilde{x}^\alpha \psi_\alpha|^2 \, dx dt \leq C E_0 + C(\lambda, \varepsilon) \int_{\partial K_{(t_2, t_1)}} \mathbf{T}_{L0} \, d\sigma.$$

---

<sup>17</sup>This is best seen using the null frame  $L$ ,  $\nabla$ , and  $\underline{L} := \partial_t + \partial_r$ ; indeed we have  $\mathbf{T}_{LL} = |\psi_L|^2$ ,  $\mathbf{T}_{L\nabla} = \langle \psi_L, \psi_{\nabla} \rangle$ , and  $\mathbf{T}_{\underline{L}\underline{L}} = -\frac{1}{2} |\psi_{\nabla}|^2$ .

Applying flux decay (46) we thus have

$$\limsup_{t_1 \rightarrow 0^-} \int_{K_{[\lambda t_1, t_1]}} \frac{\eta}{\tilde{\rho}^3} |\tilde{x}^\alpha \psi_\alpha|^2 dx dt \leq CE_0$$

uniformly for all choices of  $\lambda$  and  $\varepsilon$ . We shall not take advantage of the ability to set  $\varepsilon$  to be small, and just take  $\varepsilon := 1$ . Then by the triangle inequality we have

$$|\psi_S| = |x^\alpha \psi_\alpha| \leq |\tilde{x}^\alpha \psi_\alpha| + |t_1| |\psi_0|.$$

On the other hand, using the crude bound  $\frac{1}{\tilde{\rho}^3} \leq \frac{C}{|t|^3}$  and (45) we see that

$$\int_{K_{[\lambda t_1, t_1]}} \frac{\eta}{\tilde{\rho}^3} |t_1|^2 |\psi_0|^2 dx dt \leq CE_0$$

uniformly in  $\lambda$  and  $t_1$ , and thus

$$\limsup_{t_1 \rightarrow 0^-} \int_{K_{[\lambda t_1/2, 2t_1]}} \frac{1}{|t|^3} |\psi_S|^2 dx dt \leq CE_0.$$

In particular, for any  $\lambda$  we see that for  $t_1$  sufficiently close to 0 (depending on  $\lambda$ ) we have

$$\int_{K_{[\lambda t_1/2, 2t_1]}} \left| \frac{1}{t} \psi_S \right|^2 dx \frac{dt}{t} \leq CE_0.$$

Telescoping this, we obtain that

$$\limsup_{T \rightarrow 0^-} \frac{1}{|\log |T||} \int_{K_{[-1, T]}} \left| \frac{1}{t} \psi_S \right|^2 dx \frac{dt}{t} \leq CE_0 / \log \lambda$$

(note that one can replace the interval  $[-1, T]$  by  $[t_0, T]$  for any fixed  $t_0$  without affecting the limit superior). Letting  $\lambda \rightarrow \infty$  we obtain (42).  $\blacksquare$

*Remark 3.2.* The estimate works for all targets  $N$ , as it did not use the hypothesis that  $N$  had constant negative curvature. It is also completely covariant and did not rely the fixing of a gauge.

*Remark 3.3.* Observe that in the above proof we have in fact obtained stronger estimates than (42); for instance, we can improve the decay of  $\frac{1}{t} \psi_S$  near the light cone, and obtain good estimates on relatively narrow slabs such as  $K_{[\lambda t, t]}$  instead of the tall slabs  $K_{[-1, T]}$  in (42). For instance, it is not difficult to use the above arguments to obtain a sequence of times  $t_k \rightarrow 0^-$ , a sequence of scales  $\lambda_k \rightarrow \infty$ , and  $\varepsilon_k \rightarrow 0^+$ , such that

$$\int_{K_{[\lambda_k t_k, t_k]}} \frac{|\psi_S|_N^2}{(\rho + \varepsilon_k |t|)^3} dx dt \leq \varepsilon_k.$$

Morally speaking, this suggests that the rescaled original wave maps  $\phi(t/t_k, x/t_k)$  are becoming locally self-similar as  $k \rightarrow \infty$ . Unfortunately this is not by itself enough to deduce that these rescaled wave maps converge (weakly) to a *non-trivial* self-similar wave map, even after other rescalings, because the above estimates do not prevent

the energy from dispersing into multiple points of concentration, or disappearing into the light cone; the basic problem is that control of the  $\psi_S$  component of the energy density does not seem to directly control the other components without further structural control on  $\phi$ , even near the light cone (the presence of angular components in the stress energy seems to prevent the Gronwall inequality approach in [7], [50], etc. from being effective, and naive attempts to exploit the negative curvature of the target  $N$  seem to require more boundedness control on  $\phi$  than is currently available.)

## References

- [1] B.K. Berger, P. T. Chrusciel, V. Moncrief, *On “asymptotically flat” spacetimes with  $G_2$  invariant Cauchy surfaces*, Ann. Physics **237** (1995), 322–354.
- [2] P. Bizon, T. Chmaj, Z. Tabor, *Formation of singularities for equivariant 2+1 dimensional wave maps into two-sphere*, Nonlinearity **14** (2001), no. 5, 1041–1053.
- [3] J. Bourgain, *Global well-posedness of defocusing 3D critical NLS in the radial case*, JAMS **12** (1999), 145–171.
- [4] J. Bourgain, *New global well-posedness results for non-linear Schrödinger equations*, AMS Publications, 1999.
- [5] T. Cazenave, J. Shatah, A.S. 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), 315–349.
- [6] Y. Choquet-Bruhat, *Global existence theorems for hyperbolic harmonic maps*, Ann. Inst. H. Poincaré Phys. Théor. **46** (1987), 97–111.
- [7] D. Christodoulou, A. Tahvildar-Zadeh, *On the regularity of spherically symmetric wave maps*, Comm. Pure Appl. Math, **46** (1993), 1041–1091.
- [8] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Global well-posedness and scattering in the energy space for the critical nonlinear Schrödinger equation in  $\mathbf{R}^3$* , preprint.
- [9] P. D’Ancona, V. Georgiev, *On the continuity of the solution operator of the wave maps system*, preprint.
- [10] J. Eells, H. Sampson, *Harmonic mappings of Riemannian manifolds*, Amer. J. Math. **86** (1964), 109–160.
- [11] A. Freire, S. Müller, M. Struwe, *Weak compactness of wave maps and harmonic maps*, Ann. Inst. H. Poincaré Anal. Non Linéaire **15** (1998), no. 6, 725–754.
- [12] J. Ginibre, G. Velo, *The Cauchy problem for the  $O(N)$ ,  $CP(N - 1)$ , and  $GC(N, P)$  models*, Ann. Physics, **142** (1982), 393–415.

- [13] L. Grafakos, R. Torres, *Multilinear Calderón-Zygmund theory*, Adv. Math. **165** (1992), 124–164.
- [14] M. Grillakis, *On nonlinear Schrödinger equations.*, Comm. Partial Differential Equations **25** (2000), no. 9-10, 1827–1844.
- [15] M. Grillakis, *Classical solutions for the equivariant wave map in 1 + 2 dimensions*, to appear in Indiana Univ. Math. J.
- [16] C. Gu, *On the Cauchy problem for harmonic maps defined on two-dimensional Minkowski space*, Comm. Pure Appl. Math., **33**,(1980), 727–737.
- [17] F. Helein, *Regularite des applications faiblement harmoniques entre une surface et une variete Riemannienne*, C.R. Acad. Sci. Paris Ser. I Math., **312** (1991), 591-596.
- [18] J. Isenberg, S. Lieblich, *Singularity formation in 2 + 1 wave maps*, J. Math. Phys. **43** (2002), no. 1, 678–683.
- [19] M. Keel, *Global existence for critical power Yang-Mills-Higgs equations in  $\mathbf{R}^{3+1}$* , Commun. in PDE **22** (1997), 1167–1227.
- [20] M. Keel, T. Tao, *Endpoint Strichartz Estimates*, Amer. Math. J. 120 (1998), 955–980.
- [21] M. Keel, T. Tao, *Local and global well-posedness of wave maps on  $\mathbf{R}^{1+1}$  for rough data*, IMRN **21** (1998), 1117–1156.
- [22] C. Kenig, E. Stein, *Multilinear estimates and fractional interpolation*, Math. Res. Lett. 6 (1999), 1–15
- [23] S. Klainerman, *On the regularity of classical field theories in Minkowski space-time  $\mathbf{R}^{3+1}$* , Prog. in Nonlin. Diff. Eq. and their Applic., **29**, (1997), Birkhäuser, 113–150.
- [24] S. Klainerman, *PDE as a unified subject*, preprint.
- [25] S. Klainerman, M. Machedon, *Smoothing estimates for null forms and applications*, Duke Math. J., **81** (1995), 99–133.
- [26] S. Klainerman, M. Machedon, *On the optimal local regularity for gauge field theories*, Diff. and Integral Eq. **10** (1997), 1019–1030.
- [27] S. Klainerman, I. Rodnianski, *On the global regularity of wave maps in the critical Sobolev norm*, IMRN **13** (2001), 656–677.
- [28] S. Klainerman, I. Rodnianski, *Sharp trace theorems for null hypersurfaces on Einstein metrics with finite curvature flux*, preprint.
- [29] S. Klainerman, S. Selberg, *Remark on the optimal regularity for equations of wave maps type*, C.P.D.E., **22** (1997), 901–918.

- [30] S. Klainerman, S. Selberg, *Bilinear estimates and applications to nonlinear wave equations*, preprint.
- [31] J. Krieger, *Global regularity of wave maps from  $\mathbf{R}^{3+1}$  to  $\mathbf{H}^2$* , CMP **238** (2003), 333–366.
- [32] J. Krieger, *Global regularity of wave maps from  $\mathbf{R}^{2+1}$  to  $\mathbf{H}^2$ . Small energy*, preprint.
- [33] O.A. Ladyzhenskaya, V.I. Shubov, *Unique solvability of the Cauchy problem for the equations of the two dimensional chiral fields, taking values in complete Riemann manifolds*, J. Soviet Math., **25** (1984), 855–864. (English Trans. of 1981 Article.)
- [34] P. Li, L. Tam, *The heat equation and harmonic maps of complete manifolds*, Invent. Math. **105** (1991), no. 1, 1–46.
- [35] G. Liao, L. Tam, *On the heat equation for harmonic maps from noncompact manifolds*, Pacific J. Math. **153** (1992), no. 1, 129–145.
- [36] S. Müller, M. Struwe, *Global existence of wave maps in  $1 + 2$  dimensions with finite energy data*, Topol. Methods Nonlinear Anal. **7** (1996), 245–259.
- [37] C. Muscalu, T. Tao, C. Thiele, *Multilinear operators given by singular multipliers*, J. Amer. Math. Soc. **15** (2002), 469–496.
- [38] A. Nahmod, A. Stefanov, K. Uhlenbeck, *On the well-posedness of the wave map problem in high dimensions*, Comm. Anal. Geom. **11** (2003), 49–83.
- [39] K. Nakanishi, *Local well-posedness and Illposedness in the critical Besov spaces for semilinear wave equations with quadratic forms*, Funk. Ekvac. **42** (1999), 261–279.
- [40] K. Nakanishi, *Energy scattering for non-linear Klein-Gordon and Schrodinger equations in spatial dimensions 1 and 2*, JFA **169** (1999), 201–225.
- [41] D. Oberlin, E. Stein, *Mapping properties of the Radon transform*, Indiana U. Math. J. **31** (1982), 641–650.
- [42] K. Pohlmeyer, *Integrable Hamiltonian systems and interaction through quadratic constraints*, Comm. Math. Phys., **46** (1976), 207–221.
- [43] R. Schoen, *Analytic aspects of the harmonic map problem*, Math. Sci. Res. Inst. Publ. **2** (1984), Springer, Berlin, 321–358.
- [44] R. Schoen, S.T. Yau, *Harmonic maps and the topology of stable hypersurfaces and manifolds with non-negative Ricci curvature*, Comment. Math. Helv. **51** (1976), no. 3, 333–341.
- [45] S. Selberg, *Multilinear space-time estimates and applications to local existence theory for non-linear wave equations*, Princeton University Thesis.

- [46] J. Shatah, *Weak solutions and development of singularities of the  $SU(2)$   $\sigma$ -model*. Comm. Pure Appl. Math., **41** (1988), 459–469.
- [47] J. Shatah, *The Cauchy problem for harmonic maps on Minkowski space*, in Proceed. Inter. Congress of Math. 1994, Birkhäuser, 1126–1132.
- [48] J. Shatah, W. Strauss, *Breathers as homoclinic geometric wave maps*, Physics D **99** (1996), 113–133.
- [49] J. Shatah, M. Struwe, *Regularity results for non-linear wave equations*, Ann. of Math. **138** (1993) 503–518.
- [50] J. Shatah, M. Struwe, *Geometric Wave Equations*, Courant Lecture Notes in Mathematics **2** (1998)
- [51] J. Shatah, M. Struwe, *The Cauchy problem for wave maps*, IMRN **11** (2002) 555–571.
- [52] J. Shatah, A. Tavildar-Zadeh, *Regularity of harmonic maps from the Minkowski space into rotationally symmetric manifolds.*, Comm. Pure Appl. Math. **45** (1992), 947–971.
- [53] J. Shatah, A. Tavildar-Zadeh, *On the Cauchy problem for equivariant wave maps*, Comm. Pure Appl. Math., **47** (1994), 719 – 753.
- [54] J. Shatah, A. Tavildar-Zadeh, *On the stability of stationary wave maps*, Comm. Math. Phys., **185** (1996), 231 – 256.
- [55] T. Sideris, *Global existence of harmonic maps in Minkowski space*, Comm. Pure Appl. Math., **42** (1989),1–13.
- [56] C. D. Sogge, *Lectures on Nonlinear Wave Equations*, Monographs in Analysis II, International Press, 1995.
- [57] E. M. Stein, *Harmonic Analysis*, Princeton University Press, 1993.
- [58] M. Struwe, *On the evolution of harmonic mappings of Riemannian surfaces*, Comment. Math. Helv. **60** (1985), no. 4, 558–581.
- [59] M. Struwe, *Wave Maps, in Nonlinear Partial Differential Equations in Geometry and Physics*, Prog. in Nonlin. Diff. Eq. and their Applic., **29**, (1997), Birkhäuser, 113–150.
- [60] M. Struwe, *Radially symmetric wave maps from the  $(1+2)$ -dimensional Minkowski space to a sphere*, Math Z. **242** (2002), 407–414.
- [61] M. Struwe, *Radially symmetric wave maps from  $(1+2)$ -dimensional Minkowski space to general targets*, Calc. Var. **16** (2003), 431–437.
- [62] M. Struwe, *Equivariant wave maps in two dimensions*, preprint.



- [63] T. Tao, *Ill-posedness for one-dimensional wave maps at the critical regularity*, Amer. J. Math. **122** (2000), 451–463.
- [64] T. Tao, *Global regularity of wave maps I. Small critical Sobolev norm in high dimension*, IMRN **7** (2001), 299–328.
- [65] T. Tao, *Global regularity of wave maps II. Small energy in two dimensions*, submitted, Comm. Math. Phys.
- [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. J. Math. **123** (2001), no. 1, 37–77.
- [68] D. Tataru, *Rough solutions for the wave maps equation*, preprint.
- [69] D. Tataru, *The wave maps equation*, preprint.

DEPARTMENT OF MATHEMATICS, UCLA, LOS ANGELES CA 90095-1555

tao@math.ucla.edu