Article:

Speight, JM (2015) The adiabatic limit of wave-map flow on the two-torus. Transactions of the American Mathematical Society, 367 (12). pp. 8997-9026. ISSN 0002-9947

https://doi.org/10.1090/tran/6538

The adiabatic limit of wave map flow on a two torus

J.M. Speight*
School of Mathematics, University of Leeds
Leeds LS2 9JT, England

Abstract

The S^2 valued wave map flow on a Lorentzian domain $\mathbb{R} \times \Sigma$, where Σ is any flat two-torus, is studied. The Cauchy problem with initial data tangent to the moduli space of holomorphic maps $\Sigma \to S^2$ is considered, in the limit of small initial velocity. It is proved that wave maps, in this limit, converge in a precise sense to geodesics in the moduli space of holomorphic maps, with respect to the L^2 metric. This establishes, in a rigorous setting, a long-standing informal conjecture of Ward.

1 Introduction

Wave maps are the analogue of harmonic maps in the case where the domain is Lorentzian. They satisfy a semilinear wave equation which has been heavily studied, in the simplest non-trivial case of S^2 target space, as a model PDE system involving a manifold-valued field [21]. The wave map equation is particularly interesting in the case where the domain is $(\mathbb{R} \times \Sigma, dt^2 - g_\Sigma)$, with (Σ, g_Σ) an oriented Riemannian two-manifold. In this case, the static wave map problem is conformally invariant, so static solutions on $\Sigma = \mathbb{R}^2$ in particular have no preferred scale: they can be dilated to any size without changing their energy. This suggests that time-dependent solutions with initial data close to a static solution might collapse and form singularities in finite time, an issue which has been heavily studied both numerically and analytically mainly for $\Sigma = \mathbb{R}^2$, $N = S^2$, within a certain rotational equivariance class. Numerical studies of increasing sophistication suggested that finite-time collapse can occur, and suggested formal models of the collapse process [11, 1, 13, 17]. The first rigorous proof of blow-up came in the work of Krieger, Schlag and Tataru [8], who proved the existence of rotationally equivariant initial data, of topological degree $n = 1$, leading to finite-time collapse. Rodnianski and Sterbenz subsequently proved existence of equivariant initial data of every degree $n \geq 4$ leading to finite-time collapse, and proved that collapse is stable to small perturbations of the initial data, at least within the equivariance class [19]. These results were extended to every degree $n \geq 1$ in work of Raphael and Rodnianski [18], which also established detailed asymptotics and universality properties of the collapse mechanism. For a thorough discussion of blow-up of wave maps, see [17, 18].

*E-mail: speight@maths.leeds.ac.uk

This paper addresses a different analytic issue from singularity formation, namely the validity of the *geodesic approximation* to wave map flow. To motivate this, one should think of wave maps $\mathbb{R} \times \Sigma \to S^2$ as (formal) critical points of the action functional

$$S = \int dt\,(T - E), \qquad \text{where} \quad T = \frac{1}{2}\int_\Sigma |\phi_t|^2, \quad E = \frac{1}{2}\int_\Sigma |d_\Sigma \phi|^2. \tag{1.1}$$

It follows (from Noether's Theorem) that they conserve the total energy $T + E$. A rather general argument of Lichnerowicz [12] shows that for any map $\phi : \Sigma \to S^2$ of topological degree $n \in \mathbb{Z}$ (subject to suitable boundary conditions, if Σ is noncompact), $E \geq 4\pi|n|$, with equality if and only is ϕ is $\pm$ holomorphic. So holomorphic maps, if they exist, minimize potential energy in their homotopy class. Let us denote by M_n the moduli space of degree n holomorphic maps $\Sigma \to S^2$. Consider a wave map $\phi(t)$ with $\phi(0) \in \mathsf{M}_n$ and $\phi_t(0) \in T_{\phi(0)}\mathsf{M}_n$ with $\|\phi_t(0)\|_{L^2}$ small. By conservation of $E + \frac{1}{2}\|\phi_t(t)\|^2_{L^2}$, one expects that $\phi(t)$ will stay close to M_n, on which E attains its minimum value, for as long as the solution persists. This led Ward to suggest [29], in the specific case $\Sigma = \mathbb{R}^2$, that such wave maps should be well approximated by the dynamical system with action S, but with $\phi(t)$ *constrained* to M_n for all time. Since E is constant on M_n, this constrained system is equivalent to geodesic motion on M_n with respect to the L^2 metric (obtained by restricting the quadratic form T to $T\mathsf{M}_n$). A similar approximation had previously been proposed by Manton [15] for low energy monopole dynamics, and the geodesic approximation is now a standard technique in the study of the dynamics of topological solitons [16].

Geodesic motion on M_2 (for $\Sigma = \mathbb{R}^2$) was studied in detail in [10]. There is a technical problem: the L^2 metric is only well-defined on the leaves of a foliation of M_n and one must impose by hand that $\phi(t)$ remains on a single leaf. This turns out to be ill-justified (it precludes singularity formation for $n = 1$, for example, in contradiction of [8, 18]). This technical deficiency is removed if we choose Σ to be a compact Riemann surface. Here geodesic motion in M_n is globally well-defined, if incomplete [20], and the L^2 geometry of M_n is quite well understood, at least for some choices of Σ and n [14, 22, 23, 24].

The question remains: is geodesic motion in M_n really a good approximation to wave map flow in the adiabatic (low velocity) limit? The purpose of this paper is to prove that it *is*, for times of order (initial velocity)$^{-1}$ at least in the case where Σ is any flat two-torus. More precisely, we will prove:

Theorem 1.1 (Main Theorem)**.** *Let M_n denote the moduli space of degree $n \geq 2$ holomorphic maps from a flat two-torus Σ to S^2. For fixed $\phi_0 \in \mathsf{M}_n$ and $\phi_1 \in T_{\phi_0}\mathsf{M}_n$ consider the one parameter family of initial value problems for the wave map equation with $\phi(0) = \phi_0$, $\phi_t(0) = \varepsilon\phi_1$, parametrized by $\varepsilon > 0$. There exist constants $\tau_* > 0$ and $\varepsilon_* > 0$, depending only on the initial data, such that for all $\varepsilon \in (0, \varepsilon_*]$, the problem has a unique solution for $t \in [0, \tau_*/\varepsilon]$. Furthermore, the time re-scaled solution*

$$\phi_\varepsilon : [0, \tau_*] \times \Sigma \to S^2, \qquad \phi_\varepsilon(\tau, p) = \phi(\tau/\varepsilon, p)$$

converges uniformly in C^1 to $\psi : [0, \tau_*] \times \Sigma \to S^2$, the geodesic in M_n with the same initial data, as $\varepsilon \to 0$.

To prove this we will adapt the perturbation method devised by Stuart to prove validity of the geodesic approximation in the critically coupled abelian Higgs and Yang-Mills-Higgs models [25, 26]. The wave map problem has a key similarity with these gauge-theoretic problems, namely a moduli space of static solutions which minimize energy in their homotopy class and satisfy a system of first order "Bogomol'nyi" equations. (For wave maps, the Bogomol'nyi equation is the condition that ϕ be $\pm$ holomorphic, i.e. the Cauchy-Riemann equation.) Roughly, the idea is to decompose the solution $\phi(t)$ as $\phi(t) = \psi(t) + \varepsilon^2 Y(t)$ where $\psi(t) \in \mathsf{M}_n$, and control the growth of a suitable Sobolev norm of the error $Y(t)$ uniformly in ε by means of energy estimates. One concurrently shows that the projected trajectory $\psi(t)$ converges to a geodesic in M_n.

In comparison with Stuart's work on vortices and monopoles, the situation we study is simpler in two respects: we work on a compact domain Σ (rather than $\mathbb{R}^2$ or $\mathbb{R}^3$), and our system has no gauge symmetry. On the other hand, the wave map problem introduces two new challenges for the method.

First, our field is manifold-valued, so it is not clear a priori what the decomposition $\phi(t) = \psi(t) + \varepsilon^2 Y(t)$ really means. In preliminary work on this problem, it was suggested that the correct formulation was $\phi(t) = \exp_{\psi(t)} \varepsilon^2 Y(t)$, where $\exp : TS^2 \to S^2$ is the exponential map [5]. In fact, this turns out *not* to have the analytic properties required by Stuart's method (except for rotationally equivariant wave maps). In this paper we isometrically embed S^2 in $\mathbb{R}^3$ and use the ambient linear structure to project as usual, $\phi(t) = \psi(t) + \varepsilon^2 Y(t)$. This choice is simple, but has significant repercussions: Y is no longer tangent to the map ψ (not a section of $\psi^{-1}TS^2$), and must satisfy a nonlinear pointwise constraint to ensure that ϕ is S^2 valued. The evolution of Y is governed by a nonlinear wave equation whose (spatial) linear part is the Jacobi operator J_ψ for the harmonic map $\psi : \Sigma \to S^2$. It turns out that J_ψ is *not* self-adjoint when acting on non-tangent sections (such as Y). Since self-adjointness of (the analogue of) J_ψ is crucial for Stuart's method, we must devise a way round this: we replace J_ψ by an "improved" Jacobi operator L_ψ, which coincides with J_ψ on tangent sections, but is self-adjoint on all sections, and introduce compensating nonlinear terms into the wave equation for Y using the pointwise constraint. Further difficulties result: L_ψ, unlike J_ψ, does not define a coercive quadratic form on the L^2 orthogonal complement of $T_\psi \mathsf{M}_n$. We must work instead with a weaker near-coercivity property, which turns out to suffice for our purposes.

Second, while Σ is compact, the moduli space M_n is not. Of course, the vortex and monopole moduli spaces, dealt with by Stuart, are also noncompact, but in those cases, moving to infinity corresponds to (clusters of) solitons separating off and escaping to infinite separation, a well-controlled process. For wave maps, by contrast, approaching the boundary of M_n at infinity corresponds to one or more lumps collapsing and "bubbling off". In this process, ψ becomes singular and both geodesic motion and wave map flow become badly behaved. To handle this, we must keep careful track of the position (of $\psi \in \mathsf{M}_n$) dependence of our various estimates, and modify Stuart's a priori energy bound so that we simultaneously control the error $Y(t)$, the deviation of $\psi(t)$ from the corresponding geodesic, and the distance of $\psi(t)$ from $\partial_\infty \mathsf{M}_n$.

It is interesting to speculate to what extent Theorem 1.1 can be generalized. It is clear that the proof presented here generalizes quite easily to the case of a general compact Riemann surface, provided n is sufficiently large compared with the genus of Σ. The reason for restricting to the case $\Sigma = T^2$ is mainly one of presentation: the existence of global cartesian coordinates makes it straightforward to define the various function spaces, for example. Generalizing the target space is not so straightforward. The wave map flow $\mathbb{R} \times \Sigma \to N$ has the appropriate "Bogomol'nyi" form for Stuart's method to apply whenever Σ, N are both compact kähler manifolds (in fact, it suffices for Σ to be co-kähler). The choice $N = \mathbb{C}P^k$, $k \geq 2$, is of some interest in mathematical physics, for example. But here the reliance on an isometric embedding $N \subset \mathbb{R}^p$ becomes very problematic. It seems likely that some variant of Theorem 1.1 does remain true for general compact kähler targets, but proving it would require a rather different approach, perhaps along the lines sketched in [27].

One should note that Theorem 1.1 gives no information about singularity formation for wave maps on $\mathbb{R} \times \Sigma$ because, although there certainly are geodesics $\psi(\tau)$ which hit $\partial_\infty \mathsf{M}_n$ in time $\tau_0 < \infty$, and the corresponding wave maps do converge uniformly to $\psi(\tau)$ on some interval $[0, \tau_*]$, there is no reason to expect $\tau_* = \tau_0$. In fact Raphael and Rodnianski have shown that singularity formation of equivariant wave maps on $\mathbb{R}^2$ deviates significantly from the dynamics predicted by the (suitably regulated) geodesic approximation [18]. Since blow up is a (spatially) local phenomenon, these results presumably apply in some form on the torus, which would imply $\tau_* < \tau_0$. Nontheless, the geodesic approximation (on compact Σ or, regulated, on $\Sigma = \mathbb{R}^2$) predicted finite time blow-up of wave maps in $(2+1)$ dimensions, and this prediction turned out to be correct. The geodesic approximation also makes predictions about the *genericity* of blow up. It is not hard to prove, for example, that generic geodesics on M_1 for $\Sigma = S^2$ do *not* hit the boundary at infinity. It would be interesting to see whether the full wave map flow has this property (i.e. generic Cauchy data tangent to M_1 have global smooth solutions). The analogue of Theorem 1.1 for $\Sigma = S^2$ could provide a starting point for proving such results.

The rest of the paper is structured as follows. In section 2 the moduli space M_n of holomorphic maps is introduced and its key property, Proposition 2.1 (existence of a smooth local parametrization about any point), established. In section 3, the projection of the wave map flow to M_n is defined, and the coupled system satisfied by ψ and the error section Y is derived, equation (3.15). In section 4, some standard functional analytic definitions and results are introduced. Our aim here, and in the remainder of the paper, is to make the proof accessible to a wide mathematical physics audience, not just experts in PDE. In section 5 a local existence and uniqueness theorem for the coupled system (3.15) governing (ψ, Y) is proved, Theorem 5.1. Of course, local existence and uniqueness of *wave maps* in this context is not new; the extra, and new, information we obtain here is local existence and uniqueness of the *projection* to M_n. This is the engine underlying Stuart's method, and we go through the argument in some detail, not only because there are certain new aspects we have to deal with which Stuart did not (e.g. preservation of the pointwise constraint on Y), but also because the requirements of Picard's method for this proof determine our choice of function spaces, a point which is not obvious (to the non-analyst) in Stuart's original applications of the method [25, 26]. In section 6 the key near-coercivity property of the quadratic form associated with the improved Jacobi operator is proved, Theorem 6.5 (roughly, that $\langle L_\psi Y, L_\psi L_\psi Y \rangle_{L^2}$ controls the H^3 norm of Y). In section 7 energy estimates are established which bound the growth of $Y(t)$. Finally, in section 8, the coercivity properties and energy estimates are combined to prove long time existence of the solution (ψ, Y), and establish convergence to the corresponding geodesic in

M_n. An appendix presents the proofs of some basic analytic properties of the nonlinear terms in the coupled system for (ψ, Y) which are used repeatedly in section 5.

2 The moduli space of static n-lumps

Static wave maps are harmonic maps $\phi : \Sigma \to S^2$, that is, solutions of the harmonic map equation

$$\phi_{xx} + \phi_{yy} + (|\phi_x|^2 + |\phi_y|^2)\phi = 0 \tag{2.1}$$

or, equivalently, critical points of the Dirichlet energy

$$E(\phi) = \frac{1}{2}\int_\Sigma |\phi_x|^2 + |\phi_y|^2. \tag{2.2}$$

Maps $\Sigma \to S^2$ fall into disjoint homotopy classes labelled by their degree $n \in \mathbb{Z}$, which we may assume, without loss of generality, is non-negative. An argument of Lichnerowicz [3, p39] shows that, in the degree n class, $E(\phi) \geq 4\pi n$, with equality if and only if ϕ is holomorphic. Furthermore, all harmonic maps $\Sigma \to S^2$ of degree $n \geq 2$ are holomorphic [4]. So the moduli space of interest, M_n, is the space of degree n holomorphic maps $\Sigma \to S^2$. Such maps are called "n-lumps" by analogy with the case $\Sigma = \mathbb{C}$, where the Dirichlet energy density typically exhibits n distinct local maxima, which may loosely be thought of as smoothed out particles, or lumps of energy.

The global topology of the space M_n is quite complicated, for example, $\mathsf{M}_2 \cong [\Sigma \times PSL(2,\mathbb{C})]/(\mathbb{Z}_2 \times \mathbb{Z}_2)$, [23]. For our purposes local information will suffice, however. Given a smooth variation ϕ_s of $\phi = \phi_0 \in \mathsf{M}_n$ we have $dE(\phi_s)/ds = 0$ at $s = 0$ (since ϕ is harmonic) and

$$\frac{d^2E(\phi_s)}{ds^2} = \langle V, J_\phi V\rangle = \int_\Sigma V \cdot J_\phi V. \tag{2.3}$$

where $V = \partial_s\phi_s|_{s=0} \in \Gamma(\phi^{-1}TS^2)$ is the infinitesimal generator of the variation and

$$J_\phi V = -V_{xx} - V_{yy} - (|\phi_x|^2 + |\phi_y|^2)V - 2(\phi_x \cdot V_x + \phi_y \cdot V_y)\phi \tag{2.4}$$

is the Jacobi operator at the map ϕ [28, p155]. This operator is self-adjoint and elliptic, and its spectrum determines the stability properties of ϕ. By the Lichnerowicz argument, ϕ minimizes E in its homotopy class, so spec J_ϕ is non-negative. Given a variation ϕ_s through *harmonic* maps, that is, a curve in M_n through $\phi = \phi_0$, its infinitesimal generator $V = \partial_s\phi_s|_{s=0}$ satisfies $J_\phi V = 0$. Hence $T_\phi\mathsf{M}_n \subset \ker J_\phi$. For a general harmonic map $\phi : M \to N$ between Riemannian manifolds, the converse may be false, that is, there may be sections $V \in \ker J_\phi \subset \Gamma(\phi^{-1}TN)$ which are not tangent to any variation of ϕ through harmonic maps, and in this case the space of harmonic maps $M \to N$ may not be a smooth manifold around ϕ. It is important for us to rule out this kind of bad behaviour in our case. More precisely, we need:

Proposition 2.1. *Given any $\phi_0 \in \mathsf{M}_n$, $n \geq 2$, there exists an open set $U \subset \mathbb{R}^{4n}$ and a smooth map $\psi : U \times \Sigma \to S^2$ such that,*

(i) for each $q \in U$, $\psi(q, \cdot) \in \mathsf{M}_n$,

(ii) there exists $q_0 \in U$ *such that* $\phi_0 = \psi(q_0, \cdot)$*, and*

(iii) $\psi_\mu = \partial\psi/\partial q^\mu$, $\mu = 1, 2, \ldots, 4n$ *span* $\ker J_{\psi(q,\cdot)}$.

Proof. Choose any $p \in S^2$ such that both p and $-p$ are regular values of ϕ_0 (such p exists by Sard's Theorem). Then $\phi : \Sigma \to S^2$ is in M_n if and only if $s_p \circ \phi$, its image under stereographic projection from p, is meromorphic, of degree n, that is, a degree n elliptic function. The most general degree n elliptic function is [9]

$$(s_p \circ \phi)(z) = \lambda \frac{\sigma(z - a_1) \cdots \sigma(z - a_n)}{\sigma(z - b_1) \cdots \sigma(z - b_n)} \tag{2.5}$$

where σ is the Weierstrass sigma function, $\lambda, a_1, \ldots, a_n, b_1, \ldots, b_n$ are complex constants, $\lambda \neq 0$, $\sum a_i = \sum b_i \mod \Lambda$ and $\{a_i\} \cap \{b_j\} = \emptyset$. Hence, we may parametrize a general point $\phi \in \mathsf{M}_n$ by $4n$ real numbers q^μ, for example, the real and imaginary parts of $\lambda, a_1, \ldots, b_{n-1}$ having set $b_n = a_1 + \cdots + a_n - b_1 - \cdots - b_{n-1}$. Further, ϕ manifestly depends smoothly on q and z. Hence we have a smooth map $\psi : \widetilde{U} \times \Sigma \to S^2$ satisfying properties (i) and (ii). By our choice of p, $s_p \circ \phi_0$ has n distinct zeroes and n distinct poles, so $\{\psi_\mu\}$ at $q = q_0$ are linearly independent sections of $\psi(q, \cdot)^{-1} TS^2$, and hence, by smoothness, also linearly independent on some neighbourhood U of q_0 in $\widetilde{U}$. As explained previously, ψ_μ span a subspace of $\ker J_{\psi(q,\cdot)}$, so it remains to show that $\ker J_\phi$ has dimension $4n$ for any $\phi \in \mathsf{M}_n$.

It is known [28, p174] that $\ker J_\phi$ is isomorphic, as a complex vector space, to $H^0(\Sigma, L)$, the space of holomorphic sections of the line bundle $L = \phi^{-1} T'S^2$, where $T'S^2$ denotes the holomorphic tangent bundle of S^2. Since ϕ has degree n and $T'S^2$ has degree 2, L has degree $2n$. Now, by the Riemann-Roch formula [6]

$$\dim H^0(\Sigma, L) - \dim H^1(\Sigma, L) = \deg L = 2n \tag{2.6}$$

since Σ has genus 1. But, by Serre duality, $H^1(\Sigma, L) \cong H^0(\Sigma, K \otimes L^*)^*$ where K is the canonical bundle of Σ. Now K is trivial, so $K \otimes L^*$ has degree $-2n$, and hence has no holomorphic sections, whence $H^1(\Sigma, L) = 0$. It follows that $\ker J_\phi$ has real dimension $4n$, as required. $\square$

We can regard q^μ as local coordinates on M_n. Given the initial data $\phi_1 \in T_{\phi_0}\mathsf{M}_n$ of interest, we choose and fix such a $\psi : U \times \Sigma \to S^2$ and denote by $q_0 \in U$ and $q_1 \in \mathbb{R}^{4n}$ those vectors such that $\phi_0 = \psi(q_0, \cdot)$ and $\phi_1 = q_1^\mu \psi_\mu(q_0, \cdot)$. Here, as henceforth, we use the Einstein summation convention on repeated indices. We also choose and fix a compact neighbourhood K of q_0 in U. In a slight abuse of notation, we will also use the symbol ψ to denote the associated map $U \supset K \to \mathsf{M}_n \subset (S^2)^\Sigma$, so $\psi(q)$ will denote the holomorphic map $z \mapsto \psi(q, z)$. We will also use $\psi(t)$ as shorthand for $\psi(q(t))$, meaning a general curve in M_n.

There is a natural Riemannian metric on M_n, the L^2 metric, whose components in the local coordinate system q^μ are

$$\gamma_{\mu\nu}(q) = \langle \psi_\mu, \psi_\nu \rangle = \int_\Sigma \frac{\partial \psi}{\partial q^\mu} \cdot \frac{\partial \psi}{\partial q^\nu}. \tag{2.7}$$

The associated Christoffel symbol is

$$G^\mu_{\lambda\nu}(q) = \gamma^{\mu\alpha} \langle \psi_\alpha, \psi_{\lambda\nu} \rangle \tag{2.8}$$

where $\gamma^{\mu\nu}$ is the inverse metric and $\psi_{\mu\nu} = \partial^2\psi/\partial q^\mu \partial q^\nu$. This is the metric whose geodesics approximate wave maps in the adiabatic limit.

3 Projection of wave map flow and the coupled system

The wave map equation for $\phi : \mathbb{R} \times \Sigma \to S^2 \subset \mathbb{R}^3$ is

$$\phi_{tt} - \phi_{xx} - \phi_{yy} + (|\phi_t|^2 - |\phi_x|^2 - |\phi_y|^2)\phi = 0. \tag{3.1}$$

Given $\varepsilon > 0$, a small parameter, we decompose ϕ as

$$\phi = \psi + \varepsilon^2 Y \tag{3.2}$$

where, at each fixed time, $\psi(t, \cdot) : \Sigma \to S^2$ is a degree n harmonic map, and $Y : \Sigma \to \mathbb{R}^3$. We may think of $\psi(t)$ as a curve in M_n, the moduli space of degree n harmonic maps, and Y as the "error" incurred by projecting $\phi(t)$ to $\psi(t)$. It is useful to think of Y as a section of $\psi^{-1}\underline{\mathbb{R}^3}$, where $\underline{\mathbb{R}^3} = S^2 \times \mathbb{R}^3$ is the trivial $\mathbb{R}^3$ bundle over S^2, and $\psi^{-1}\underline{\mathbb{R}^3}$ is its pullback to Σ. With this is mind, we refer to Y as the "error section", and to any $Z : \Sigma \to \mathbb{R}^3$ with $Z \cdot \psi = 0$ pointwise as a "tangent section" (in bundle language, Z is a section of $\psi^{-1}TS^2 \subset \psi^{-1}\underline{\mathbb{R}^3}$). Clearly, Y is *not* a tangent section (unless $Y = 0$). Since both ψ and ϕ are S^2 valued, Y must satisfy the pointwise constraint

$$\psi \cdot Y = -\frac{1}{2}\varepsilon^2 |Y|^2. \tag{3.3}$$

For a given curve $\psi(t)$, if ϕ is a wave map then Y must satisfy the PDE obtained by substituting (3.2) into (3.1),

$$Y_{tt} + J_\psi Y = k + \varepsilon j \tag{3.4}$$

where J_ψ is the Jacobi operator associated with the harmonic map $\psi(t) : \Sigma \to S^2$ and the terms on the right hand side are

$$\begin{aligned} k &= -(\psi_{\tau\tau} + |\psi_\tau|^2\psi) \\ j &= 2(\psi_\tau \cdot Y_t)\psi + \varepsilon\{(|Y_t|^2 - |Y_x|^2 - |Y_y|^2)\psi + (|\psi_\tau|^2 - 2\psi_x \cdot Y_x - 2\psi_y \cdot Y_y)Y\} \\ &\quad +2\varepsilon^2(\psi_\tau \cdot Y_t)Y + \varepsilon^3(|Y_t|^2 - |Y_x|^2 - |Y_y|^2)Y. \end{aligned} \tag{3.5}$$

We have here introduced the slow time variable $\tau = \varepsilon t$ as a book-keeping device. The precise expression for j is not important. What matters is its qualitative form: it depends only on ψ, ψ_τ and Y and its first derivatives, and the dependence is smooth (polynomial, in fact).

Superficially (3.4) looks exactly analogous to the corresponding equation in Stuart's analysis of vortex dynamics [25], but this is deceptive. As already noted, the Jacobi operator is a self-adjoint, elliptic, second order linear operator $J_\psi : \Gamma(\psi^{-1}TS^2) \to \Gamma(\psi^{-1}TS^2)$ whose spectrum determines the stability properties of the harmonic map ψ [28]. It is important to realize, however, that in (3.4) J_ψ is acting on Y, which is *not* a tangent section. So in (3.4), J_ψ is precisely the same operator defined above (2.4), but extended to act on sections of $\psi^{-1}\underline{\mathbb{R}^3}$. But J_ψ is *not* self-adjoint (with respect to L^2) as an operator on $\psi^{-1}\underline{\mathbb{R}^3}$, and self-adjointness of (the analogue of) J_ψ is a crucial ingredient in Stuart's method. In fact

$$J_\psi Z = -\Delta Z - (|\psi_x|^2 + |\psi_y|^2)Z + A_\psi Z \tag{3.6}$$

where the non-self-adjoint piece and its adjoint are

$$\begin{aligned} A_\psi Z &= -2(\psi_x \cdot Z_x + \psi_y \cdot Z_y)\psi \\ A_\psi^\dagger Z &= -2\left\{(\psi\cdot Z)\Delta\psi + (\psi\cdot Z)_x\psi_x + (\psi\cdot Z)_y\psi_y\right\}. \end{aligned} \tag{3.7}$$

and we have adopted the analysts' convention for the Laplacian, that is, $\Delta Z = Z_{xx} + Z_{yy}$. To remedy this deficiency, we make the following definition:

Definition 3.1. Given a harmonic map $\psi : \Sigma \to S^2$, we define its *improved Jacobi operator* to be

$$L_\psi : \Gamma(\psi^{-1}\underline{\mathbb{R}^3}) \to \Gamma(\psi^{-1}\underline{\mathbb{R}^3}), \qquad L_\psi = J_\psi + A_\psi^\dagger.$$

Note that L_ψ coincides with J_ψ on $\Gamma(\psi^{-1}TS^2)$, and hence L_ψ maps tangent sections to tangent sections. Its principal part is the Laplacian, so it is elliptic, and it is manifestly self adjoint.

Remark 3.2. Any section can be decomposed into tangent and normal components. As just observed, L_ψ maps a tangent section Z to the tangent section $J_\psi Z$, so an alternative way of characterizing L_ψ is by specifying how it acts on normal sections, $\alpha\psi$ where $\alpha : \Sigma \to \mathbb{R}$. A short calculation, using harmonicity of ψ, yields

$$L_\psi(\alpha\psi) = -(\Delta\alpha)\psi - 4(\alpha_x\psi_x + \alpha_y\psi_y). \tag{3.8}$$

It follows immediately that $\ker L_\psi = \ker J_\psi \oplus \langle\psi\rangle$. Note that, in general, L_ψ does not map normal sections to normal sections.

Now, for any Y satisfying the pointwise constraint,

$$A_\psi^\dagger Y = \varepsilon^2\{|Y|^2\Delta\psi + 2(Y\cdot Y_x)\psi_x + 2(Y\cdot Y_y)\psi_y\} =: \widehat{j} \tag{3.9}$$

and so, for any curve $\psi(t)$, if ϕ is a wave map then Y satisfies the PDE

$$Y_{tt} + L_\psi Y = k + \varepsilon j' \tag{3.10}$$

where $j' = j + \widehat{j}$. Note that j' has the same qualitative analytic properties as j (specifically, no higher than first derivatives of Y appear).

We have yet to specify the curve $\psi(t)$ in M_n. We do this by demanding that the error section Y should at all times be L^2 orthogonal to $T_\psi\mathsf{M}_n$. In this case ($\Sigma = T^2$, $n \geq 2$), $T_\psi\mathsf{M}_n = \ker J_\psi$ so, in terms of the local coordinate system q on M_n provided by Proposition 2.1, this amounts to requiring

$$\langle Y, \psi_\mu\rangle = 0, \qquad \mu = 1, 2, \ldots, 4n. \tag{3.11}$$

We convert this into an evolution equation for q by differentiating the orthogonality constraint (3.11) twice with respect to time and using (3.10),

$$\langle -L_\psi Y + k + \varepsilon j', \psi_\mu\rangle + 2\varepsilon Y_t, \psi_{\mu\nu}\dot{q}^\nu + \varepsilon^2\langle Y, \psi_{\mu\nu\lambda}\rangle\dot{q}^\nu\dot{q}^\lambda + \varepsilon^2\langle Y, \psi_{\mu\nu}\rangle\ddot{q}^\nu = 0, \tag{3.12}$$

where an overdot denotes differentiation with respect to τ. Now L_ψ is self adjoint and $\psi_\mu \in \ker J_\psi \subset \ker L_\psi$, so this equation simplifies to $\langle k, \psi_\mu \rangle = O(\varepsilon)$, or, more explicitly,

$$\ddot{q}^\mu + G^\mu_{\nu\lambda}(q)\dot{q}^\nu \dot{q}^\lambda = \varepsilon h^\mu(\varepsilon, q, \dot{q}, Y, Y_t) + \varepsilon^2 \gamma^{\mu\nu}\langle Y, \psi_{\nu\lambda}\rangle \ddot{q}^\lambda \tag{3.13}$$

where γ and G are the L^2 metric and its Christoffel symbol in the coordinate system q, (2.7), (2.8), and the function h is

$$\begin{aligned} h^\mu &= \gamma^{\mu\nu}\Big\{ \langle Y_t, \psi_{\nu\lambda}\rangle \dot{q}^\lambda + \varepsilon \langle Y, \psi_{\nu\lambda\rho}\rangle \dot{q}^\lambda \dot{q}^\rho + \varepsilon \langle (\psi_\lambda \cdot \psi_\rho) Y, \psi_\nu \rangle \dot{q}^\lambda \dot{q}^\rho \\ &\quad -2\varepsilon \langle (\psi_x \cdot Y_x + \psi_y \cdot Y_y) Y, \psi_\nu \rangle + 2\varepsilon^2 \langle (\psi_\lambda \cdot Y_t) Y, \psi_\nu \rangle \dot{q}^\lambda \\ &\quad + \varepsilon^3 \langle (|Y_t|^2 - |Y_x|^2 - |Y_y|^2) Y, \psi_\nu \rangle \Big\}. \end{aligned} \tag{3.14}$$

Taking the formal limit $\varepsilon \to 0$, (3.13) reduces to the geodesic equation on (M_n, γ), as one would hope.

To summarize, if ϕ is a wave map, and $q(t)$ is a curve in M_n such that $Y = \varepsilon^{-2}(\phi - \psi(q))$ satisfies the orthogonality constraint (3.11) at all times, then (Y, q) satisfies the coupled system

$$Y_{tt} + LY = k + \varepsilon j', \qquad \ddot{q}^\mu + G^\mu_{\nu\lambda}\dot{q}^\nu \dot{q}^\lambda = \varepsilon h^\mu + \varepsilon^2 \gamma^{\mu\nu}\langle Y, \psi_{\nu\lambda}\rangle \ddot{q}^\lambda \tag{3.15}$$

and the pointwise constraint (3.3). Conversely, if (Y, q) satisfies the constraints (3.3) and (3.11) and the coupled system (3.15), then $\phi = \psi(q) + \varepsilon^2 Y$ is a wave map. Our goal is to prove that (3.15) with fixed initial data $q(0) = q_0$, $\dot{q}(0) = q_1$, $Y(0) = 0$, $Y_t(0) = 0$ has solutions with $\|Y\|_{C^1}$ bounded uniformly in ε for times of order ε^{-1}. It follows immediately that, in the limit $\varepsilon \to 0$, $\phi(\tau/\varepsilon)$ converges uniformly to a curve $\psi(\tau)$ in M_n. In the course of the proof, we will simultaneously show that $\psi(\tau)$ is the geodesic with initial data q_0, q_1.

4 Analytic prelimaries

In this section we set up the function spaces we will use, and collect some standard functional analytic results which we will appeal to repeatedly. More details can be found in [2], and references therein. Let H^k denote the set of real-valued functions on Σ whose partial derivatives up to order k are square integrable. This is a Hilbert space with respect to the inner product

$$\langle f, g\rangle_k = \sum_{|\alpha|\le k} \int_\Sigma D_\alpha f \, D_\alpha g \tag{4.1}$$

where α is a multi-index taking values from $\{x, y\}$, $|\alpha|$ is its length and $D_\alpha = \partial_{\alpha_1}\partial_{\alpha_2}\cdots\partial_{\alpha_{|\alpha|}}$, so $D_{(x,x,y)} = \partial_x^2 \partial_y$, for example. We denote the corresponding norm by $\|\cdot\|_k$,

$$\|f\|_k^2 = \langle f, f\rangle_k. \tag{4.2}$$

Let $H^k = \mathsf{H}^k \oplus \mathsf{H}^k \oplus \mathsf{H}^k$, the space of $\mathbb{R}^3$-valued functions on Σ whose components are in H^k. This is a Hilbert space with respect to the inner product

$$\langle Y, Z\rangle_k = \langle Y_1, Z_1\rangle_k + \langle Y_2, Z_2\rangle_k + \langle Y_3, Z_3\rangle_k \tag{4.3}$$

whose norm will again be denoted $\| \cdot \|_k$. We adopt the convention that $\| \cdot \| = \| \cdot \|_0$ and $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_0$, that is, undecorated norms and inner products refer to L^2. We will frequently, and without further comment, use the Cauchy-Schwarz inequality

$$\langle Y, Z \rangle_k \leq \|Y\|_k \|Z\|_k \tag{4.4}$$

and the trivial bound $\|Y\|_k \geq \|Y\|_{k'}$ if $k \geq k'$.

In the sequel, we will prove existence of a solution of the coupled system (3.15) with $Y \in H^k$, $Y_t \in H^{k-1}$, for $k = 3$. This choice of k is motivated by the following fundamental fact about H^k on a compact 2-manifold.

Proposition 4.1 (Algebra property of H^k, $k \geq 2$)**.** *The Banach space $(\mathsf{H}^k, \| \cdot \|_k)$ is a Banach algebra for all $k \geq 2$. That is, if $f \in \mathsf{H}^k$ and $g \in \mathsf{H}^k$ then $fg \in \mathsf{H}^k$, and there exists a constant $\alpha_k > 0$, depending only on Σ and k, such that $\|fg\|_k \leq \alpha_k \|f\|_k \|g\|_k$.*

It follows directly from this that, if $(Y, Y_t) \in H^3 \oplus H^2$, then the nonlinear term j' in the coupled system is in H^2, and $\|j'\|_2$ can be bounded by a polynomial in $\|Y\|_3, \|Y_t\|_2$ (the point being that j' contains no derivatives of Y higher than first, and Y_x, Y_y, Y_t are all in H^2). This is crucial, not only for proving the local existence result for (3.15), but also in later sections where we prove that $\|Y\|_3$ is controlled by $\langle LY, LLY \rangle$, and make energy estimates for the solution. So the fact that we choose $k = 3$ is not just motivated by a desire to get strong bounds on the error section Y; the method will not work for any lower k. Indeed, to uniformly bound Y on Σ, it would suffice to control $\|Y\|_2$, as we have the following Sobolev inequality.

Proposition 4.2 (Sobolev inequalities)**.** *Let C^k denote the Banach space of continuous maps $\Sigma \to \mathbb{R}^3$ with the usual norm $\|Y\|_{C^k} = \sup\{|D_\alpha Y(p)| \, : \, |\alpha| \leq k, p \in \Sigma\}$. Then $H^2 \subset C^0$, $H^3 \subset C^1$, and there is a constant $\alpha > 0$, depending only on Σ, such that $\|Y\|_{C^k} \leq \alpha \|Y\|_{k+2}$ for all $Y \in H^2$, $k = 0, 1$. (More briefly, the inclusions $\iota : H^2 \to C^0$ and $\iota : H^3 \to C^1$ are continuous.)*

In later sections we will need to bound $\|L_\psi Y\|_k$ in terms of $\|Y\|_{k+2}$. Of course, since L_ψ is a linear second order operator we have trivially, for all $q \in K$, the upper bound

$$\|L_\psi Y\|_k \leq C \|Y\|_{k+2} \tag{4.5}$$

where C is a constant depending only on Σ and K. For a lower bound, we use the fact that L_ψ is elliptic.

Proposition 4.3 (Standard elliptic estimate)**.** *Let D be an elliptic linear differential operator of order r acting on sections of a vector bundle V over Σ. Then there exist constants α_k, β_k depending only on Σ and k, such that*

$$\|DY\|_k + \alpha_k \|Y\|_0 \geq \beta_k \|Y\|_{k+r}.$$

If we consider only sections which are L^2 orthogonal to $\ker D$, the same inequality holds with $\alpha_k = 0$.

The reason for quoting this result in the context of a general vector bundle V over Σ is that we will want to apply it to both the ordinary Laplacian on $V = \underline{\mathbb{R}^3}$, and the classical Jacobi operator J_ψ on $V = \psi^{-1} TS^2 \subset \underline{\mathbb{R}^3}$, where the H^k norm is defined by inclusion.

5 Local existence theorem

Theorem 5.1 (Local existence for the coupled system)**.** *Consider the coupled system (3.15) with initial data* $q(0) = q_0 \in K$, $q_t(0) = \varepsilon q_1 \in \mathbb{R}^{4n}$, $Y(0) = Y_0 \in H^3$, $Y_t(0) = Y_1 \in H^2$ *such that*

$$\text{dist}\,(q_0, \partial K) > d, \qquad |q_1|, \|Y_0\|_3, \|Y_1\|_2 < \Gamma$$

where Γ, d *are positive constants. Then there exist constants* $C(K) > 0$ *and* $T(K, \Gamma, d) > 0$ *such that for all* $\varepsilon \in (0, C(K)/\sqrt{\Gamma})$*, this initial value problem has a unique solution on* $[0, T]$ *with*

$$\begin{aligned} q &\in C^3([0,T], K) \\ Y &\in C^0([0,T], H^3) \cap C^1([0,T], H^2) \cap C^2([0,T], H^1). \end{aligned}$$

If the initial data are tangent to the L^2 *orthogonality constraint (3.11) and the pointwise constraint (3.3) then the solution preserves these constraints.*

We will prove this using Picard's method: we iteratively define a sequence $(q^i, Y^i) \in C^0([0,T], K \times H^3)$ which converges to a solution of the initial value problem. To establish that the iteration scheme is well-defined and convergent, the following standard energy estimate for the driven wave equation is key:

Theorem 5.2 (Existence and energy estimate for the wave equation)**.** *The driven wave equation on* $[0, T] \times \Sigma$

$$Y_{tt} - \Delta Y = \Xi$$

with $\Xi : [0,T] \times \Sigma \to \mathbb{R}^3$ *smooth and smooth initial data* $Y_0 = Y(0)$ *and* $Y_1 = Y_t(0)$ *has a unique global solution. The solution is smooth, and there exists an absolute constant* $c(\Sigma) \geq 1$*, depending only on the choice of torus* Σ*, such that*

$$\max\{\|Y_t(t)\|_2, \|Y(t)\|_3\} \leq c(\Sigma) e^t \left\{ \|Y_1\|_2 + \|Y_0\|_3 + \left(\int_0^t \|\Xi(s)\|_2^2 ds \right)^{\frac{1}{2}} \right\}.$$

Proof. Existence, uniqueness and smoothness follow from [7]. Let $E(t) = \|Y_t(t)\|_0^2 + \|Y(t)\|_1^2$. Then

$$\begin{aligned} E'(t) &= 2\langle Y_t, \Delta Y + \Xi \rangle_0 + 2\langle Y, Y_t \rangle_1 = 2\langle Y_t, \Xi \rangle_0 + 2\langle Y_t, Y \rangle_0 \\ &\leq 2\|Y_t\|_0^2 + \|Y\|_0^2 + \|\Xi\|_0^2 \leq 2E(t) + \|\Xi(t)\|_0^2 \qquad (5.1) \\ \Rightarrow \quad \frac{d}{dt}(e^{-2t}E(t)) &\leq e^{-2t}\|\Xi(t)\|_0^2 \leq \|\Xi(t)\|_0^2 \\ \Rightarrow \quad E(t) &\leq e^{2t}\left[E(0) + \int_0^t \|\Xi\|_0^2 \right]. \qquad (5.2) \end{aligned}$$

Now consider $Z = \Delta Y$. This is also smooth and satisfies the wave equation with source $\Delta\Xi$. Applying the above estimate to Z yields

$$\|\Delta Y_t\|_0^2 + \|\Delta Y\|_1^2 \leq e^{2t}\left[\|\Delta Y_1\|_0^2 + \|\Delta Y_0\|_1^2 + \int_0^t \|\Delta\Xi\|_0^2 \right] \leq 2e^{2t}\left[\|Y_1\|_2^2 + \|Y_0\|_3^2 + \int_0^t \|\Xi\|_2^2 \right].$$

Since Δ is an elliptic operator, there exist positive constants α_k, β_k depending only on k and Σ such that

$$\|\Delta Y\|_k^2 + \alpha_k \|Y\|_0^2 \geq \beta_k \|Y\|_{k+2}^2,$$

by the standard elliptic estimate, Proposition 4.3. The result immediately follows. $\square$

To prove Theorem 5.1 we must first write the coupled system (3.15) as an explicit evolution system (note that both equations have $\ddot{q}$ on the right hand side). Let $X = \mathbb{R}^{4n} \times \mathbb{R}^{4n} \times H^3 \times H^2$ given the norm $\|(q, p, Y, Z)\|_X = \max\{|q|, \varepsilon^{-1}|p|, \|Y\|_3, \|Z\|_2\}$. The ε dependence of the norm is chosen so that $\|(0, q_t, 0, 0)\|_X = \dot{q}$. Given any $\Gamma > 0$ let $X_\Gamma = \{(q, p, Y, Z) \in X : q \in K, \|(0, p, Y, Z)\| \leq 8c(\Sigma)\Gamma\}$ where $c(\Sigma) \geq 1$ is the absolute constant obtained from Theorem 5.2. Note that X_Γ is a closed subset of a Banach space, and hence is a complete metric space with respect to the metric induced by $\|\cdot\|_X$. Consider the matrix valued function $M : \mathbb{R} \times X_\Gamma \to End(\mathbb{R}^{4n})$,

$$M(\varepsilon, q, Y)^\mu_\nu = \delta^\mu_\nu - \varepsilon^2 \gamma^{\mu\lambda} \langle Y, \psi_{\lambda\nu} \rangle. \tag{5.3}$$

Since the matrix $(\gamma^{\mu\lambda})$ is postive definite, K is compact, all q-dependence is smooth, and $\|Y\|_0 \leq 8c(\Sigma)\Gamma$ there exists a constant $c(K) > 0$ such that $M : [0, c(K)/\sqrt{\Gamma}] \times X_\Gamma \to GL(4n, \mathbb{R})$ and $M^{-1} : [0, c(K)/\sqrt{\Gamma}] \times X_\Gamma \to GL(4n, \mathbb{R})$ is C^1 and bounded. Hence, for all $\varepsilon \in [0, \varepsilon_*(K, \Gamma)]$, where $\varepsilon_* = c(K)/\sqrt{\Gamma}$ the coupled system can be rewritten

$$\begin{aligned} q_{tt} &= \varepsilon^2 f(\varepsilon, q, q_t, Y, Y_t) && (5.4) \\ Y_{tt} - \Delta Y &= g(\varepsilon, q, q_t, Y, Y_t) && (5.5) \end{aligned}$$

where

$$\begin{aligned} f(\varepsilon, q, q_t, Y, Y_t) &= M^{-1}(\varepsilon, q, Y)(-G(q, q_t, q_t) + \varepsilon h(\varepsilon, q, \varepsilon^{-1} q_t, Y, Y_t)) \\ G(q, u, v)^\mu &= G^\mu_{\nu\lambda}(q) u^\nu v^\lambda && (5.6) \\ g(\varepsilon, q, q_t, Y, Y_t) &= -B_\psi Y - \psi_\mu f^\mu - \psi_{\mu\nu} \frac{q_t^\mu}{\varepsilon} \frac{q_t^\nu}{\varepsilon} + \varepsilon j'(\varepsilon, q, \varepsilon^{-1} q_t, Y, Y_t), && (5.7) \end{aligned}$$

B_ψ denotes the first and zeroth order piece of L_ψ, so $L_\psi = -\Delta + B_\psi$, explicitly

$$\begin{aligned} B_\psi Y &= -(|\psi_x|^2 + |\psi_y|^2) Y - 2(\psi_x \cdot Y_x + \psi_y \cdot Y_y)\psi - 2(\psi \cdot Y)\Delta\psi \\ &\quad -2(\psi \cdot Y)_x \psi_x - 2(\psi \cdot Y)_y \psi_y, && (5.8) \end{aligned}$$

and h and $j' = j + \widehat{j}$ are as defined in (3.14), (3.5), (3.9).

It is convenient henceforth to consider ε as a fixed parameter in $[0, \varepsilon_*(K, \Gamma)]$ and supress the dependence of f, g on ε. The proof of existence will use Picard's method, which requires that f, g be bounded and Lipschitz on X_Γ. This follows quickly from the following proposition, whose proof is straightforward but lengthy, and so is deferred to the appendix:

Proposition 5.3. *The functions f, g are continuously differentiable maps $f : X_\Gamma \to \mathbb{R}^{4n}$ and $g : X_\Gamma \to H^2$. Their differentials $\mathrm{d}f : X_\Gamma \to \mathcal{L}(X, \mathbb{R}^{4n})$, $\mathrm{d}g : X_\Gamma \to \mathcal{L}(X, H^2)$ are bounded, uniformly in ε. That is, there exist constants $\Lambda_f(K, \Gamma), \Lambda_g(K, \Gamma) > 0$ such that*

$$|\mathrm{d}f_x \omega| \leq \Lambda_f \|\omega\|_X, \qquad \|\mathrm{d}f_x \omega\|_2 \leq \Lambda_g \|\omega\|_X$$

for all $x \in X_\Gamma$, $\omega \in X$.

Note that, for Banach spaces B, C, $\mathcal{L}(B, C)$ denotes the space of bounded linear maps $B \to C$, which is itself a Banach space with respect to the norm $\|S\|_{\mathcal{L}(B,C)} = \sup\{\|S(x)\|_C / \|x\|_B : x \in B, x \neq 0\}$.

Corollary 5.4. *The functions $f : X_\Gamma \to \mathbb{R}^{4n}$ and $g : X_\Gamma \to H^2$ are Lipschitz and bounded, uniformly in ε. That is, there exist constants $\Lambda_f, \Lambda_g, C_f, C_g$, depending only on K and Γ, such that for all $x, x' \in X_\Gamma$,*

$$|f(x) - f(x')| \leq \Lambda_f \|x - x'\|_X, \quad \|g(x) - g(x')\|_2 \leq \Lambda_g \|x - x'\|_X, \quad |f(x)| \leq C_f, \quad \|g(x)\|_2 \leq C_g$$

Proof. Let $x_1, x_2 \in X_\Gamma$. Since X_Γ is convex, the curve $x(t) = x_1 + t(x_2 - x_1)$, $0 \leq t \leq 1$ remains in X_Γ. Hence

$$|f(x_1) - f(x_2)| = \left| \int_0^1 \mathrm{d}f_{x(t)}(x_2 - x_1)\, dt \right| \leq \int_0^1 \Lambda_f \|x_2 - x_1\|_X \, dt = \Lambda_f \|x_1 - x_2\|_X. \tag{5.9}$$

From the definition of f one sees that $f(q, 0, 0, 0) = 0$ for all $q \in K$. Hence, for all $x = (q, p, Y, Z) \in X_\Gamma$

$$|f(x)| = |f(q, p, Y, Z) - f(q, 0, 0, 0)| \leq \Lambda_f \|(0, p, Y, Z)\|_X \leq 8c(\Sigma)\Gamma\Lambda_f. \tag{5.10}$$

The proof for g follows mutatis mutandis. ☐

To establish uniqueness of the solution, and to show that q is three times continuously differentiable, we will need the following extension property of $\mathrm{d}f$ and $\mathrm{d}g$, whose proof is also deferred to the appendix:

Proposition 5.5. *The differentials $\mathrm{d}f : X_\Gamma \to \mathcal{L}(X, \mathbb{R}^{4n})$ and $\mathrm{d}g : X_\Gamma \to \mathcal{L}(X, H^2)$ of f and g extend continuously to maps $\mathrm{d}f^{ext} : X_\Gamma \to \mathcal{L}(\mathbb{R}^{4n} \times \mathbb{R}^{4n} \times H^1 \times L^2, \mathbb{R}^{4n})$ and $\mathrm{d}g^{ext} : X_\Gamma \to \mathcal{L}(\mathbb{R}^{4n} \times \mathbb{R}^{4n} \times H^1 \times L^2, L^2)$, bounded by Λ_f and Λ_g respectively.*

The rest of this section is devoted to the proof of Theorem 5.1. Let $T > 0$ be chosen such that

$$T \leq \log 2, \quad T \leq \frac{1}{\varepsilon}, \quad T \leq \frac{d}{\varepsilon(2\Gamma + C_f(K, \Gamma))}, \quad T \leq \frac{\Gamma}{\varepsilon C_f(K, \Gamma)},$$
$$\sqrt{T} \leq \frac{\Gamma}{C_g(K, \Gamma)}, \quad T \leq \frac{1}{4\varepsilon\Lambda_f(K, \Gamma)}, \quad \sqrt{T} \leq \frac{1}{8c(\Sigma)\Lambda_g}. \tag{5.11}$$

Given a complete subset B of a Banach space with norm $\|\cdot\|_B$, denote by $C_T B$ the space of continuous maps $[0, T] \to B$ equipped with the sup norm $|||b||| = \sup\{\|b(t)\|_B : t \in [0, T]\}$. $(C_T B, |||\cdot|||)$ is itself a complete subset of a Banach space.

5.1 Definition of the iteration scheme

We will produce a sequence $\omega^i \in C_T X_\Gamma$ converging to a solution of the initial value problem for the coupled system. Choose and fix $\delta \in (0, \Gamma/4)$, and let $Y_0^i, Y_1^i \in C^\infty(\Sigma, \mathbb{R}^3)$ be sequences such that

$$\|Y_0^i - Y_0\|_3 < \frac{\delta}{2^i}, \qquad \|Y_1^i - Y_1\|_2 < \frac{\delta}{2^i}. \tag{5.12}$$

Such sequences exist since C^∞ is dense in H^k for all $k \geq 0$. Let $\omega^0 = (q_0, \varepsilon q_1, Y_0^0, Y_1^0)$, which is constant in t and smooth on Σ, and trivially lies in $C_T X_\Gamma$. Given ω^i, we define the next iterate to be the solution of the initial value problem $q^{i+1}(0) = q_0, q_t^{i+1}(0) = \varepsilon q_1, Y^{i+1}(0) = Y_0^{i+1}, Y_t^{i+1}(0) = Y_1^{i+1}$ for

$$\begin{aligned} q_{tt}^{i+1} &= \varepsilon^2 f(\omega^i) && (5.13)\\ Y_{tt}^{i+1} - \Delta Y^{i+1} &= g(\omega^i). && (5.14) \end{aligned}$$

We must first check that the sequence ω^i is well defined. So, assume that ω^i is smooth and lies in $C_T X_\Gamma$. Then

$$q^{i+1}(t) = q_0 + \varepsilon t q_1 + \varepsilon^2 \int_0^t \left(\int_0^s f(\omega^i(r)) dr \right) ds, \tag{5.15}$$

which exists since $f \circ \omega^i$ is continuous. Now

$$|q^{i+1}(t) - q_0| \leq \varepsilon T |q_1| + \frac{1}{2} \varepsilon^2 T^2 C_f(K, \Gamma) \leq \varepsilon T \left(\Gamma + \frac{C_f(K, \Gamma)}{2} \right) \leq \frac{d}{2} \tag{5.16}$$

by our choice of T, so $q^{i+1}(t)$ remains in K. Further,

$$\varepsilon^{-1} |q_t^{i+1}(t)| \leq |q_1| + \varepsilon \int_0^t |f(\omega^i(s))| ds \leq \Gamma + T \varepsilon C_f(K, \Gamma) \leq 2\Gamma < 8c(\Sigma)\Gamma \tag{5.17}$$

by our choice of T. Turning to Y^{i+1}, we see by inspection that if Y^i, q^i are smooth, then $g(\omega^i)$ is smooth, so the solution $Y^{i+1}(t)$ exists, is unique and smooth, by Theorem 5.2, which also yields the energy estimate

$$\begin{aligned} \max\{\|Y_t^{i+1}(t)\|_2, \|Y^{i+1}(t)\|_3\} &\leq C(\Sigma) e^t \{\|Y_0^{i+1}\|_3 + \|Y_1^{i+1}\|_2 + \sqrt{t} C_g(K, \Gamma)\} \\ &\leq 2C(\Sigma)\{2(\Gamma + \delta) + \sqrt{T} C_g(K, \Gamma)\} < 8C(\Sigma)\Gamma && (5.18) \end{aligned}$$

by our choice of T and δ. Hence, if ω^i is smooth and in $C_T X_\Gamma$, so is ω^{i+1}. We have already observed that ω^0 is smooth and in $C_T X_\Gamma$, so, by induction, the sequence $\omega^i \in C_T X_\Gamma$ is well-defined.

5.2 Convergence of the iteration scheme

We will now show that ω^i is Cauchy, and hence converges in $C_T X_\Gamma$. From (5.15) one has

$$\begin{aligned} |q^{i+1}(t) - q^i(t)| &= \varepsilon^2 \left| \int_0^t \int_0^s (f(\omega^i(r)) - f(\omega^{i-1}(r))) dr\, ds \right| \\ &\leq \frac{\varepsilon^2}{2} T^2 \Lambda_f(K, \Gamma) |||\omega^i - \omega^{i-1}||| \leq \frac{1}{8} |||\omega^i - \omega^{i-1}||| && (5.19) \end{aligned}$$

by our choice of T. Similarly

$$\begin{aligned} \varepsilon^{-1} |q_t^{i+1}(t) - q_t^i(t)| &= \varepsilon \left| \int_0^t (f(\omega^i(r)) - f(\omega^{i-1}(r))) dr \right| \\ &\leq \varepsilon T \Lambda_f(K, \Gamma) |||\omega^i - \omega^{i-1}||| \leq \frac{1}{4} |||\omega^i - \omega^{i-1}|||. && (5.20) \end{aligned}$$

Now $Z = Y^{i+1} - Y^i$ satisfies the wave equation with source $g(\omega^i) - g(\omega^{i-1})$ and small smooth initial data $\|Z(0)\|_3, \|Z_t(0)\|_2 \leq \delta/2^{i-1}$. Hence, by Theorem 5.2, for each $t \in [0,T]$,

$$\begin{aligned} \max\{\|Z(t)\|_3, \|Z_t(t)\|_2\} &\leq c(\Sigma)e^t \left\{ \frac{\delta}{2^{i-2}} + \left(\int_0^t \|g(\omega^i(s)) - g(\omega^{i-1}(s))\|_2^2 ds \right)^{\frac{1}{2}} \right\} \\ &\leq 2c(\Sigma) \left\{ \frac{\delta}{2^{i-2}} + \sqrt{T}\Lambda_g(K,\Gamma)|||\omega^i - \omega^{i-1}||| \right\} \\ &\leq \frac{c(\Sigma)\delta}{2^{i-3}} + \frac{1}{4}|||\omega^i - \omega^{i-1}|||. \end{aligned} \tag{5.21}$$

Assembling these inequalities, one sees that

$$|||\omega^{i+1} - \omega^i||| \leq \frac{1}{4}|||\omega^i - \omega^{i-1}||| + \frac{\alpha}{2^i} \tag{5.22}$$

where $\alpha = c(\Sigma)\delta/8$. It follows that

$$|||\omega^{i+1} - \omega^i||| \leq \frac{1}{4^i}|||\omega^1 - \omega^0||| + \frac{\alpha}{2^{i-1}}, \tag{5.23}$$

and hence, for all $k \geq 1$,

$$|||\omega^{i+k} - \omega^i||| \leq \sum_{j=1}^{k} |||\omega^{i+j} - \omega^{i+j-1}||| \leq \frac{1}{4^i}|||\omega^1 - \omega^0||| \sum_{j=1}^{\infty} \frac{1}{4^j} + \frac{\alpha}{2^i} \sum_{j=0}^{\infty} \frac{1}{2^j}. \tag{5.24}$$

Hence ω^i is Cauchy with respect to $|||\cdot|||$, so $\omega^i \to \omega = (q,p,Y,Z) \in C_T X_\Gamma$.

5.3 The limit solves the initial value problem

We have established that

$$\begin{aligned} q^i &\to q \quad \text{in} \quad C_T K, & (5.25)\\ q_t^i &\to p \quad \text{in} \quad C_T \mathbb{R}^{4n}, & (5.26)\\ Y^i &\to Y \quad \text{in} \quad C_T H_3, & (5.27)\\ Y_t^i &\to Z \quad \text{in} \quad C_T H^2. & (5.28) \end{aligned}$$

Now, for all i,

$$\begin{aligned} \|\omega(0) - (q_0, \varepsilon q_1, Y_0, Y_1)\|_X &\leq \|\omega(0) - \omega^i(0)\|_X + \|\omega^i(0) - (q_0, \varepsilon q_1, Y_0, Y_1)\|_X \\ &\leq |||\omega - \omega^i||| + \frac{\delta}{2^i} \to 0 \end{aligned} \tag{5.29}$$

as $i \to \infty$. Hence $\omega(0) = (q_0, \varepsilon q_1, Y_0, Y_1)$, that is, the limit has the correct initial data.

We will now show that the limit solves the coupled system and has the differentiability properties claimed. Let $\widetilde{Y}(t) = Y_0 + \int_0^t Z(s)ds$. Note that $\widetilde{Y}$ is manifestly in $C^1([0,T], H^2)$, with derivative $\widetilde{Y}_t = Z$. Now

$$\begin{aligned} |||Y^i - \widetilde{Y}|||_{C_TH^2} &= |||Y^i(0) - Y_0 + \int_0^t (Y_t^i(s) - Z(s))ds|||_{C_TH^2} \\ &\leq \frac{\delta}{2^i} + T|||Y_t^i - Z|||_{C_TH^2} \to 0 \end{aligned} \tag{5.30}$$

as $i \to \infty$. Hence, $Y^i \to \widetilde{Y}$ in C_TH^2. But $Y^i \to Y$ in C_TH^3, hence also in C_TH^2, so $Y = \widetilde{Y}$. Hence, $Y \in C^1([0,T], H^2)$ and $Y_t = Z$.

Consider $Y_{tt}^{i+k} - Y_{tt}^i$. This is smooth, and satisfies the wave equation with source $g(\omega^{i+k-1}) - g(\omega^{i-1})$. Hence

$$\begin{aligned} \|Y_{tt}^{i+k} - Y_{tt}^i\|_1 &= \|\Delta(Y^{i+k} - Y^i) + g(\omega^{i+k-1}) - g(\omega^{i-1})\|_1 \\ &\leq 2\|Y^{i+k} - Y^i\|_3 + \|g(\omega^{i+k-1}) - g(\omega^{i-1})\|_2 \\ &\leq 2\|Y^{i+k} - Y^i\|_3 + \Lambda_g(K,\Gamma)\|\omega^{i+k-1} - \omega^{i-1})\|_X \\ \Rightarrow \quad |||Y_{tt}^{i+k} - Y_{tt}^i|||_{C_TH^1} &\leq [2 + \Lambda_g(K,\Gamma)]|||\omega^{i+k-1} - \omega^{i-1}|||_{C_TX_\Gamma}. \end{aligned} \tag{5.31}$$

Since ω^i is Cauchy in C_TX_Γ, it follows that Y_{tt}^i is Cauchy in C_TH^1. Hence $Y_{tt}^i \to W$ in C_TH^1. Let $\widetilde{Z}(t) = Y_1 + \int_0^t W$. Note that $\widetilde{Z}$ is manifestly in $C^1([0,T], H^1)$ and $\widetilde{Z}_t = W$. Now

$$\begin{aligned} |||Y_t^i - \widetilde{Z}|||_{C_TH^1} &\leq |||Y_1^i - Y_1|||_{C_TH^1} + |||\int_0^t (Y_{tt}^i - W)|||_{C_TH^1} \\ &\leq \frac{\delta}{2^i} + T|||Y_{tt}^i - W|||_{C_TH^1} \to 0 \end{aligned} \tag{5.32}$$

as $i \to \infty$. Hence $Y_t^i \to \widetilde{Z}$ in C_TH^1. But $Y_t^i \to Z$ in C_TH^2, hence also in C_TH^1, so $Z = \widetilde{Z}$. But $Y_t = Z$. Hence, $Y \in C^2([0,T], H^1)$ and $Y_{tt} = W$.

By similar reasoning, $q_{tt}^i \to m$ in $C_T\mathbb{R}^{4n}$ and $q \in C^2([0,T], \mathbb{R}^{4n})$ with $q_t = p$ and $q_{tt} = m$.

We can now show that ω solves the coupled system:

$$\begin{aligned} |||Y_{tt} - \Delta Y - g(\omega)|||_{C_TH^1} &\leq |||Y_{tt} - Y_{tt}^i|||_{C_TH^1} + |||\Delta(Y - Y^i)|||_{C_TH^1} + |||g(\omega) - g(\omega^{i-1})|||_{C_TH^1} \\ &\leq |||Y_{tt} - Y_{tt}^i|||_{C_TH^1} + 2|||Y - Y^i|||_{C_TH^3} + |||g(\omega) - g(\omega^{i-1})|||_{C_TH^2}. \end{aligned} \tag{5.33}$$

Now $Y_{tt}^i \to Y_{tt}$ in C_TH^1, $Y^i \to Y$ in C_TH^3, $g : X_\Gamma \to H^2$ is continuous, and $\omega^i \to \omega$ in C_TX_Γ, so $g(\omega^i) \to g(\omega)$ in C_TH^2. Hence

$$|||Y_{tt} - \Delta Y - g(\omega)|||_{C_TH^1} = 0. \tag{5.34}$$

Similarly $\|q_{tt} - \varepsilon^2 f(\omega)\|_{C_T\mathbb{R}^{4n}} = 0$, that is, $q_{tt} = \varepsilon^2 f(\omega)$.

It remains to establish the higher differentiability of q. Differentiating the equation for q_{tt}^{i+1} gives

$$q_{ttt}^{i+1} = \varepsilon^2 df_{\omega^i}^{ext}\omega_t^i. \tag{5.35}$$

Now $\omega^i \to \omega$ in $C_T X_\Gamma$, $\omega^i_t \to \omega_t$ in $C_T(\mathbb{R}^{4n} \times \mathbb{R}^{4n} \times H^2 \times H^1)$, and df^{ext} is continuous, so $q^{i+1}_{ttt} \to \ell$, say, in $C_T\mathbb{R}^{4n}$. Let $\widetilde{m}(t) = \varepsilon^2 f(q_0, \varepsilon q_1, Y_0, Y_1) + \int_0^t \ell = q_{tt}(0) + \int_0^t \ell$. Note that $\widetilde{m} \in C^1([0,T], \mathbb{R}^{4n})$ and $\widetilde{m}_t = \ell$. Then

$$\|q^i_{tt} - \widetilde{m}\|_{C_T\mathbb{R}^{4n}} = \|\int_0^t (q^i_{ttt} - \ell)\|_{C_T\mathbb{R}^{4n}} \leq T\|q^i_{ttt} - \ell\|_{C_T\mathbb{R}^{4n}} \to 0 \tag{5.36}$$

so $q^i_{tt} \to \widetilde{m}$ in $C_T\mathbb{R}^{4n}$. But $q^i_{tt} \to q_{tt}$ in $C_T\mathbb{R}^{4n}$, so $q_{tt} = \widetilde{m}$. Hence, $q_{tt} \in C^1([0,T], \mathbb{R}^{4n})$, as claimed.

5.4 Uniqueness of the solution

Assume that $(\widetilde{q}, \widetilde{Y})$ is another solution of (5.4), (5.5) with the same initial data and regularity as (q, Y), and let $(p, Z) = (q - \widetilde{q}, Y - \widetilde{Y})$. Then (p, Z) satisfies the system

$$Z_{tt} - \Delta Z = \Xi(t), \qquad p_{tt} = \varepsilon^2 \Upsilon(t) \tag{5.37}$$

with initial data $Z(0) = Z_t(0) = 0$, $p(0) = p_t(0) = 0$, where

$$\Xi(t) = g(\omega) - g(\widetilde{\omega}), \quad \Upsilon(t) = f(\omega) - f(\widetilde{\omega}) \tag{5.38}$$

and $\omega = (q, q_t, Y, Y_t)$, $\widetilde{\omega} = (\widetilde{q}, \widetilde{q}_t, \widetilde{Y}, \widetilde{Y}_t)$. Define

$$E(t) = \|Z(t)\|_1^2 + \|Z_t(t)\|_0^2 + |p|^2 + \frac{1}{\varepsilon^2}|p_t|^2, \tag{5.39}$$

which, by the regularity properties of (p, Z), is continuously differentiable, and has $E(0) = 0$. Reprising the argument in (5.1), which requires only that $Z \in H^2$ and $Z_t \in H^1$, one sees that

$$E'(t) \leq 2E(t) + \|\Xi(t)\|_0^2 + |\Upsilon(t)|^2. \tag{5.40}$$

Now, arguing as in the proof of Corollary 5.4, with $\omega(s) = \omega + s(\widetilde{\omega} - \omega)$,

$$\|\Xi(t)\|_0 = \|g(\omega) - g(\widetilde{\omega})|_0 = \left\| \int_0^1 \mathrm{d}g_{\omega(s)}(\widetilde{\omega} - \omega)\, ds \right\|_0 \leq \Lambda_g \max\{|p|, \varepsilon^{-1}|p_t|, \|Z\|_1, \|Z_t\|_0\} \tag{5.41}$$

by Proposition 5.5. Similarly $|\Upsilon(t)| \leq \Lambda_f \max\{|p|, \varepsilon^{-1}|p_t|, \|Z\|_1, \|Z_t\|_0\}$. Hence

$$E'(t) \leq \kappa E(t) \tag{5.42}$$

where $\kappa = 2 + \Lambda_f + \Lambda_g$, whence it follows that

$$\frac{d}{dt} e^{-\kappa t} E(t) \leq 0. \tag{5.43}$$

So $e^{-\kappa t} E(t)$ is a nonincreasing, non-negative function which is zero at $t = 0$. Hence $E(t) = 0$ for all t, and we conclude that $(p, Z) = (0, 0)$ for all t, that is, $(q, Y) = (\widetilde{q}, \widetilde{Y})$.

5.5 Preservation of constraints

Given the solution (q, Y) produced above, define for each $\mu \in \{1, 2, \dots, 4n\}$

$$a_\mu(t) = \langle Y, \frac{\partial \psi}{\partial q^\mu} \rangle. \tag{5.44}$$

The L^2 orthogonality constraint is that $a_\mu(t) = 0$ for all μ, t. By construction, the coupled system implies that $\ddot{a}_\mu = 0$. If the initial data are tangent to the constraint then $a_\mu(0) = \dot{a}_\mu(0)$, and hence $a_\mu(t) = 0$ for all $t \in [0, T]$.

Similarly, given the solution (q, Y) produced above, define $\chi : [0, T] \times \Sigma \to \mathbb{R}$ by

$$\chi = Y \cdot \psi(q) + \frac{1}{2}\varepsilon^2 |Y|^2. \tag{5.45}$$

The pointwise constraint is that $\chi = 0$ everywhere on Σ. Note that $\chi(t) \in \mathsf{H}^3$ for all $t \in [0, T]$ so $\chi(t) : \Sigma \to \mathbb{R}$ is continuous. Assume that $\chi(0) = 0$ and $\chi_t(0) = 0$, that is the initial data are tangent to the constraint. A straightforward, if lengthy, calculation using the coupled system and the harmonic map equation for ψ shows that χ satsifies the linear PDE

$$\begin{aligned} \chi_{tt} - \Delta\chi &= 2\varepsilon^2 \Big\{ [2\psi_x \cdot Y_x + 2\psi_y \cdot Y_y - |\psi_\tau|^2 + Y \cdot \Delta\psi - 2\varepsilon\psi_\tau \cdot Y_t \\ &\quad -\varepsilon^4(|Y_t|^2 - |Y_x|^2 - |Y_y|^2)]\chi + (Y \cdot \psi_x)\chi_x + (Y \cdot \psi_y)\chi_y \Big\} \\ &=: a\chi + b_1\chi_x + b_2\chi_y \end{aligned} \tag{5.46}$$

where $a(t) \in H^2$, $b_1(t), b_2(t) \in H^3$ for all t. Let $E(t) = \|\chi(t)\|_1^2 + \|\chi_t(t)\|_0^2$. Then

$$\begin{aligned} E'(t) &= 2\langle \chi_t, \Delta\chi + a\chi + b_1\chi_x + b_2\chi_y \rangle_0 + 2\langle \chi_t, \chi \rangle_1 = 2\langle \chi_t, \chi + a\chi + b_1\chi_x + b_2\chi_y \rangle_0 \\ &\leq \|\chi_t\|_0^2 + \|\chi + a\chi + b_1\chi_x + b_2\chi_y\|_0^2 \leq \|\chi_t\|_0^2 + \kappa\|\chi\|_1^2 \leq \kappa E(t) \end{aligned} \tag{5.47}$$

where

$$\kappa = \max_{0 \leq t \leq T} 8(1 + \|a(t)\|_2^2 + \|b_1(t)\|_2^2 + \|b_2(t)\|_2^2). \tag{5.48}$$

Hence $e^{-\kappa t}E(t)$ is a nonincreasing, non-negative function which is zero at $t = 0$, so $E(t) = 0$, whence $\|\chi(t)\|_1 = 0$ for all t. Since we already know that $\chi(t) : \Sigma \to \mathbb{R}$ is continuous, it follows that $\chi = 0$ everywhere. This completes the proof of Theorem 5.1.

6 Near coercivity of the improved Hessian

By repeatedly applying the local existence theorem, we can extend the solution of the coupled system whilever q remains in K and $\varepsilon^{-1}q_t$, $\|Y\|_3$ and $\|Y_t\|_2$ remain bounded. So to prove long time existence, we must, among other things, bound the growth of $\|Y\|_3$. The first step is to show that $\|Y\|_3$ is controlled by the quadratic form $\langle LY, LLY \rangle$ or, more precisely, by the quadratic form $Q_2 : H^3 \to \mathbb{R}$ defined next.

Definition 6.1. For a fixed harmonic map $\psi(q)$ we denote by Q_1, Q_2 the quadratic forms

$$\begin{aligned} Q_1 : H^1 \to \mathbb{R}, \qquad Q_1(Y) &= \int_\Sigma \left\{ |Y_x|^2 + |Y_y|^2 - (|\psi_x|^2 + |\psi_y|^2)|Y|^2 - 4(\psi_x \cdot Y_x + \psi_y \cdot Y_y)\psi \cdot Y \right\} \\ Q_2 : H^3 \to \mathbb{R}, \qquad Q_2(Y) &= Q_1(LY). \end{aligned}$$

Note that both Q_1 and Q_2 are continuous, and that $Q_1(Y) = \langle Y, LY\rangle$ for all $Y \in H^2$. It is also convenient to define the projection map $P : H^k \to H^k$

$$P(Y) = Y - (\psi \cdot Y)\psi$$

which pointwise orthogonally projects $Y(p)$ to $T_{\psi(p)}S^2$.

Lemma 6.2. *For all $Y \in H^2$, $Q_1(Y) \geq Q_1(P(Y))$.*

Proof. $Y = P(Y) + f\psi$, where $f = -\psi \cdot Y \in \mathsf{H}^2$. Now

$$\begin{aligned} Q_1(Y) &= \langle Y, LY\rangle = \langle P(Y), LP(Y)\rangle + 2\langle f\psi, LP(Y)\rangle + \langle f\psi, L(f\psi)\rangle \\ &= \langle P(Y), LP(Y)\rangle + \langle f\psi, L(f\psi)\rangle \end{aligned} \tag{6.1}$$

since L is self-adjoint and maps tangent sections to tangent sections. But, as we saw in Remark 3.2,

$$L(f\psi) = -(\Delta f)\psi - 4(f_x\psi_x + f_y\psi_y), \tag{6.2}$$

so $\langle f\psi, L(f\psi)\rangle = -\langle f, \Delta f\rangle \geq 0$. □

If our error section Y were a tangent section, the results of [5] would immediately imply that Q_1 is coercive, that is, $Q_1(Y) \geq c(q)\|Y\|_1^2$, orthogonal to $\ker J$:

Theorem 6.3 (Haskins-Speight, [5]). *There exists a constant $c(q) > 0$, depending continuously on q, such that*

$$Q_1(Y) \geq c(q)\|Y\|_1^2$$

for all $Y \in H^1$ satisfying $\psi \cdot Y = 0$, L^2 orthogonal to $\ker J_{\psi(q)}$.

Unfortunately, $\psi \cdot Y \neq 0$ in our set-up, but is small (of order ε^2). This means we can only establish the following "near coercivity" property for Q_1. This will suffice for our purposes, however.

Theorem 6.4 (Near coercivity of Q_1). *There exist constants $c(q), \widetilde{c}(q) > 0$, depending continuously on q, such that*

$$Q_1(Y) \geq c(q)\|Y\|_1^2 - \varepsilon^2\widetilde{c}(q)\|Y\|_1\|Y\|_2^2$$

for all $Y \in H^2$ satisfying the pointwise constraint (3.3), L^2 orthogonal to $\ker J_{\psi(q)}$.

Proof. By Lemma 6.2, $Q_1(Y) \geq \langle P(Y), LP(Y)\rangle = \langle P(Y), JP(Y)\rangle$ since $L \equiv J$ on tangent sections. Given any $Z \in \ker J$, $\langle Z, P(Y)\rangle = \langle Z, Y + f\psi\rangle = \langle Z, Y\rangle = 0$, since Z is pointwise orthogonal to ψ and Y is L^2 orthogonal to $\ker J$. Hence $P(Y)$ is L^2 orthogonal to $\ker J$, and so, by Theorem 6.3, there exists a constant $\widetilde{c}(q) > 0$ such that

$$Q_1(Y) \geq \widetilde{c}(q)\|P(Y)\|_1^2. \tag{6.3}$$

Now, since Y satisfies (3.3),

$$\|P(Y)\|_1^2 \;=\; \|Y + \frac{1}{2}\varepsilon^2|Y|^2\psi\|_1^2 \geq \|Y\|_1^2 - \varepsilon^2\langle Y, |Y|^2\psi\rangle_1 \tag{6.4}$$

and, by the algebra property of H^2 (Proposition 4.1),

$$\langle Y, |Y|^2\psi\rangle_1 \;\leq\; \|Y\|_1\||Y|^2\psi\|_2 \leq C\|Y\|_1\|Y\|_2^2\|\psi\|_2 \tag{6.5}$$

where $C > 0$ is a constant depending only on Σ. Combining (6.3), (6.4) and (6.5), and noting that $\|\psi\|_2$ depends continuously (in fact smoothly) on q, the result immediately follows. ☐

Theorem 6.5 (Near coercivity of Q_2)**.** *There exist constants $c(q), \widetilde{c}(q) > 0$, depending continuously on q, such that*

$$Q_2(Y) \geq c(q)\|Y\|_3^2 - \varepsilon^2\widetilde{c}(q)(\|Y\|_3^3 + \varepsilon^2\|Y\|_3^4)$$

for all $Y \in H^3$ satisfying the pointwise constraint (3.3), L^2 orthogonal to $\ker J_{\psi(q)}$.

Proof. Recall that the pointwise constraint (3.3) is equivalent to $|\psi + \varepsilon^2 Y| \equiv 1$. The set of smooth maps $\Sigma \to S^2$ is dense in the Banach manifold of H^3 maps $\Sigma \to S^2$, ψ and $\ker J$ are smooth, and $Q_2 : H^3 \to \mathbb{R}$ is continuous, so it suffices to prove the inequality in the case that Y is smooth. So, let Y be smooth, L^2 orthogonal to $\ker J$ and satisfy the pointwise constraint (3.3). Define the smooth section $Z = LY$ and the smooth real functions $\alpha = |Y|^2$ and $\beta = Z \cdot Y$. Since Y satisfies (3.3), and ψ is harmonic, it follows that

$$\beta = \varepsilon^2\left\{Y \cdot \Delta Y + |Y_x|^2 + |Y_y|^2 + 2(|\psi_x|^2 + |\psi_y|^2)|Y|^2\right\}. \tag{6.6}$$

In the following, $c_1(q), c_2(q), \dots$ denote positive functions depending continuously on q. We have the following elementary estimate,

$$\|\beta\psi\|_1^2 \leq c_1(q)\|\beta\|_1^2 \leq \varepsilon^4 c_2(q)\left\{\|Y\|_{C^1}^2\|Y\|_2^2 + \|Y\|_{C^0}^2\|Y\|_3^2\right\}. \tag{6.7}$$

Applying the Sobolev inequalities (Proposition 4.2) gives

$$\|\beta\psi\|_1 \leq \varepsilon^2 c_3(q)\|Y\|_3^2. \tag{6.8}$$

We will also need to estimate $\|\alpha\|_3$. Again, we have an elementary estimate

$$\|\alpha\|_3^2 \leq c\left\{\|Y\|_{C^1}^2\|Y\|_2^2 + \|Y\|_{C^0}^2\|Y\|_3^2\right\} \tag{6.9}$$

which, on appealing to Proposition 4.2 yields

$$\|\alpha\|_3 \leq c\|Y\|_3^2. \tag{6.10}$$

By Lemma 6.2,

$$Q_2(Y) = Q_1(Z) \leq Q_1(P(Z)) = \langle P(Z), JP(Z)\rangle. \tag{6.11}$$

It is in the last step that we have used the smoothness of Y (assuming only $Y \in H^3$ gives $P(Z) \in H^1$, which is not sufficiently regular to make sense of $JP(Z)$). Since L is self-adjoint and $\ker J \subset \ker L$, $Z = LY$ is automatically L^2 orthogonal to $\ker J$, as is $P(Z) = Z + \beta\psi$ (since ψ is *pointwise* orthogonal to anything in $\ker J$). Hence, by Theorem 6.3 and the estimate (6.8)

$$\begin{aligned} Q_2(Y) &\geq c_4(q)\|P(Z)\|_1^2 = c_4(q)\|Z - \beta\psi\|_1^2 \geq c_4(q)\left\{\|Z\|_1^2 - \|Z\|_1\|\beta\psi\|_1\right\} \\ &\geq c_4(q)\left\{\|Z\|_1^2 - \varepsilon^2 c_3(q)\|Y\|_3^2\|Z\|_1\right\}. \end{aligned} \tag{6.12}$$

We next estimate $\|Z\|_1 = \|LY\|_1$ in terms of $\|Y\|_3$. Note that Y is *not* L^2 orthogonal to $\ker L$, since it has a component in the direction of ψ, so we cannot apply the standard elliptic estimate for L directly. We must decompose

$$Y = P(Y) + (\psi \cdot Y)\psi = P(Y) - \frac{1}{2}\varepsilon^2\alpha\psi, \tag{6.13}$$

using (3.3), and handle the two terms separately. Then, by Proposition 4.3 (for the lower bound on $\|JP(Y)\|_1$), and an elementary estimate (for the upper bound on $\|JP(Y)\|_1$),

$$\begin{aligned} \|JP(Y)\|_1 - \frac{1}{2}\varepsilon^2\|L(\alpha\psi)\|_1 \leq \quad &\|Z\|_1 \quad \leq \|JP(Y)\|_1 + \frac{1}{2}\varepsilon^2\|L(\alpha\psi)\|_1 \\ c_5(q)\|P(Y)\|_3 - \varepsilon^2 c_6(q)\|\alpha\|_3 \leq \quad &\|Z\|_1 \quad \leq c_7(q)\|P(Y)\|_3 + \varepsilon^2 c_6(q)\|\alpha\|_3 \\ c_5(q)\|Y\|_3 - \varepsilon^2 c_8(q)\|\alpha\|_3 \leq \quad &\|Z\|_1 \quad \leq c_7(q)\|Y\|_3 + \varepsilon^2 c_8(q)\|\alpha\|_3 \\ c_5(q)\|Y\|_3 - \varepsilon^2 c_9(q)\|Y\|_3^2 \leq \quad &\|Z\|_1 \quad \leq c_7(q)\|Y\|_3 + \varepsilon^2 c_9(q)\|Y\|_3^2 \end{aligned} \tag{6.14}$$

where we have used (6.10) in the last line. Combining (6.12) and (6.14), the result immediately follows. □

Remark 6.6. Since K is compact, we can replace $c(q)$, $\widetilde{c}(q)$ in Theorems 6.4, 6.5 by global constants $C, \widetilde{C} > 0$, under the extra assumption that $q \in K$.

7 Energy estimates for the coupled system

Having shown that $Q_2(Y)$ controls $\|Y\|_3^2$, for small ε, we must now bound the growth of $Q_2(Y)$ for a solution (q, Y) of the coupled system. We do this by establishing quasi-conservation of energies E_1, E_2, related to Q_1, Q_2:

Definition 7.1. Let $(q, Y) : [0, T] \to K \times H^3$ with the regularity of Theorem 5.1. Associated to (q, Y) we define the energies $E_1, E_2 : [0, T] \to \mathbb{R}$,

$$E_1(t) = \frac{1}{2}\|Y_t\|_0^2 + \frac{1}{2}Q_1(Y), \qquad E_2(t) = \frac{1}{2}\|(LY)_t\|_0^2 + \frac{1}{2}Q_2(Y).$$

Note that E_1 is C^1 and E_2 is continuous.

Throughout this section we will use the following

Convention 7.2. C will denote a positive constant depending (at most) on the choice of Σ and K. $c(a_1, a_2, \ldots, a_p)$ will denote a smooth positive bounding function of p non-negative real arguments, which may also depend (implicitly) on Σ, K, and which is increasing in each of its arguments. C_0, c_0 will denote that the constant or bounding function depends, in addition, on the initial data q_0, q_1, Y_0, Y_1. The value of C, C_0, c, c_0 may vary from line to line.

Theorem 7.3 (Quasi-conservation of E_1)**.** *Let $(q, Y) : [0, T] \to K \times H^3$ be a solution of the coupled system with the initial data and regularity of Theorem 5.1. Then*

$$E_1(t) \le C_0 + C(|\ddot{q}| + |\dot{q}|^2)\|Y(t)\|_0 + \varepsilon \int_0^t c(|\dot{q}|, |\ddot{q}|, |\dddot{q}|, \|Y\|_3, \|Y_t\|_2).$$

Proof. The solution satisfies (3.15) and has $Y \in H^3$, $Y_t \in H^2$, $Y_{tt} \in H^1$, so

$$\begin{aligned}
\frac{dE_1}{dt} &= \langle Y_t, Y_{tt}\rangle + \langle Y_t, LY\rangle + \frac{1}{2}\varepsilon\langle Y, L_\tau Y\rangle = \langle Y_t, k + \varepsilon j'\rangle + \frac{1}{2}\varepsilon\langle Y, L_\tau Y\rangle \\
&= \frac{d}{dt}\langle Y, k\rangle - \varepsilon\langle Y, k_\tau - \frac{1}{2}L_\tau Y\rangle + \varepsilon\langle Y_t, j'\rangle \\
\Rightarrow E_1(t) &= E_1(0) - \langle Y(0), k(0)\rangle + \langle Y(t), k(t)\rangle + \varepsilon \int_0^t \left\{\langle Y, \frac{1}{2}L_\tau Y - k_\tau\rangle + \langle Y_t, j'\rangle\right\} \\
&\le C_0 + \|k(t)\|_0\|Y(t)\|_0 + 2\varepsilon \int_0^t \left\{\|Y\|_0^2 + \|L_\tau Y\|_0^2 + \|k_\tau\|_0^2 + \|Y_t\|_0^2 + \|j'\|_0^2\right\}.
\end{aligned} \tag{7.1}$$

Now, only the first and zeroth order parts of L depend on time, so it is clear that

$$\|L_\tau Y\|_0 \le C|\dot{q}|\|Y\|_1. \tag{7.2}$$

Recall that $k = -\psi_{\tau\tau}$, so

$$\|k(t)\|_0 \le C\|k(t)\|_{C^0} \le C(|\ddot{q}| + |\dot{q}|^2), \tag{7.3}$$

and

$$\|k_\tau(t)\|_0 \le C\|k_\tau(t)\|_{C^0} \le C(|\dddot{q}| + |\ddot{q}|^2 + |\dot{q}|^3 + |\dot{q}|^2). \tag{7.4}$$

Finally, it follows immediately from the algebra property of H^2 (Proposition 4.1) that $\|j'\|_0 \le \|j'\|_2 \le c(|\dot{q}|, \|Y\|_3, \|Y_t\|_2)$, and the result directly follows. □

We will need a similar result bounding the growth of $E_2(t)$. Formally, this is obtained by applying the argument above with Y replaced by LY (which formally solves a PDE of the form $(LY)_{tt} + L(LY) = Lk + O(\varepsilon)$). Unfortunately, this argument is not rigorous since Y is insufficiently regular to make sense of expressions like LY_{tt} (recall Y_{tt} is only H^1).

Theorem 7.4 (Quasi-conservation of E_2)**.** *Let $(q, Y) : [0, T] \to K \times H^3$ be a solution of the coupled system with the initial data and regularity of Theorem 5.1. Then*

$$E_2(t) \le C_0 + C(|\ddot{q}| + |\dot{q}|^2)\|Y(t)\|_2 + \varepsilon \int_0^t c(|\dot{q}|, |\ddot{q}|, |\dddot{q}|, \|Y\|_3, \|Y_t\|_2).$$

Proof. By uniqueness, (q,Y) must arise as the limit of an iteratively defined sequence of smooth functions (q^i,Y^i), as constructed in the proof of Theorem 5.1. Recall that the sections Y^i satisfy the PDEs

$$Y^{i+1}_{tt} - \Delta Y^{i+1} + B^i Y^i = k^i + \varepsilon (j')^i \tag{7.5}$$

where B denotes the first and zeroth order piece of L (so $L = -\Delta + B$) and the superscript i on B^i, k^i, $(j')^i$ denotes that the quantity is evaluated on the iterate (q^i,Y^i). Recall also $(Y^i, Y^i_t, Y^i_{tt}) \to (Y, Y_t, Y_{tt})$ in $C^0(H^3 \oplus H^2 \oplus H^1)$. Now, for each i define

$$\begin{aligned} Z^i &= -\Delta Y^{i+1} + B^i Y^i \\ E^i(t) &= \frac{1}{2}\|Z^i_t\|^2 + \frac{1}{2}\langle Z^i, -\Delta Z^i + B^i Z^i\rangle. \end{aligned} \tag{7.6}$$

Each $E^i : [0,T] \to \mathbb{R}$ is smooth, $E^i \to E_2$ uniformly on $[0,T]$, $Z^i \to Z = LY$ in H^1 and $Z^i_t \to (LY)_t$ in L^2. Note that Z^i satisfies the PDE

$$Z^i_{tt} - \Delta Z^i + B^i Z^i = B^i(Z^i - Z^{i-1}) + \hat{k}^i + \varepsilon \hat{j}^i \tag{7.7}$$

where $\hat{k}^i = -\Delta k^i + B^i k^{i-1}$ and $\hat{j}^i = -\Delta (j')^i + B^i (j')^{i-1}$. Since B^i is self-adjoint, one has

$$\begin{aligned} \frac{dE^i}{dt} &= \langle Z^i_t, Z^i_{tt} - \Delta Z^i + B^i Z^i\rangle + \frac{\varepsilon}{2}\langle Z^i, B^i_\tau Z^i\rangle \\ &= \langle Z^i_t, B^i(Z^i - Z^{i-1}) + \hat{k}^i\rangle + \varepsilon\left\{\langle Z^i_t, \hat{j}^i\rangle + \frac{1}{2}\langle Z^i, B^i_\tau Z^i\rangle\right\} \\ &= \frac{d}{dt}\langle Z^i, B^i(Z^i - Z^{i-1}) + \hat{k}^i\rangle - \langle B^i Z^i, Z^i_t - Z^{i-1}_t\rangle \\ &\quad + \varepsilon\{\langle Z^i, B^i_\tau(\tfrac{1}{2}Z^i - Z^{i-1}) + \hat{k}^i_\tau\rangle + \langle Z^i_t, \hat{j}^i\rangle\}. \end{aligned} \tag{7.8}$$

Integrating this from 0 to t and taking the limit $i \to \infty$ yields

$$E_2(t) - E_2(0) \le C_0 + \|Z(t)\|_0 \|\hat{k}(t)\|_0 + C\varepsilon \int_0^t \{\|Z\|_0(\|\hat{k}_\tau\|_0 + |\dot{q}|\|Z\|_1) + \|Z_t\|_0\|\hat{j}\|_0\} \tag{7.9}$$

where we have used the facts that $Z^i_t - Z^{i-1}_t \to 0$ in L^2, $Z^i - Z^{i-1} \to 0$ in H^1, $\hat{k}^i \to \hat{k} = Lk$ in H^3, $\hat{k}^i_\tau \to \hat{k}_\tau$ in H^3 and $\hat{j}^i \to \hat{j} = Lj'$ in L^2 (since $L : H^2 \to L^2$ and $j : K \times \mathbb{R}^{4n} \times H^3 \times H^2 \to H^2$ are continuous). Now, we have the elementary estimates

$$\begin{aligned} \|Z\|_k &= \|LY\|_k \le C\|Y\|_{k+2}, \qquad k = 1,2, \\ \|Z_t\|_0 &= \|LY_t + \varepsilon B_\tau Y\|_0 \le C(\|Y_t\|_2 + \varepsilon|\dot{q}|\|Y\|_1), \\ \|\hat{k}\|_0 &\le C(|\ddot{q}| + |\dot{q}|^2), \\ \|\hat{k}_\tau\|_0 &\le C(|\dddot{q}| + |\ddot{q}|^2 + |\dot{q}|^2 + |\dot{q}|^3) \\ \|\hat{j}\|_0 &\le C\|j'\|_2 \le c(|\dot{q}|, \|Y\|_3, \|Y_t\|_2). \end{aligned}$$

Combining these with (7.9), the result follows. ∎

8 Long time existence and proof of the main theorem

Throughout this section we choose and fix $q_0 \in K$ and $q_1 \in \mathbb{R}^{4n}$, and denote by (q, Y) the solution of the coupled system (3.15) with initial data $q(0) = q_0$, $\dot{q}(0) = q_1$, $Y(0) = Y_t(0) = 0$. By Theorem 5.1, provided $\varepsilon < C(K)/\sqrt{|q_1|}$, this solution exists at least for time $t \in [0, T_0]$, where T_0 depends on the initial data, but is independent of ε. Moreover, the solution is unique, has the advertised regularity, satisfies the pointwise and L^2 orthogonality constraints (3.3), (3.11), and obeys the energy estimates of section 7. Denote by $q_*(\tau)$ the geodesic in (M_n, γ) with the same initial data, $q_*(0) = q_0$, $\dot{q}_*(0) = q_1$. Note that q_*, considered as a function of rescaled time τ, is independent of ε. Since geodesic flow conserves speed $\gamma(\dot{q}_*, \dot{q}_*)$, which uniformly bounds $|\dot{q}_*|^2$ on K, there exist $\tau_0 > 0$, $\alpha_0 > 0$, depending only on the initial data, such that q_* exists and has

$$|\dot{q}_*| \leq \alpha_0, \quad |\ddot{q}_*| \leq \alpha_0, \quad \mathrm{dist}\,(q(\tau), \partial K)) < d/2, \tag{8.1}$$

for all $\tau \in [0, \tau_0]$, where $d = \mathrm{dist}\,(q_0, \partial K)$. Hence, the geodesic q_* exists for time $t \in [0, \varepsilon^{-1}\tau_0]$ which, for ε small, exceeds T_0. Whilever q, q_* both exist, we define $\varepsilon^2\widetilde{q}(t)$ to be the error between them, that is

$$q = q_* + \varepsilon^2\widetilde{q}, \tag{8.2}$$

and

$$M(s) = \max_{0\leq t\leq s} \left\{\varepsilon^2|\widetilde{q}(t)|^2 + |\widetilde{q}'(t)|^2 + |\widetilde{q}''(t)|^2 + \|Y(t)\|_3^2 + \|Y_t(t)\|_2^2\right\}, \tag{8.3}$$

where primes denote differentiation with respect to t. This function, which measures the total error in replacing the wave map $\phi = \psi(q) + \varepsilon^2 Y$ with the geodesic $\psi(q_*)$, is continuous, manifestly increasing, and has initial value $M(0) = 0$. Our next task is to bound its growth. Before doing so, we define another absolute constant (depending only on K and Σ), which will appear frequently in this section:

$$\alpha_a = \sup\{\|\gamma^{\mu\nu}\psi_{\nu\lambda}\| \;:\; q \in K, 0 \leq \mu, \lambda \leq 4n\}. \tag{8.4}$$

We will again use Convention 7.2 regarding bounding constants and functions.

Theorem 8.1 (A priori bound)**.** *Whilever (q, Y) exists, and $t < \tau_0/\varepsilon$, and $M(t) < \varepsilon^{-4}\alpha_a^{-2}$,*

$$M(t) \leq C_0 + C_0 M(t)^{\frac{1}{2}} + (\varepsilon^2 + \varepsilon t + \varepsilon^2 t^2 + \varepsilon^4 t^4)\frac{c_0(M(t))}{1 - \varepsilon^2\alpha_a M(t)^{\frac{1}{2}}}.$$

Proof. We first derive the ODE satisfied by $\widetilde{q}$. The curve $q(\tau)$ satisfies the ODE

$$\ddot{q} + G(q, \dot{q}, \dot{q}) = \varepsilon h(\varepsilon, q, \dot{q}, Y, Y_t) + \varepsilon^2 a(q, \ddot{q}, Y) \tag{8.5}$$

where G, h are defined in (3.14), (5.7), and

$$a : K \times \mathbb{R}^{4n} \times L^2 \to \mathbb{R}^{4n}, \quad a(q, v, Y)^\mu = \gamma^{\mu\nu}\langle\psi_{\nu\lambda}, Y\rangle v^\lambda, \tag{8.6}$$

which is smooth with respect to q and linear with respect to v and Y. The geodesic satisfies the ODE

$$\ddot{q}_* + G(q_*, \dot{q}_*, \dot{q}_*) = 0 \tag{8.7}$$

with the same initial data. Substituting $q = q_* + \varepsilon^2 \widetilde{q}$ into (8.5), and using (8.7), we see that $\widetilde{q}$ satisfies

$$\begin{aligned} \widetilde{q}'' &= [G(q_*, \dot{q}_*, \dot{q}_*) - G(q_* + \varepsilon^2 \widetilde{q}, \dot{q}_*, \dot{q}_*)] + [G(q, \dot{q}_*, \dot{q}_* - G(q, \dot{q}_* + \varepsilon \widetilde{q}', \dot{q}_* + \varepsilon q')] \\ &\quad + \varepsilon h(\varepsilon, q, \dot{q}, Y, Y_t) + \varepsilon^2 a(q, Y, \ddot{q}_* + \widetilde{q}''). \end{aligned} \tag{8.8}$$

Now, q, by assumption, remains in K, and for all $t \in [0, \tau_0/\varepsilon]$, q_* remains in K and $|\dot{q}_*| \leq \alpha_0$, so, since $G(q, u, v)$ is Lipshitz with respect to q (on K) and bilinear in (u, v), we have

$$\begin{aligned} |\widetilde{q}''| &\leq C\alpha_0^2 \varepsilon^2 |\widetilde{q}| + C[\varepsilon^2 |\widetilde{q}'|^2 + \varepsilon \alpha_0 |\widetilde{q}'| + \varepsilon |\widetilde{q}'|]\varepsilon |h| + \varepsilon^2 |a| \\ &\leq C_0 \varepsilon M^{\frac{1}{2}} + C\varepsilon^2 M + \varepsilon |h| + \varepsilon^2 |a|. \end{aligned} \tag{8.9}$$

To estimate the h term, we note that it is smooth in q and polynomial in $\dot{q}$ and Y and its (first) derivatives, so in light of Proposition 4.1 we have (for $\varepsilon \leq 1$) the crude bound

$$|h(\varepsilon, q, \dot{q}, Y, Y_t)| \leq c(|\dot{q}|, \|Y\|_3, \|Y_t\|_2). \tag{8.10}$$

Now, by the definition of M,

$$|\dot{q}| \leq |\dot{q}_*| + \varepsilon |\widetilde{q}'| \leq \alpha_0 + \varepsilon M^{\frac{1}{2}}, \tag{8.11}$$

so $|h| \leq c_0(M)$. Turning to $|a|$, we have by linearity,

$$|a(q, Y, \ddot{q}_* + \widetilde{q}'')| \leq \alpha_a \|Y\|(|\ddot{q}_*| + |\widetilde{q}''|) \leq \alpha_a M^{\frac{1}{2}}(\alpha_0 + M^{\frac{1}{2}}) \leq c_0(M). \tag{8.12}$$

Hence,

$$|\widetilde{q}''(t)| \leq \varepsilon c_0(M(t)). \tag{8.13}$$

Now $\widetilde{q}'(t) = \widetilde{q}'(0) + \int_0^t \widetilde{q}'' = \int_0^t \widetilde{q}''$, so

$$|\widetilde{q}'(t)| \leq \varepsilon \int_0^t c_0(M(s))\, ds \leq \varepsilon t c_0(M(t)) \tag{8.14}$$

since c_0 and M are, by definition, increasing. Similarly

$$\varepsilon |\widetilde{q}(t)| \leq \varepsilon \int_0^t |\widetilde{q}'| \leq \varepsilon^2 t^2 c_0(M(t)). \tag{8.15}$$

We have now bounded the growth of all the $\widetilde{q}$ terms in M. To bound the growth of $\|Y\|_3$ and $\|Y_t\|_2$, we will use the energy estimates of section 7 and the near coercivity property of Q_2 (Theorem 6.5). But to do this, we need to control $|\ddot{q}|$ and $|\dddot{q}|$, which appear in the energy estimates for $E_1(t)$ and $E_2(t)$, so we have not yet finished with the ODE for q. For $\ddot{q}$ we have from (8.13) the obvious bounds

$$|\ddot{q}| \leq |\ddot{q}_*| + |\widetilde{q}''| \leq \alpha_0 + \varepsilon c_0(M(t)). \tag{8.16}$$

For $\dddot{q}$ we must work harder. So, for fixed ε, let h_i, $i = 1, 2, 3, 4$, denote the partial derivatives of $h : K \times \mathbb{R}^{4n} \times H^3 \times H^2 \to \mathbb{R}^{4n}$ with respect to each of its four entries. Similarly, let G_1,

a_1 denote the derivatives of G, a with respect to their first entries. Differentiating (8.5) with respect to τ one finds

$$\dddot{q} + G_1(q,\dot q,\dot q)\dot q + 2G(q,\dot q,\dot q)\ddot q \;=\; \varepsilon h_1 \dot q + h_2 \ddot q + h_3 Y_t + h_4 Y_{tt} + \varepsilon^2(a_1(q,\ddot q, Y)\dot q + a(q,\dddot{q},Y)) + \varepsilon a(q,\ddot q, Y_t) \quad (8.17)$$

where we have used the bilinearity properties of G and a. Inspecting the formula for h (3.14), we obtain, using (8.16) and (8.11) a crude bound

$$\begin{aligned}
|\dddot{q}| &\le c_0(M) + c_0(M)\|Y_{tt}\| + \varepsilon^2\alpha_a|\dddot{q}|\|Y\| \\
&\le c_0(M) + c_0(M)(\|LY\| + \|k\| + \varepsilon\|j'\|) + \varepsilon^2\alpha_a M^{\frac12}|\dddot{q}| \\
&\le c_0(M) + c_0(M)(\|Y\|_2 + |\ddot q| + |\dot q|^2 + c(|\dot q|, \|Y\|_3, \|Y_t\|_2)) + \varepsilon^2\alpha_a M^{\frac12}|\dddot{q}| \\
&\le c_0(M) + \varepsilon^2\alpha_a M^{\frac12}|\dddot{q}|. \qquad (8.18)
\end{aligned}$$

Hence

$$|\dddot{q}| \le \frac{c_0(M)}{1-\varepsilon^2\alpha_a M^{\frac12}} \qquad (8.19)$$

whilever $M(t) < \varepsilon^{-4}\alpha_a^{-2}$. We can now turn to bounding Y.

Since Q_2 is nearly coercive (Theorem 6.5),

$$\|Y\|_3^2 \le C\{Q_2(Y) + \varepsilon^2 c(\|Y\|_3)\}, \qquad (8.20)$$

and, by the standard elliptic estimate for L (Proposition 4.3)

$$\begin{aligned}
\|Y_t\|_2^2 &\le C\{\|LY_t\|^2 + \|Y_t\|^2\} \le C\{\|(LY)_t\|^2 + \varepsilon^2\|L_\tau Y\|^2 + \|Y_t\|^2\} \\
&\le C\{\|(LY)_t\|^2 + \varepsilon^2|\dot q|^2\|Y\|_1^2 + \|Y_t\|^2\} \qquad (8.21)
\end{aligned}$$

since the principal part of L does not depend on time. Adding (8.20) and (8.21), one sees that

$$\|Y\|_3^2 + \|Y_t\|_2^2 \le C\{E_1(t) + E_2(t) + \varepsilon^2 c(\|Y\|_3) + \varepsilon^2|\dot q|^2\|Y\|_1^2\} \le C\{E_1(t)+E_2(t)+\varepsilon^2 c_0(M)\}, \qquad (8.22)$$

where E_1, E_2 are as in Definition 7.1. Then, by Theorem 7.3 and 7.4, and the estimates (8.11), (8.16), (8.19),

$$\begin{aligned}
\|Y\|_3^2 + \|Y_t\|_2^2 &\le C_0 + (C_0 + \varepsilon c_0(M))M^{\frac12} + \varepsilon\int_0^t \left(1 + \frac{c_0(M)}{1-\varepsilon^2\alpha_a M^{\frac12}}\right) c_0(M) + \varepsilon^2 c_0(M) \\
&\le C_0 + C_0 M^{\frac12} + \varepsilon c_0(M) + \frac{\varepsilon t c_0(M)}{1-\varepsilon^2\alpha_a M^{\frac12}}, \qquad (8.23)
\end{aligned}$$

provided $\varepsilon \le 1$ and $M(t) < \varepsilon^{-4}\alpha_a^{-2}$. Combining (8.23), (8.15), (8.14) and (8.13) gives the a priori bound claimed. □

Theorem 8.2 (Long time existence)**.** *There exist $\varepsilon_* > 0$ and $\tau_* > 0$, depending only on the initial data, such that for all $\varepsilon \in (0, \varepsilon_*)$ the solution (q, Y) persists for all $t \in [0, \tau_*/\varepsilon]$ and has $M(t)$ bounded, independent of ε.*

Proof. Choose and fix a constant $M_* > 4\alpha_0^2$ so large that $M_* > C_0 + C_0 M_*^{\frac{1}{2}}$, where C_0 is the specific constant, depending only on initial data, appearing in the a priori bound. Assume $\varepsilon > 0$ is so small that

$$\varepsilon^2 < \frac{1}{\alpha_a M_*^{\frac{1}{2}}}, \quad \varepsilon < \frac{d}{4M_*^{\frac{1}{2}}}, \quad \varepsilon \leq \frac{1}{2}, \quad \varepsilon < \frac{C(K)}{M_*^{\frac{1}{4}}} \tag{8.24}$$

where, as before, $d = \mathrm{dist}\,(q_0, \partial K)$ and $C(K) > 0$ is the constant quoted in Theorem 5.1. Then, whilever $M(t) \leq M_*$, $\|Y\|_3 < M_*^{\frac{1}{2}}$, $\|Y_t\|_2 \leq M_*^{\frac{1}{2}}$,

$$|\dot{q}| \leq \alpha_0 + \varepsilon M_*^{\frac{1}{2}} < M_*^{\frac{1}{2}}, \tag{8.25}$$

and

$$\varepsilon^2 \|\widetilde{q}\| \leq \varepsilon M^{\frac{1}{2}} < \varepsilon M_*^{\frac{1}{2}} < \frac{d}{4}. \tag{8.26}$$

Hence, whilever $M(t) \leq M_*$, the value of the solution $(q, q_t, Y, Y_t)(t)$ satisfies the conditions of the initial data for the local existence theorem 5.1, with $\Gamma = M_*^{\frac{1}{2}}$ and $\mathrm{dist}\,(q(t), \partial K) < \frac{d}{4}$. Given the last condition on ε, (8.24), it follows that we may apply Theorem 5.1 and extend the solution for a time $\delta T > 0$ depending only on M_* and d, independent of ε. It follows that the solution persists for as long as $M(t) \leq M_*$. Furthermore, by the first condition on ε, (8.24), whilever $M(t) \leq M_*$ the solution obeys the a priori bound, Theorem 8.1.

For each $\varepsilon > 0$ let $t_\varepsilon = \sup\{t \,:\, M(t) \leq M_*\}$. We claim that there exists $\varepsilon_* > 0$ such that $\varepsilon t_\varepsilon$ is bounded away from zero on $[0, \varepsilon_*]$. Note that this immediately implies the statement in the theorem since then there exists $\tau_* > 0$ such that $t_\varepsilon \geq \tau_*/\varepsilon$ for all $\varepsilon \in (0, \varepsilon_*)$, and the solutions exists, with $M(t) \leq M_*$ on $[0, \tau_*/\varepsilon]$. Assume, towards a contradiction, that no such ε_* exists. Then there is a positive sequence $\varepsilon_i \to 0$ such that $\varepsilon_i t_{\varepsilon_i} \to 0$. But then the a priori bound at time t_{ε_i} gives (recall M is continuous, so $M(t_{\varepsilon_i}) = M_*$), in the limit $i \to \infty$,

$$M_* < C_0 + C_0 M_*^{\frac{1}{2}}, \tag{8.27}$$

a contradiction, by our choice of M_*. □

Proof of Main Theorem. By Theorem 8.2, for all $\varepsilon \in (0, \varepsilon_*]$ the solution exists for $t \in [0, \tau_*/\varepsilon]$, and coincides with $\psi(q_*(\tau) + \varepsilon^2 \widetilde{q}(t)) + \varepsilon^2 Y(t)$, with $\varepsilon|\widetilde{q}(t)|$, $\|Y(t)\|_3$, and hence (by Proposition 4.2) $\|Y(t)\|_{C^0}$ uniformly bounded in t and ε. The rescaled solution is

$$\phi^\varepsilon : [0, \tau_*] \times \Sigma \to S^2 \subset \mathbb{R}^3, \qquad \phi^\varepsilon(\tau, x, y) = \psi(q_*(\tau) + \varepsilon^2 \widetilde{q}(\tau/\varepsilon), x, y) + \varepsilon^2 Y(\tau/\varepsilon, x, y), \tag{8.28}$$

and the geodesic with the same initial data is

$$\psi_* : [0, \tau_*] \times \Sigma, \qquad \psi_*(\tau, x, y) = \psi(q_*(\tau), x, y). \tag{8.29}$$

Now $\psi : K \times \Sigma \to S^2$ is smooth, hence uniformly continuous (since $K \times \Sigma$ is compact). Hence, as $\varepsilon \to 0$, ϕ^ε converges uniformly on $[0, \tau_*] \times \Sigma$ to ψ_*. Furthermore,

$$\begin{aligned}
\phi^\varepsilon_x(\tau, x, y) &= \psi_x(q_*(\tau) + \varepsilon^2 \widetilde{q}(\tau/\varepsilon), x, y) + \varepsilon^2 Y_x \\
\phi^\varepsilon_y(\tau, x, y) &= \psi_y(q_*(\tau) + \varepsilon^2 \widetilde{q}(\tau/\varepsilon), x, y) + \varepsilon^2 Y_y \\
\phi^\varepsilon_\tau(\tau, x, y) &= (\dot{q}_*^\mu + \varepsilon \widetilde{q}'(\tau/\varepsilon))\psi_\mu(q_*(\tau) + \varepsilon^2 \widetilde{q}(\tau/\varepsilon), x, y) + \varepsilon Y_t
\end{aligned} \tag{8.30}$$

and $|\tilde{q}'(t)|$, $\|Y_x(t)\|_{C^0}$, $\|Y_y(t)\|_{C^0}$, $\|Y_t(t)\|_{C^0}$ are bounded uniformly in t and ε (again using Proposition 4.2), so $\phi^\varepsilon_x, \phi^\varepsilon_y, \phi^\varepsilon_\tau$ converge uniformly on $[0,\tau_*]\times\Sigma$ to $\psi_{*x}, \psi_{*y}, \psi_{*\tau}$. Hence, ϕ^ε converges to ψ_* in C^1. □

Appendix: Analytic properties of the nonlinear terms

The proof of the local existence theorem 5.1 makes fundamental use of certain basic analytic properties (Propositions 5.3 and 5.5) of the right hand sides f, g of the evolution system (5.4),(5.5). These properties are established by a long chain of elementary arguments which we sketch in this appendix. We regard ε as a fixed parameter in $(0,1)$ and choose a local parametrization $\psi : U\times\Sigma \to S^2$ of M_n as given by Proposition 2.1 and a compact convex neighbourhood $K\subset U$. We begin by showing that the associated maps $\Psi_k : U \to H^k$, $q\mapsto \psi(q,\cdot)$ are smooth for all $k\in\mathbb{N}$.

Lemma A.1. *Let $k\in\mathbb{N}$ and $f : U\times\Sigma\to\mathbb{R}$ be smooth. Then $F : U\to \mathsf{H}^k$, $F(q) = f(q,\cdot)$, is smooth.*

Proof. As usual, we will denote partial derivatives with respect to q^μ by a subscript μ. It suffices to show that F is everywhere differentiable, since all partial derivatives $f_{\mu_1\mu_2\cdots\mu_r}$ are, like f, smooth maps $U\times\Sigma\to\mathbb{R}$. By the mean value theorem there exist $t_1,\ldots,t_r\in[0,1]$ such that

$$\begin{aligned}\|F(q+p)-F(q)-f_\mu(q,\cdot)p^\mu\|_k^2 \;\; &\leq \;\; |p|^2\sum_{\mu=1}^{4n}\int_\Sigma \big\{(f_\mu(q+t_1p)-f_\mu(q))^2 \\ &\quad +(f_{\mu x}(q+t_2p)-f_{\mu x}(q))^2+\cdots \\ &\quad \cdots+(f_{\mu y\cdots y}(q+t_rp)-f_{\mu y\cdots y}(q))^2\big\} \qquad \text{(A.1)}\end{aligned}$$

since f and all its partial derivatives are C^1 functions of q. But $f_\mu,\ldots,f_{\mu y\cdots y}$ are uniformly continuous on $B_\delta(q)\times\Sigma$, for $\delta>0$ sufficiently small, so

$$\lim_{p\to 0}\frac{1}{|p|}\|F(q+p)-F(q)-\left.\frac{\partial f}{\partial q^\mu}\right|_q p^\mu\|_k = 0 \qquad \text{(A.2)}$$

as was to be proved. □

Now, by the definition of H^k, $\psi : U\to H^k$ is differentiable if and only if $\psi_i : U\to\mathsf{H}^k$ is differentiable for $i=1,2,3$, so we immediately obtain:

Corollary A.2. $\Psi_k : U\to H^k$, $q\mapsto\psi(q,\cdot)$, *is smooth for all* $k\in\mathbb{N}$.

The error terms are $j'(q,\dot q, Y, Y_t)$ and $h(q,\dot q,Y,Y_t)$ where

$$\begin{aligned} j'(q,p,Y,Z) \;\;&=\;\; 2p^\mu(\psi_\mu\cdot Z)\psi+\varepsilon(|Z|^2-|Y_x|^2-|Y_y|^2)\psi+\varepsilon p^\mu p^\nu(\psi_\mu\cdot\psi_\nu)Y \\ &\quad -2\varepsilon(\psi_x\cdot Y_x+\psi_y\cdot Y_y)Y+\varepsilon^2\{|Y|^2\Delta\psi+2(Y\cdot Y_x)\psi_x+2(Y\cdot Y_y)\psi_y\} \\ &\quad +2\varepsilon^2p^\mu(\psi_\mu\cdot Z)Y+\varepsilon^3(|Z|^2-|Y_x|^2-|Y_y|^2)Y \\ h(q,p,Y,Z)^\mu \;\;&=\;\; \gamma^{\mu\nu}\{\langle Z,\psi_{\nu\lambda}\rangle p^\lambda+\varepsilon\langle Y,\psi_{\lambda\nu\rho}\rangle p^\lambda p^\rho+\varepsilon\langle\psi_\lambda\cdot\psi_\rho Y,\psi_\nu\rangle p^\lambda p^\rho \\ &\quad -2\varepsilon\langle(\psi_x\cdot Y_x+\psi_y\cdot Y_y)Y,\psi_\nu\rangle+2\varepsilon^2\langle(\psi_\lambda\cdot Z)Y,\psi_\nu\rangle p^\lambda \\ &\quad +\varepsilon^3\langle(|Z|^2-|Y_x|^2-|Y_y|^2)Y,\psi_\nu\rangle\}. \qquad \text{(A.3)}\end{aligned}$$

We will also need to consider the quantities

$$M^\mu{}_\nu(q,Y) = \delta^\mu{}_\nu - \varepsilon^2\gamma^{\mu\lambda}(q)\langle Y, \psi_{\lambda\nu}\rangle \tag{A.4}$$

$$\begin{aligned} A(q,Y) &= -(|\psi_x|^2+|\psi_y|^2)Y - 2(\psi_x\cdot Y_x + \psi_y\cdot Y_y)\psi - 2(\psi\cdot Y)\Delta\psi \\ &\quad -2(\psi\cdot Y)_x\psi_x - 2(\psi\cdot Y)_y\psi_y. \end{aligned} \tag{A.5}$$

Let B denote the Banach space $\mathbb{R}^{4n}\times\mathbb{R}^{4n}\times H^3\times H^2$ with norm

$$\|(q,p,Y,Z)\|_B = \max\{|q|,|p|,\|Y\|_3,\|Z\|_2\}, \tag{A.6}$$

$B_U\subset B$ denote the open set on which $q\in U$ and for each $\Gamma\geq 0$, $B_\Gamma\subset B_U$ denote the closed convex subset on which $q\in K$ and $\|(0,p,Y,Z)\|_B\leq\Gamma$.

Proposition A.3. *The quantities defined above are smooth maps $j':B_U\to H^2$, $h:B_U\to\mathbb{R}^{4n}$, $M^\mu{}_\nu:B_U\to\mathbb{R}$, $A:B_U\to H^2$*

Proof. That j',h define maps $B_U\to H^2$ and $B_U\to\mathbb{R}^{4n}$ follows immediately from the algebra property of H^2 (Proposition 4.1). Now j' is a linear combination of terms formed by composing the maps

$$\begin{aligned} \Psi_k : U\to H^k, &\qquad q\mapsto\psi(q,\cdot) \\ \mathrm{d}\Psi_k : U\times\mathbb{R}^{4n}\to H^k, &\qquad (q,p)\mapsto p^\mu\psi_\mu(q,\cdot) \end{aligned} \tag{A.7}$$

which are smooth by Corollary A.2, and the manifestly smooth maps

$$\begin{aligned} H^k\to\mathsf{H}^k, &\qquad Y\mapsto Y_i \\ \mathsf{H}^k\to H^k, &\qquad f\mapsto fe_i \\ H^k\to H^{k-1}, &\qquad Y\mapsto Y,\ Y\mapsto Y_x,\ Y\mapsto Y_y \\ \mathsf{H}^2\times\mathsf{H}^2\to\mathsf{H}^2, &\qquad (f,g)\to fg, \end{aligned} \tag{A.8}$$

where $e_1=(1,0,0)$, $e_2=(0,1,0)$ and $e_3=(0,0,1)$. Hence j' is smooth.

The map h is handled similarly, after noting that the inverse metric coefficients $\gamma^{\mu\nu}$ are smooth $U\to\mathbb{R}$, the higher derivatives

$$\begin{aligned} \mathrm{d}^2\Psi_k : U\times\mathbb{R}^{4n}\times\mathbb{R}^{4n}\to H^k, &\qquad (q,p_1,p_2)\mapsto p_1^\mu p_2^\nu\psi_{\mu\nu}(q,\cdot) \\ \mathrm{d}^3\Psi_k : U\times\mathbb{R}^{4n}\times\mathbb{R}^{4n}\times\mathbb{R}^{4n}\to H^k, &\qquad (q,p_1,p_2,p_3)\mapsto p_1^\mu p_2^\nu p_3^\lambda\psi_{\mu\nu\lambda}(q,\cdot) \end{aligned} \tag{A.9}$$

are smooth by Lemma A.2 and, in addition to the maps in (A.8) above, the map

$$H^0\times H^0\to\mathbb{R}, \qquad (Y,Z)\mapsto\langle Y,Z\rangle \tag{A.10}$$

is manifestly smooth.

That $M^\mu{}_\nu$ and A define smooth maps on B_U is clear, since their q-dependence is smooth, and they depend linearly on Y (and are independent of p and Z). □

We can now assemble these pieces to show that f and g are smooth functions on X_Γ, the space defined in section 5. To do so, we note that $f = \widehat{f} \circ \iota$ and $g = \widehat{g} \circ \iota$ where $\iota : X_\Gamma \to B_{8c(\Sigma)\Gamma}$ is the linear isometry

$$\iota : (q, p, Y, Z) \to (q, \varepsilon^{-1} p, Y, Z)$$

and $\widehat{f} : B_{\Gamma'} \to \mathbb{R}^{4n}$,$\widehat{g} : B_{\Gamma'} \to H^2$ are

$$\widehat{f}(q, p, Y, Z) = M^{-1}(q, Y)(-\varepsilon^2 G(q, p, p) + \varepsilon h(q, p, Y, Z)) \tag{A.11}$$

$$\widehat{g}(q, p, Y, Z) = -A(q, Y) - \psi_\mu \widehat{f}^\mu(q, p, Y, Z) - \psi_{\mu\nu} p^\mu p^\nu + \varepsilon j'(q, p, Y, Z), \tag{A.12}$$

and G is defined in (5.6). The point is that f, g are defined as functions of q_t (and q, Y, Y_t) on a space (X_Γ) with ε-dependent norm, but it is more covenient here to think of them as functions of $\dot{q} = \varepsilon^{-1} q_t$, on a space $(B_{\Gamma'}$, $\Gamma' = 8c(\Sigma)\Gamma)$ with fixed norm. Since ι is a linear isometry, f, g are smooth, bounded, Lipschitz, etc. if and only if $\widehat{f}, \widehat{g}$ are.

Proposition A.4. *There exists $\varepsilon_* = O(1/\sqrt{\Gamma})$ such that for all $\varepsilon \in (0, \varepsilon_*)$, $\widehat{f} : B_\Gamma \to \mathbb{R}^{4n}$ and $\widehat{g} : B_\Gamma \to H^2$ are smooth.*

Proof. Since $\gamma^{\mu\lambda}(q)$, $\psi_{\lambda\nu}(q)$ are smooth and K is compact, there exists $\varepsilon_* > 0$ such that the matrix $M(q, Y)$ is uniformly invertible on B_Γ for all $\varepsilon \in (0, \varepsilon_*)$. The components of M^{-1} are rational in $M^\mu{}_\nu$ with denominator $\det M$, which is bounded away from 0 on B_Γ. Hence, M^{-1} is smooth on B_Γ, and the proposition follows immediately from the Leibniz rule and Proposition A.3. □

Since $\widehat{f}$ and $\widehat{g}$ are smooth, they are certainly continuously differentiable. To complete the proof of Proposition 5.3, it remains to show that their differentials are bounded.

Proposition A.5. *For all $\varepsilon \in (0, \varepsilon_*)$, the derivatives $\mathrm{d}\widehat{g} : B_\Gamma \to \mathcal{L}(B, H^2)$ and $\mathrm{d}\widehat{f} : B_\Gamma \to \mathcal{L}(B, \mathbb{R}^{4n})$ are bounded, independent of ε.*

Proof. This is established by estimating the operator norm of the derivatives termwise. For example, the first term of j',

$$J(q, p, Y, Z) = 2p^\mu(\psi_\mu \cdot Z)\psi \tag{A.13}$$

has derivative

$$dJ_{(q,p,Y,Z)} : (\hat{q}, \hat{p}, \hat{Y}, \hat{Z}) \mapsto 2\hat{q}^\nu p^\mu[(\psi_{\nu\mu} \cdot Z)\psi + (\psi_\mu \cdot Z)\psi_\nu] + 2\hat{p}^\mu(\psi_\mu \cdot Z)\psi + 2p^\mu(\psi_\mu \cdot \hat{Z})\psi, \tag{A.14}$$

and so, by the algebra property of H^2, for all $(q, p, Y, Z) \in B_\Gamma$,

$$\begin{aligned} \|dJ_{(q,p,Y,Z)}(\hat{q}, \hat{p}, \hat{Y}, \hat{Z})\|_2 &\leq C(K)\{|\hat{q}||p|\|Z\|_2 + |\hat{p}|\|Z\|_2 + \|Z\|_2|p|\} \\ &\leq C(K)\{2\Gamma + \Gamma^2\}\|(\hat{q}, \hat{p}, \hat{Y}, \hat{Z})\|_B \end{aligned} \tag{A.15}$$

where $C(K) > 0$ is a constant depending only on K. Hence, for all $(q, p, Y, Z) \in B_\Gamma$,

$$\|dJ_{(p,q,Y,Z)}\|_{\mathcal{L}(B,H^2)} \leq C(K)\{2\Gamma + \Gamma^2\}. \tag{A.16}$$

The other terms of j', h and A are handled similarly. To bound $\mathrm{d}M^{-1}$ we bound $\mathrm{d}M$ and appeal to the Leibniz rule and uniform invertibility of M. Boundedness of the differentials of functions depending only on q and p is immediate by compactness and finiteness of dimension. □

Finally, we turn to Proposition 5.5, which is equivalent to:

Proposition A.6. *The differentials of the maps* $\widehat{f} : B_\Gamma \to \mathbb{R}^{4n}$ $\widehat{g} : B_\Gamma \to H^2$ *extend to maps* $\mathrm{d}\widehat{f}^{ext} : B_\Gamma \to \mathcal{L}(\mathbb{R}^{4n} \times \mathbb{R}^{4n} \times H^1 \times L^2, \mathbb{R}^{4n})$ *and* $\mathrm{d}\widehat{g}^{ext}g : B_\Gamma \to \mathcal{L}(\mathbb{R}^{4n} \times \mathbb{R}^{4n} \times H^1 \times L^2, L^2)$ *bounded by* Λ_f, Λ_g *respectively.*

Proof. This follows from explicit termwise computation. For example, the last term of h is

$$m_\mu(q, Y, Z) = \langle (|Z|^2 - |Y_x|^2 - |Y_y|^2)Y, \psi_\mu \rangle$$

whose differential at (q, Y, Z) is the linear map $\mathbb{R}^{4n} \times H^3 \times H^2 \to \mathbb{R}$

$$\begin{aligned} \mathrm{d}m_\mu : (\hat{q}, \hat{Y}, \hat{Z}) \;\mapsto\; & \langle (|Z|^2 - |Y_x|^2 - |Y_y|^2)Y, \psi_{\mu\nu} \rangle \hat{q}^\nu + \langle (|Z|^2 - |Y_x|^2 - |Y_y|^2)\psi_\mu, \hat{Y} \rangle \\ & -2\langle (Y \cdot \psi_\mu)Y_x, \hat{Y}_x \rangle - 2\langle (Y \cdot \psi_\mu)Y_y, \hat{Y}_y \rangle + 2\langle (Y \cdot \psi_\mu)Z, \hat{Z} \rangle \end{aligned}$$

which clearly extends to a bounded linear map $\mathbb{R}^{4n} \times H^1 \times L^2 \to \mathbb{R}$. □

Acknowledgements

This work was financially supported by EPSRC. It began as a collaboration with Mark Haskins, and I gratefully acknowledge many hours of valuable discussion of the problem with him. I also benefited from conversations with David Stuart and Matthew Daws. The work was largely completed while I was on sabbatical leave at the Isaac Newton Institute for Mathematical Sciences in Cambridge.

References

[1] P. Bizon, T. Chmaj and Z. Tabor, "Formation of singularities for equivariant (2 + 1)-dimensional wave maps into the 2-sphere" *Nonlinearity* **14** (2001) 1041-53.

[2] S. K. Donaldson and P.B. Kronheimer, *The Geometry of Four-Manifolds* (Clarendon Press, Oxford, UK, 1990) pp420-425.

[3] J. Eells and L. Lemaire, "A Report on Harmonic Maps" *Bull. London Math. Soc.* **10** (1978) 1-68.

[4] J. Eells and J.C. Wood, "Restrictions on harmonic maps of surfaces" *Topology* **15** (1976) 263-266.

[5] M. Haskins and J.M. Speight, "The geodesic approximation for lump dynamics and coercivity of the Hessian for harmonic maps" *J. Math. Phys.* **44** (2003) 3470-3494.

[6] N.J. Hitchin, G.B. Segal and R.S. Ward, *Integrable Systems: Twistors, Loop Groups and Riemann Surfaces* (Clarendon Press, Oxford, UK, 1999) pp17-26.

[7] F. John, *Partial Differential Equations* (Springer-Verlag, Berlin, Germany, 1978) pp112-115.

[8] J. Krieger, W. Schlag and D. Tataru, "Renormalization and blow up for charge one equivariant critical wave maps" *Invent. Math.* **171** (2008), 543–615.

[9] D.F. Lawden, *Elliptic Functions and Applications* (Springer-Verlag, Berlin, Germany, 1989) pp226-7.

[10] R.A. Leese, "Low energy scattering of solitons in the $\mathbb{C}P^1$ model" *Nucl. Phys.* **B344** (1990) 33-72.

[11] R.A. Leese, M. Peyrard and W.J. Zakrzewski, "Soliton stability in the $O(3)$ σ-model in $(2+1)$ dimensions" *Nonlinearity* **3** (1990) 387412.

[12] A. Lichnerowicz, "Applications harmoniques et variétés kähleriennes" *Symp. Math. Bologna* **3** (1970) 341-402.

[13] J.M. Linhart and L.A. Sadun, "Fast and slow blowup in the S^2 σ model and the $(4+1)$-dimensional Yang-Mills model" *Nonlinearity* **15** (2002) 219-38.

[14] J.A. McGlade and J.M. Speight, "Slow equivariant lump dynamics on the two-sphere" *Nonlinearity* **19** (2006) 441-452.

[15] N.S. Manton, "A remark on the scattering of BPS monopoles" *Phys. Lett.* **110B** (1982) 54-6.

[16] N.S. Manton and P.M. Sutcliffe, *Topological Solitons* (Cambridge University Press, Cambridge, UK, 2004).

[17] Y.N. Ovchinnikov and I.M. Sigal, "On collapse of wave maps" *Physica D* **240** (2011) 1311-1324.

[18] P. Raphaël and I. Rodnianski, "Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang-Mills problems" *Publ. Math. Inst. Hautes Études Sci.* (2012), 1–122.

[19] I. Rodnianski and J. Sterbenz, "On the formation of singularities in the critical $O(3)$ σ-model" *Ann. of Math. (2)* **172** (2010) 187242.

[20] L. Sadun and J.M. Speight, "Geodesic incompleteness in the $\mathbb{C}P^1$ model on a compact Riemann surface" *Lett. Math. Phys.* **43** (1998) 329-34.

[21] J. Shatah and M. Struwe, *Geometric Wave Equations* (American Mathematical Society, Providence RI, USA, 1998).

[22] J.M. Speight, "Low energy dynamics of a $\mathbb{C}P^1$ lump on the sphere" *J. Math. Phys.* **36** (1995) 796-813.

[23] J.M. Speight, "Lump dynamics in the $\mathbb{C}P^1$ model on the torus" *Commun. Math. Phys.* **194** (1998) 513-539.

[24] J.M. Speight, "The L^2 geometry of spaces of harmonic maps $S^2 \to S^2$ and $\mathbb{R}P^2 \to \mathbb{R}P^2$" to appear in *J. Geom. Phys.*

[25] D. Stuart, "Dynamics of abelian Higgs vortices in the near Bogomolny regime" *Commun. Math. Phys.* **159** (1994) 51-91.

[26] D. Stuart, "The geodesic approximation for the Yang-Mills-Higgs equations" *Commun. Math. Phys.* **166** (1994) 149-90.

[27] D. Stuart, "Analysis of the adiabatic limit for solitons in classical field theory" *Proc. R. Soc. A* **463** (2007) 2753-2781.

[28] H. Urakawa, *Calculus of Variations and Harmonic Maps* (AMS, Providence RI, USA, 1993).

[29] R.S. Ward, "Slowly moving lumps in the $\mathbb{C}P^1$ model in $(2+1)$ dimensions" *Phys. Lett.* **158B** (1985) 424-8.