Quasi-local energy and isometric embedding

Karsten Gimre

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2016

ABSTRACT

Quasi-local energy and isometric embedding

Karsten Gimre

In this thesis, we consider the recent definition of gravitational energy at the quasilocal level provided by Mu-Tao Wang and Shing-Tung Yau. Their definition poses a variational question predicated on isometric embedding of Riemannian surfaces into the Minkowski space $\mathbb{R}^{3,1}$; as such, there is a naturally associated Euler-Lagrange equation, which is a fourth-order system of partial differential equations for the embedding functions. We prove a perturbation result for solutions of this Euler-Lagrange equation.

Table of Contents

List of Tables

Acknowledgments

I am greatly indebted to my advisor, Mu-Tao Wang, for incredibly much helpful insight and guidance. Over my years as a graduate student, I am also grateful to Simon Brendle, Toti Daskalopoulos, Mikhael Gromov, Richard Hamilton, Sergiu Klainerman, Duong Phong, and Ovidiu Savin for teaching me most of what I know about geometry and analysis.

I'm also very grateful to Connor Mooney, Ye-Kai Wang, Alex Waldron, Pei-Ken Hung, Sebastien Picard, Rahul Krishna, Zhijie Huang, Xuan Wu, and Dan Rubin for helpful conversations related to this thesis, and for many other illuminating conversations about other mathematical topics. I'd also like to thank my friends Noah, Stephane, Rob, Ioan, Joao, Andrea, Jordan, Thomas, Vivek, Natasha, Petr, Andrey, and Sara for making my time as a grad student so enjoyable.

For the last twelve years, my undergraduate advisor Christine Guenther has been a consistent source of support, as well as a valued collaborator.

Lastly, I'd like to thank my family for much support.

Part I

Motivation and Background

Chapter 1

Introduction

In this thesis we will study quasilocal gravitational energy. This will be in the physical context of classical general relativity and in the mathematical context of geometric partial differential equations.

In 1982, Penrose [28] asked for a definition of quasilocal gravitational energy, as a quantity to be associated to two-dimensional surfaces within $(3 + 1)$-dimensional spacetimes. In the years since, there have been many proposed definitions, each of which satisfy various properties and have various uses. For example, the Hawking mass, although not considered to be a physically viable energy, plays a fundamental role in the proof of the global nonlinear stability of the Minkowski space [7], as well as in the proof of the Riemannian Penrose inequality [18].

In recent years, there has been a new definition proposed by M.-T. Wang and S.-T. Yau [36]. The precise definition will be reviewed in the following chapter. For now, we will only say that the definition is based on the minimization of an integral quantity over an open class of solutions of the isometric embedding problem.

There are now many reasons to consider the Wang-Yau definition as a 'good' definition of quasilocal energy, namely (all true under some restrictions):

- it is positive [36]
- it is zero for any surface in Minkowski space [36]
- it recovers classical [1; 2] definitions of total energy [37; 4]

- it can be extended to definitions of quasilocal linear momentum, angular momentum, and center of mass [6]

We remark that the integral quantity in the Wang-Yau definition is closely related to a well-known quasilocal energy proposed by J. Brown and J. York [3], with the significant distinction that there is no minimization present in the Brown-York definition, where there is a uniquely determined isometric embedding for which the corresponding integral quantity is evaluated. As an artifact of this, the Brown-York energy will usually not vanish for surfaces in Minkowski space. Moreover, the value of the Brown-York energy depends on the identification of an initial data set, which could be considered to be an artificial piece of information. There is a modification due to C.-C. Liu and S.-T. Yau [20] of the Brown-York energy which does not require this extra data, although for the same reason as for the Brown-York energy, it will also usually be nonzero for surfaces in the Minkowski space.

The minimization in the Wang-Yau definition is thus key for its viability as a physically plausible definition. It is this minimization problem that we study in this thesis.

In particular, we will prove a perturbation result for the solvability of the local minimization. That is to say, if there is a surface in a spacetime whose submanifold geometry is close to the submanifold geometry of another surface in another spacetime for which the minimum is attained, then the minimum is also attained for the nearby surface. This will be stated more precisely in the third chapter. As will be discussed there, this is a generalization of a result of P. Miao, L.-F. Tam and N. Xie [22].

There are a few natural contexts for a result of this type. One is to compare the gravitational energy contained by a membrane under small perturbations of the ambient spacetime metric – for example, during the course of the propagation of a gravitational wave. Another is to study the asymptotic geometry of initial data sets with general asymptotics. The now-classical positive energy theorem of Schoen-Yau and Witten [30; 31; 38] gives the nonnegativity of the ADM and Bondi energies for asymptotically flat spacetimes [1; 2]. The validity of the positive energy theorem requires certain decay rates for the metrics and second fundamental forms of initial data sets. A perturbation result of this type could be useful for studying more general asymptotics, by viewing large coordinate spheres as small perturbations of spheres in Minkowski space, or of spheres in even more general model spaces. This would involve a generalization of the limiting process involved in the aforementioned recovery of the ADM and Bondi energies by the Wang-Yau energy [37; 4].

At the very end of this thesis, we collect some of the notations used, for ease of reference.

Chapter 2

Wang-Yau quasilocal energy

Let $(N^{3,1}, \overline{g})$ be a spacetime, i.e. a $(3+1)$-dimensional Lorentzian manifold. We restrict our attention to spacetimes satisfying the dominant energy condition, which is to say that, for any timelike vector v, $T(v,v) \geq 0$ and $T(v,\cdot)$ is non-spacelike; here T is the 2-tensor given by $\mathrm{Ric} - \frac{1}{2}R\overline{g}$ where Ric and R are respectively the Ricci and scalar curvatures of $\overline{g}$. This is a standard restriction to make [34]. In particular, it is automatically satisfied for any vacuum spacetime (one which is Ricci-flat).

Let $\iota : S^2 \hookrightarrow (N, \overline{g})$ be an embedding inducing the Riemannian metric σ on S^2, which has spacelike mean curvature vector H. Let α_H denote the 1-form on S^2 given by

$$\alpha_H(X) = -\overline{g}\left(\overline{\nabla}_{\iota_* X}\frac{H}{|H|}, J\right) \qquad (X \in TS^2)$$

where J is the future-directed unit timelike normal vector field along $\iota(S^2)$ which is orthogonal to H. Altogether we refer to $(\sigma, |H|, \alpha_H)$ as the *physical data* of our problem, and as shorthand we also denote this triple by Σ. Crucially, Σ contains information about the submanifold geometry of $\iota(S^2) \subset N$.

We are concerned with isometric embeddings of (S^2, σ) into the Minkowski space $\mathbb{R}^{3,1}$. The fundamental solution of the 'Weyl problem,' due to Nirenberg [26] and Pogorelov [29], is as follows:

Nirenberg-Pogorelov theorem. *Let σ be a positively curved Riemannian metric on S^2, of Hölder regularity $C^{k,\alpha}$ (with $k \geq 4$ and $0 < \alpha < 1$). There is an isometric embedding of (S^2, σ) into $\mathbb{R}^3$ equipped with its Euclidean metric, uniquely determined up to rigid motions of $\mathbb{R}^3$, and it is of regularity $C^{k,\alpha}$.*

Due to the uniqueness of the resulting isometric embedding, there is a well-defined second fundamental form and mean curvature of the embedded surface.

The Nirenberg-Pogorelov result has seen very little improvement since its proof over a half-century ago. The underlying reason is that any significant extension would require understanding of a non-elliptic problem. P. Guan and Y. Li [13] have provided an extension to nonnegative curvature, but as this leads to a degenerate elliptic problem, the regularity of the resulting isometric embedding is not sufficient for many purposes, in particular for well-definedness of the second fundamental form of the embedded surface, which will be crucial for the definition of the Wang-Yau energy. Essentially nothing is known for isometric embedding of higher-dimensional spheres (even in the case of positive curvature), for the same reason of lack of ellipticity.

Recall that we are concerned with isometric embeddings of (S^2, σ) into $\mathbb{R}^{3,1}$. The link to the Nirenberg-Pogorelov theorem is provided by the elementary observation that $(\tau, \widehat{X}) : S^2 \to \mathbb{R}^{3,1}$ is an isometric embedding of σ if and only if $\widehat{X} : S^2 \to \mathbb{R}^3$ is an isometric embedding of (the necessarily Riemannian metric) $\sigma + d\tau \otimes d\tau$. Hence, provided that we can *choose* a function τ such that $\sigma + d\tau \otimes d\tau$ has positive curvature, we obtain from the Nirenberg-Pogorelov theorem a wide variety of isometric embeddings of σ into $\mathbb{R}^{3,1}$, parametrized by all such choices of τ.

Now take as given some smooth physical data $\Sigma = (\sigma, |H|, \alpha_H)$. Suppose we can choose $\tau \in C^\infty(S^2)$ such that $\widehat{\sigma} \equiv \sigma + d\tau \otimes d\tau$ has positive curvature. According to the Nirenberg-Pogorelov theorem, let $\widehat{X} : S^2 \to \mathbb{R}^3$ be an isometric embedding of $\widehat{\sigma}$, with $\widehat{h}$ and $\widehat{H}$ the second fundamental form and mean curvature of the embedded surface. Define the functional

$$E(\Sigma, \tau) = \int_{S^2} \widehat{H}\, d\widehat{\mu} - \int_{S^2} \Big[\cosh\theta |H| \sqrt{1 + |\nabla\tau|^2} + \theta\Delta\tau - \alpha_H(\nabla\tau) \Big]\, d\mu \tag{2.1}$$

where θ denotes the expression

$$\sinh^{-1} \frac{-\Delta\tau}{|H|\sqrt{1 + |\nabla\tau|^2}}$$

and $d\mu$ (resp. $d\widehat{\mu}$) is the volume form determined by σ (resp. $\widehat{\sigma}$).

We will not motivate the definition 2.1 here, only noting that it is based on an analysis of the Hamiltonian formulation of general relativity, the starting point being the Hawking-Horowitz action for the Einstein equations on a manifold with boundary; this provides a key difference with the classical energies of Arnowitt-Deser-Misner and Bondi, which were based on the classical Einstein-Hilbert action on a manifold without boundary.

Remark. *Although they will not be relevant in this thesis, here we recall the Brown-York and Liu-Yau definitions of quasilocal energy referred to in the introduction. Given the additional data of a spacelike hypersurface M^3 in N with boundary $\iota(S^2)$, we define the Brown-York energy by*

$$E(\Sigma, M^3) = \int_{S^2} (H_0 - H')\, d\mu$$

where here H' is the mean curvature of $\iota(S^2) \subset M$ while H_0 is the mean curvature of an isometric embedding of (S^2, σ) into $\mathbb{R}^3$; according to the Nirenberg-Pogorelov theorem, we should restrict to the setting of σ having positive curvature. We note that this definition depends in a definite way on the selection of M, as this selection directly influences H. The Liu-Yau energy does not depend on a choice of M, and is defined by

$$E(\Sigma) = \int_{S^2} (H_0 - |H|)\, d\mu$$

for H_0 defined as for the Brown-York energy. The Liu-Yau energy is generally nonnegative, but is strictly positive for most surfaces in the Minkowski space.

In order to prove the nonnegativity of $E(\Sigma, \tau)$, Wang and Yau add, in addition to the prior restrictions of H being spacelike and $\widehat{\sigma} \equiv \sigma + d\tau \otimes d\tau$ having positive curvature, the additional restriction that there exist a spacelike hypersurface M of N, with boundary $\iota(S^2)$, such that the boundary value problem for the 'Jang equation'

$$\left(g^{ij} - \frac{f^i f^j}{1 + |Df|^2}\right)\left(\frac{D_i D_j f}{\sqrt{1 + |Df|^2}} - p_{ij}\right) = 0 \qquad \text{on } M$$

$$f = \tau \qquad \text{on } \iota(S^2) = \partial M$$

is solvable; here g_{ij} denotes the metric induced on M by $\overline{g}$, and p_{ij} denotes the second fundamental form of $M \subset N$. We make the final restriction that

$$-\sqrt{1 + |\nabla\tau|^2}\langle H, e_3'\rangle - \alpha_{e_3'}(\nabla\tau) > 0,$$

where

$$e_3' = \sqrt{1 + \frac{(e_3 f)^2}{1 + |\nabla\tau|^2}}\, e_3 - \frac{e_3 f}{\sqrt{1 + |\nabla\tau|^2}}\, e_4,$$

for e_3, e_4 the oriented orthonormal frame of the normal bundle of $\iota(S^2) \subset N$ given by e_3 being the outward normal of $\iota(S^2) \subset M$. The one-form $\alpha_{e_3'}$ is defined by

$$\alpha_{e_3'}(X) = \overline{g}\left(\overline{\nabla}_{\iota_* X} e_3', e_4'\right) \qquad (X \in TS^2)$$

for e_4' defined so that e_3', e_4' is an oriented orthonormal frame of the normal bundle of $\iota(S^2) \subset N$.

In the context of the restrictions stated in the previous paragraph, Wang-Yau proved the nonnegativity of $E(\Sigma, \tau)$. As remarked in the introduction, these restrictions restrict the admissible τ to an open subset of $C^\infty(S^2)$.

Theorem 1 (Wang-Yau)**.** *Under the above restrictions,* $E(\Sigma, \tau) \geq 0$.

In order to recover a pure functional of Σ that does not depend on a choice of τ (equivalently, on a choice of an isometric embedding of (S^2, σ) into $\mathbb{R}^{3,1}$), we can infimize over the restricted class of τ:

$$E(\Sigma) = \inf_{\tau} E(\Sigma, \tau). \tag{2.2}$$

It is clear that the infimized quantity inherits the same nonnegativity as $E(\Sigma, \cdot)$.

As indicated in the introduction, when $N = \mathbb{R}^{3,1}$ and τ is the time component of ι, it can be calculated that $E(\Sigma, \tau) = 0$. Given the nonnegativity of $E(\Sigma, \cdot)$, it is then clear that $E(\Sigma) = 0$ as well.

Furthermore, in the setting of asymptotically flat spacetimes, Wang-Yau and Chen-Wang-Yau have shown that $E(\Sigma, 0)$ recovers the ADM and Bondi energies at spatial at null infinity. Lastly, the definition 2.1 can be modified so that the cosmological de Sitter or anti-de Sitter spacetimes are the inherent reference, rather than $\mathbb{R}^{3,1}$. This is to say that the modified energies are still nonnegative, but vanish when the surface is embedded in de Sitter or anti-de Sitter.

In their paper introducing 2.1 and 2.2, Wang-Yau calculated the Euler-Lagrange equation of $E(\Sigma, \cdot)$ to be

$$-(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd})\frac{\nabla_a\nabla_b\tau}{\sqrt{1+|\nabla\tau|^2}} + \operatorname{div}\left(\frac{\cosh\theta|H|\nabla\tau}{\sqrt{1+|\nabla\tau|^2}} - \nabla\theta - \alpha_H\right) = 0. \tag{2.3}$$

As written, this is not a partial differential equation for τ since the terms $\widehat{H}$ and $\widehat{h}_{cd}$ depend on τ, but only through the solution of isometric embedding problem, not through partial derivatives. The Euler-Lagrange equation can be interpreted as a system of partial differential equations for the full isometric embedding $(\tau, \widehat{X})$ by coupling it with the isometric embedding equation

$$-d\tau \otimes d\tau + \langle d\widehat{X}, d\widehat{X}\rangle = \sigma,$$

and we take this to be implicit whenever we refer to 2.3. We remark that this system is second order in $\widehat{X}$, through the terms $\widehat{H}$ and $\widehat{h}_{cd}$, and fourth order in τ, through the term $\operatorname{div}\nabla\theta$, θ

containing two derivatives of τ. Furthermore, at top order in τ the equation is biharmonic, the principal symbol being

$$\xi \mapsto \frac{-|\xi|^4}{\sqrt{|H|^2(1+|\nabla\tau|^2)+(\Delta\tau)^2}}.$$

Part II

Local solvability

Chapter 3

Introduction

Suppose we are given some physical data $\Sigma_0 = (\sigma, |H|, \alpha_H)$, together with a critical point τ_0 of $E(\Sigma_0, \cdot)$. Given some new physical data Σ which is close to Σ_0 (as measured by closeness in normed spaces of the constituent parts of Σ and Σ_0) we would like to find a critical point τ of $E(\Sigma, \cdot)$. This is naturally posed as an implicit function theorem-type result.

In a relatively recent paper, P. Miao, L.-F. Tam and N. Xie [22] studied this problem in the case that τ_0 is the zero function, with Σ and Σ_0 being physical data which are literally near to each other as the submanifold data of some joint ambient spacetime, the 'nearness' measured by requiring Σ to be the graph of small functions over the normal bundle of Σ_0, via the exponential function of the ambient spacetime. In this context the linearized optimal isometric embedding equation is a fourth order scalar elliptic *partial* differential equation, and can be solved by essentially Hodge-theoretic means. The implicit function theorem is then applied to the graphing functions in the normal bundle. A key simplification offered by the restriction that τ_0 is the zero function is that the linearized Euler-Lagrange operator becomes particularly simple:

$$v \mapsto -(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd})\nabla_a\nabla_b v + \operatorname{div}(|H|\nabla v) + \Delta\frac{\Delta v}{|H|}.$$

As remarked, this is a partial differential operator in v, and in particular there is no genuine presence of the isometric embedding problem, except in as much as it is present in the point being linearized around.

Here we plan to remove these restrictions. Specifically, we allow:

1. τ_0 to be nearly any critical point, with

2. Σ and Σ' not required to be the submanifold data of the same ambient spacetime.

The first of these is achieved by understanding the linearized Euler-Lagrange operator in a greater generality than Miao-Tam-Xie. As the linearized Euler-Lagrange operator is very algebraically complicated when linearized at a general point, we do this by studying the second variation of the energy functional rather than by studying the Euler-Lagrange operator itself. The fact that we do not require Σ and Σ' to be submanifolds of the same spacetime is a consequence of the fact that we will apply the implicit function theorem where the parameters are the physical data itself, rather than graphing functions in the normal bundle, as they are for Miao-Tam-Xie. Our main theorem is the following

Theorem 2. *Let $(\sigma, |H|, \alpha_H)$ be some smooth physical data for which there is a smooth critical point τ_0, such that $|H| < |H_{\tau_0}|$. If $\Sigma' = (\sigma', |H|', \alpha'_H)$ is sufficiently close to $(\sigma, |H|, \alpha_H)$ in $C^{3,\alpha} \times C^{2,\alpha} \times C^{1,\alpha}$, then the Euler-Lagrange equation for $E(\Sigma', \cdot)$ has a solution τ which is close in $C^{4,\alpha}$ to τ_0. If Σ' is smooth then τ is as well.*

The only significant restriction here is $|H| < |H_{\tau_0}|$, which is also a restriction for Miao-Tam-Xie. We recall [4]

Theorem 3 (Chen-Wang-Yau)**.** *If τ_0 is a critical point of $E(\Sigma, \cdot)$ such that $|H| < |H_{\tau_0}|$, then τ_0 is a local minimum of $E(\Sigma, \cdot)$.*

We will revisit the proof of this result in proving theorem 2. It will be seen that $|H| < |H_{\tau_0}|$ is a particularly natural condition under which to have the local minimization, although it is not clear to what extent it is necessary. At any rate, it is satisfied in a number of cases, for example for large coordinate spheres in asymptotically Schwarzschild initial data sets in the case of $\tau_0 = 0$.

In theorem 2, we require the data we perturb around to be smooth in order that the linearized operator is Fredholm. It is not clear that this would be the case if we perturbed around data in $C^{3,\alpha} \times C^{2,\alpha} \times C^{1,\alpha}$, since in this case the pseudodifferential part of the linearization is of the same order as the partial differential operator part, and the ellipticity is unclear. This will be explained further in the next chapter.

We will prove theorem 2 in the next chapter. In the last chapter, we will prove the following

Theorem 4. *Let $(\sigma, |H|, \alpha_H)$ be some smooth physical data for which there is a smooth critical point*

τ_0, such that $|H| < |H_{\tau_0}|$. If Σ' is sufficiently close to Σ in the smooth topology, then the solution of the Euler-Lagrange equation for $E(\Sigma', \cdot)$ can be constructed from the Newton-Nash-Moser iteration.

The solution in the setting of theorem 4 is already known to exist from theorem 2. Thus theorem 4 only provides an alternative construction of the solution. The convergence of the Newton-Nash-Moser iteration does not follow from the applicability of the implicit function theorem, since the local solvability does not provide the necessary quantitative estimates for the linearized operator at nearby points. Thus the last chapter is entirely devoted to proving these estimates. For the application of these estimates to the convergence of the iteration, we refer to the general literature on the Nash-Moser theorem [11; 16; 24; 25; 32; 39; 40], and in particular [14].

We remark that although the implicit function theorem in Banach spaces was applicable, we need the Nash-Moser modification of the Newton method since, as indicated above, in studying the linearized problem, we needed the point we perturbed around to be of higher regularity than the perturbation we take in the inversion of the linearization.

Chapter 4

Linearized equation

In this chapter we establish the general solvability of the linearized optimal isometric embedding equation. In order to establish the linearized operator as an isomorphism between appropriate function spaces, we will identify exactly the kernel and cokernel of the linearization. Then we will establish the invertibility in the smooth category, and will end by finding the general regularity of the inverted linearization, by invoking Schauder estimates.

We recall that the optimal isometric embedding equation is given by

$$P(\sigma, |H|, \alpha_H; \tau) \equiv -(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd})\frac{\nabla_a\nabla_b\tau}{\sqrt{1+|\nabla\tau|^2}} + \mathrm{div}\left(\frac{\cosh\theta|H|\nabla\tau}{\sqrt{1+|\nabla\tau|^2}} - \nabla\theta - \alpha_H\right) = 0,$$

where θ is shorthand for the expression

$$\sinh^{-1}\frac{-\Delta\tau}{|H|\sqrt{1+|\nabla\tau|^2}}.$$

We also recall that $\widehat{\sigma} = \sigma + d\tau\otimes d\tau$, so that according to the Nirenberg-Pogorelov theorem there is an isometric embedding $\widehat{X}$ of $\widehat{\sigma}$ into $\mathbb{R}^3$ if the curvature of $\widehat{\sigma}$, considered as a Riemannian metric, has positive curvature. So if σ is of regularity $C^{k,\alpha}$ and τ is of regularity $C^{l,\alpha}$, then $\widehat{\sigma}$ is of regularity $C^{\min\{k,l-1\},\alpha}$. According to the regularity statement in the Nirenberg-Pogorelov theorem, $\widehat{X}$ is then also of regularity $C^{\min\{k,l-1\},\alpha}$. Then we also see

$$\begin{aligned}
\widehat{\Gamma}^c_{ab} &= \frac{1}{2}\widehat{\sigma}^{cd}(\partial_a\widehat{\sigma}_{bd} + \partial_b\widehat{\sigma}_{ad} - \partial_d\widehat{\sigma}_{ab}) \\
&\in C^{\min\{k,l-1\}-1,\alpha} \\
\widehat{h}_{ab} &= \widehat{\nabla}_a\widehat{\nabla}_b\widehat{X} \\
&= \partial_a\partial_b\widehat{X} - \widehat{\Gamma}^c_{ab}\partial_c\widehat{X} \\
&\in C^{\min\{k,l-1\}-2,\alpha} \\
\widehat{H} &= \widehat{\sigma}^{ab}\widehat{h}_{ab} \\
&\in C^{\min\{k,l-1\}-2,\alpha}.
\end{aligned}$$

Going through the constituent parts of $P(\tau)$, there is:

$$\begin{aligned}
\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd} &\in C^{\min\{k,l-1\}-2,\alpha} \\
\nabla_a\nabla_b\tau &= \partial_a\partial_b\tau - \Gamma^c_{ab}\partial_c\tau \\
&\in C^{\min\{k-1,l-2\},\alpha} \\
|\nabla\tau|^2 &= \sigma^{ab}\partial_a\tau\partial_b\tau \\
&\in C^{\min\{k,l-1\},\alpha} \\
\frac{1}{\sqrt{1+|\nabla\tau|^2}} &\in C^{\min\{k,l-1\},\alpha} \\
(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd})\frac{\nabla_a\nabla_b\tau}{\sqrt{1+|\nabla\tau|^2}} &\in C^{\min\{k-2,l-3\},\alpha}.
\end{aligned}$$

For the remaining terms, we also need to place regularity assumptions on $|H|$ and α_H; we say

$|H| \in C^{m,\alpha}$ and $\alpha_H \in C^{n,\alpha}$. Then:

$$\begin{aligned}
\Delta\tau &= \sigma^{ab}\nabla_a\nabla_b\tau \\
&\in C^{\min\{k-1,l-2\},\alpha} \\
\theta &\in C^{\min\{k-1,l-2,m\},\alpha} \\
\cosh\theta|H| &\in C^{\min\{k-1,l-2,m\},\alpha} \\
\frac{\cosh\theta|H|\nabla\tau}{\sqrt{1+|\nabla\tau|^2}} &\in C^{\min\{k-1,l-2,m\},\alpha} \\
\operatorname{div}\frac{\cosh\theta|H|\nabla\tau}{\sqrt{1+|\nabla\tau|^2}} &= \sigma^{ab}\left(\partial_a\omega_b - \Gamma^c_{ab}\omega_c\right) \\
&\in C^{\min\{k-2,l-3,m-1\},\alpha} \\
\nabla_a\nabla_b\theta &= \partial_a\partial_b\theta - \Gamma^c_{ab}\partial_c\theta \\
&\in C^{\min\{k-3,l-4,m-2\},\alpha} \\
\Delta\theta &\in C^{\min\{k-3,l-4,m-2\},\alpha} \\
\operatorname{div}\alpha_H &= \sigma^{ab}(\partial_a(\alpha_H)_b - \Gamma^c_{ab}(\alpha_H)_c) \\
&\in C^{\min\{k-1,n-1\},\alpha}.
\end{aligned}$$

So, all put together, $P(\sigma, |H|, \alpha_H; \tau)$ is a map

$$C^{k,\alpha} \times C^{m,\alpha} \times C^{n,\alpha} \times C^{l,\alpha} \to C^{\min\{k-3,l-4,m-2,n-1\},\alpha}$$

and is C^1.

4.1 Quantitative comparison estimate

To recall the context: let $\Sigma = (\sigma, |H|, \alpha_H)$ be fixed smooth physical data, together with a smooth critical point τ_0 of $E(\Sigma, \cdot)$. We are then necessarily equipped with an embedding $\widehat{X} : S^2 \hookrightarrow \mathbb{R}^3$ so that $X \equiv (\tau_0, \widehat{X}) : S^2 \hookrightarrow \mathbb{R}^{3,1}$ is an isometric embedding of σ. We denote the mean curvature vector of $(\tau_0, \widehat{X})$ by H_{τ_0}.

Assumption/Restriction. $|H| < |H_{\tau_0}|$.

As remarked in the introduction, this is a fundamental assumption for us. We will see its relevance in the following calculations.

We consider also the physical data $\Sigma_{\tau_0} = (\sigma, |H_{\tau_0}|, \alpha_{H_{\tau_0}})$ arising from the submanifold data of $X(S^2) \subset \mathbb{R}^{3,1}$. A key calculation of P.-N. Chen, M.-T. Wang and S.-T. Yau [5] is that

$$\begin{aligned} A &\equiv E(\Sigma,\tau) - E(\Sigma_{\tau_0},\tau) \\ &= \int \sqrt{1+|\nabla\tau|^2}\cdot f\left(\frac{\Delta\tau}{\sqrt{1+|\nabla\tau|^2}}\right) d\mu \\ &\qquad - \int \frac{\nabla\tau_0\cdot\nabla\tau}{\sqrt{1+|\nabla\tau_0|^2}}\left(\sqrt{|H_{\tau_0}|^2+\frac{(\Delta\tau_0)^2}{1+|\nabla\tau_0|^2}} - \sqrt{|H|^2+\frac{(\Delta\tau_0)^2}{1+|\nabla\tau_0|^2}}\right) d\mu \end{aligned}$$

where f is the single-variable function (to be precise, it associates to each point of S^2 a single-variable function)

$$f(x) = \sqrt{|H_{\tau_0}|^2 + x^2} - \sqrt{|H|^2+x^2} - x\left(\sinh^{-1}\frac{x}{|H_{\tau_0}|} - \sinh^{-1}\frac{x}{|H|} - \sinh^{-1}\frac{x_0}{|H_{\tau_0}|} + \sinh^{-1}\frac{x_0}{|H|}\right)$$

with

$$x_0 = \frac{\Delta\tau_0}{\sqrt{1+|\nabla\tau_0|^2}}.$$

The reason for taking the difference $E(\Sigma,\tau) - E(\Sigma_{\tau_0},\tau)$ is that the Riemannian metrics underlying both Σ and Σ_{τ_0} are the same; hence exactly the same isometric embedding problem is considered in both $E(\Sigma,\tau)$ and $E(\Sigma_{\tau_0},\tau)$, and in taking the difference there are purely terms to do with τ and the physical datas. If we did not have this cancellation, we would have to find a way to compare the mean curvatures for different isometric embedding problems, and this may be rather involved.

Under the assumption $|H| < |H_{\tau_0}|$ as above, it is a straightforward matter of single-variable calculus to say that f achieves its global minimum at x_0. We calculate

$$f''(x_0) = \frac{1}{\sqrt{|H|^2 + x_0^2}} - \frac{1}{\sqrt{|H_{\tau_0}|^2 + x_0^2}}$$

and we apply Taylor's theorem to estimate A from below by its quadratic Taylor polynomial. The presence of the quadratic term will allow us to identify the kernel of the linearization $DP(0,0,0,v)$. We carry this out as follows: let $v \in C^\infty$ be arbitrary, and take $\tau = \tau_0 + \varepsilon v$ to say, denoting

$$x = \frac{\Delta\tau}{\sqrt{1+|\nabla\tau|^2}},$$

and making repeated use of Taylor's theorem, that

$$x - x_0 = \left(\frac{\Delta v}{\sqrt{1+|\nabla\tau_0|^2}} - \frac{\Delta\tau_0\langle\nabla\tau_0,\nabla v\rangle}{(1+|\nabla\tau_0|^2)^{3/2}}\right)\varepsilon + \mathrm{O}(\varepsilon^2),$$

and then

$$f(x) = \sqrt{|H_{\tau_0}|^2 + x_0^2} - \sqrt{|H|^2 + x_0^2} + \frac{\varepsilon^2}{2}\left(\frac{1}{\sqrt{|H|^2+x_0^2}} - \frac{1}{\sqrt{|H_{\tau_0}|^2+x_0^2}}\right)\left(\frac{\Delta v}{\sqrt{1+|\nabla\tau_0|^2}} - \frac{\Delta\tau_0\langle\nabla\tau_0,\nabla v\rangle}{(1+|\nabla\tau_0|^2)^{3/2}}\right)^2 + \mathrm{O}(\varepsilon^3).$$

Substituting this into the Chen-Wang-Yau calculation for $A \equiv E(\Sigma,\tau) - E(\Sigma_{\tau_0},\tau)$, we find

$$\begin{aligned}
A = \int \Bigg[& \sqrt{1+|\nabla\tau|^2}\left(\sqrt{|H_{\tau_0}|^2+x_0^2} - \sqrt{|H|^2+x_0^2}\right) \\
& + \frac{\varepsilon^2\sqrt{1+|\nabla\tau|^2}}{2(1+|\nabla\tau_0|^2)}\left(\frac{1}{\sqrt{|H|^2+x_0^2}} - \frac{1}{\sqrt{|H_{\tau_0}|^2+x_0^2}}\right)\left(\Delta v - \frac{\Delta\tau_0\langle\nabla\tau_0,\nabla v\rangle}{1+|\nabla\tau_0|^2}\right)^2 \\
& - \frac{\nabla\tau_0\cdot\nabla\tau}{\sqrt{1+|\nabla\tau_0|^2}}\left(\sqrt{|H_{\tau_0}|^2+x_0^2} - \sqrt{|H|^2+x_0^2}\right)\Bigg]\, d\mu \\
& + \mathrm{O}(\varepsilon^3) \\
\geq E(\Sigma,\tau_0) + \int & \frac{\varepsilon^2}{4\sqrt{1+|\nabla\tau_0|^2}}\left(\frac{1}{\sqrt{|H|^2+x_0^2}} - \frac{1}{\sqrt{|H_{\tau_0}|^2+x_0^2}}\right)\left(\Delta v - \frac{\Delta\tau_0\langle\nabla\tau_0,\nabla v\rangle}{1+|\nabla\tau_0|^2}\right)^2 d\mu.
\end{aligned}$$

The estimation above is as follows: in the second block term of the integrand, we use the Taylor estimation

$$\sqrt{1+|\nabla\tau|^2} = \sqrt{1+|\nabla\tau_0|^2} + \mathrm{O}(\varepsilon)$$

to replace the entire term by the integrand appearing in the last line above. The estimation of the first and third terms to give the $E(\Sigma,\tau_0)$ in the last line is exactly as in Chen-Wang-Yau, which is to say that we use Cauchy-Schwarz to estimate the factor

$$\sqrt{1+|\nabla\tau|^2} - \frac{\nabla\tau_0\cdot\nabla\tau}{\sqrt{1+|\nabla\tau_0|^2}} \geq \frac{1}{\sqrt{1+|\nabla\tau_0|^2}},$$

with

$$E(\Sigma,\tau_0) = \int \frac{1}{\sqrt{1+|\nabla\tau_0|^2}}\left(\sqrt{|H_{\tau_0}|^2+x_0^2} - \sqrt{|H|^2+x_0^2}\right) d\mu,$$

which can be recognized after rewriting the expression for $E(\Sigma,\tau_0)$ using the optimal isometric embedding equations.

For later reference we state the above as the following lemma:

Quantitative Comparison Lemma. *Let $\Sigma = (\sigma, |H|, \alpha_H)$ be physical data with τ_0 a critical point of $E(\Sigma,\cdot)$. Let $\Sigma_{\tau_0} = (\sigma, |H_{\tau_0}|, \alpha_{H_{\tau_0}})$ be the physical data of the isometrically embedded surface*

$X(\Sigma) \subset \mathbb{R}^{3,1}$. Suppose that $|H| < |H_{\tau_0}|$. Then

$$\delta^2 E(\Sigma,\tau)(v)\Big|_{\tau=\tau_0} \geq \delta^2 E(\Sigma_{\tau_0},\tau)(v)\Big|_{\tau=\tau_0} + \int \Phi\left(\Delta v - \frac{\Delta\tau_0\langle\nabla\tau_0,\nabla v\rangle}{1+|\nabla\tau_0|^2}\right)^2 d\mu$$

for some strictly positive expression Φ.

Note that the statement of the lemma is just an infinitesimal version (at second order) of the previous estimate, where for simplicity we only refer to

$$\frac{1}{2\sqrt{1+|\nabla\tau_0|^2}}\left(\frac{1}{\sqrt{|H|^2+x_0^2}} - \frac{1}{\sqrt{|H_{\tau_0}|^2+x_0^2}}\right)$$

as some positive expression.

As remarked in the introduction, we have $E(\Sigma_{\tau_0},\tau_0) = 0$ and $E(\Sigma_{\tau_0},\tau) \geq 0$ for τ close to τ_0. Hence we have

$$\delta^2 E(\Sigma_{\tau_0},\tau)(v)\Big|_{\tau=\tau_0} \geq 0$$

and so it is clear from the lemma that we also have

$$\delta^2 E(\Sigma,\tau)(v)\Big|_{\tau=\tau_0} \geq 0,$$

with equality only if

$$\Delta v - \frac{\Delta\tau_0\langle\nabla\tau_0,\nabla v\rangle}{1+|\nabla\tau_0|^2} = 0.$$

By the strong maximum principle, this is the case only when v is constant.

4.2 Solvability

Fix some smooth physical data $\Sigma \in C^\infty \times C^\infty \times C^\infty$, and some $\tau \in C^\infty$. We consider the structure of the linearization $DP_{\sigma,|H|,\alpha_H;\tau}(0,0,0;v)$, which for convenience we denote by $DP_\tau(v)$. We do *not* suppose that τ is a critical point of $E(\Sigma,\cdot)$. The linearization is

$$DP_\tau(v) = -\delta\Big(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd}\Big)(v)\frac{\nabla_a\nabla_b\tau}{\sqrt{1+|\nabla\tau|^2}} + E_\tau v.$$

Here the first term on the left denotes the linearization of the mapping

$$\tau \mapsto \sigma + d\tau\otimes d\tau \equiv \widehat{\sigma} \mapsto (\widehat{\sigma},\widehat{X}) \mapsto \widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd},$$

which is notably *not* a partial differential operator in τ; this failure is caused by factoring through the solution of the isometric embedding problem at $\widehat{\sigma} \mapsto (\tau, \widehat{X})$. The second term of $DP_\tau(v)$, $E_\tau v$, is a fourth order linear partial differential operator with (smooth) coefficients which are nonlinearly dependent upon our choice of the physical data and τ.

Lemma. *$DP_\tau : C^{l,\alpha} \to C^{l-4,\alpha}$ is a Fredholm map for any choice of $l \gg 1$.*

We recall that a 'Fredholm' map between Banach spaces is one with closed range and finite-dimensional kernel. We note that, as a fourth order linear elliptic partial differential operator with smooth coefficients, the E_τ part of DP_τ is already known to be a Fredholm map.

So we investigate more closely the term

$$-\delta\Big(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd}\Big)(v)\frac{\nabla_a\nabla_b\tau}{\sqrt{1+|\nabla\tau|^2}},$$

which we denote $P_\tau \circ Q_\tau \circ R_\tau \circ S_\tau(v)$ where

- $S_\tau : C^{l,\alpha} \to \mathrm{Sym}^2_{l-1,\alpha}\, S^2$ is the linearization of $\tau \mapsto \widehat{\sigma}$, i.e. is defined by

$$v \mapsto d\tau \circ dv + dv \circ d\tau$$

- $R_\tau : \mathrm{Sym}^2_{l-1,\alpha}\, S^2 \to C^{l-1,\alpha}(S^2, \mathbb{R}^3)$ is a solution operator for the isometric embedding problem $S^2 \hookrightarrow \mathbb{R}^3$ when linearized at $\widehat{X}$, i.e. is a right inverse for $Y \mapsto 2\langle d\widehat{X}, dY\rangle$
- $Q_\tau : C^{l-1,\alpha}(S^2, \mathbb{R}^3) \to \mathrm{Sym}_2^{l-3,\alpha}\, S^2$ is the linearization at $\widehat{X}$ of the map sending an immersion $Z : S^2 \hookrightarrow \mathbb{R}^3$ to $\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd}$
- $P_\tau : \mathrm{Sym}_2^{l-3,\alpha}(S^2) \to C^{l-3,\alpha}(S^2)$ is defined by

$$T^{ab} \mapsto \frac{T^{ab}\nabla_a\nabla_b\tau}{\sqrt{1+|\nabla\tau|^2}}.$$

We remark that all four of the above are bounded linear maps (bounded with respect to the indicated Banach spaces with their standard norms). Only this statement for R_τ is non-obvious, and it is a result of Nirenberg's [26]. It also follows from the later, more general, result of Douglis and Nirenberg [9]. In the context of this thesis, it can be seen from the estimates in the next chapter.

Remark. *In the above, we used* $\mathrm{Sym}^2_{l-1,\alpha} S^2$ *to denote the space of* $C^{l-1,\alpha}$ $(2,0)$*-tensors and* $\mathrm{Sym}_2^{l-1,\alpha} S^2$ *to denote the space of* $C^{l-1,\alpha}$ $(0,2)$*-tensors.*

Thus $P_\tau \circ Q_\tau \circ R_\tau \circ S_\tau$ is a compact operator, and since E_τ is a Fredholm operator, we find that DP_τ is Fredholm as claimed.

Remark. *A naive application of the Schauder estimates may seem to say that* R_τ *should map into* $C^{l,\alpha}$*, as it is the inversion of a first-order operator and according to the usual Schauder estimates, we should gain one derivative. However we need to be careful in choosing our spaces in such a way that* R_τ *is the inverse of an elliptic operator. As currently stated, this cannot be the case since on the symbol level, infinitesimal variations coming from the isometry group of* $\mathbb{R}^3$ *will lead to degenerate directions. However, we do have an elliptic operator once we project (via the normal vector field) every infinitesimal variation of* $\widehat{X}$ *to the tangent bundle, while modding out in* Sym^2 *by the subbundle generated by* $\widehat{h}$*. Relative to these spaces, we do gain the expected one derivative; this is to say that the projection of the inversion to the tangent bundle is in* $C^{l,\alpha}$*. The recovery of the full infinitesimal variation from its projection to the tangent bundle involves both zeroth order terms in the metric variation and first order terms in the tangent-direction variations, either of which, by themselves, would be enough to place our final variation in* $C^{l-1,\alpha}$*. In the next section, where we derive estimates for the linearized equation, we will use the weighted Douglis-Nirenberg Schauder estimates to avoid these projections, which would cause algebraically messy terms in our analysis.*

This remark may be clearer in the context of the estimates in the next chapter.

Remark. *Here we note why we restricted ourselves to perturbing around smooth data. If we were to suppose* τ *to be a* $C^{l,\alpha}$ *critical point rather than a smooth one, then* $\widehat{\sigma} = \sigma + d\tau^2$ *would be* $C^{l-1,\alpha}$ *and then* $\widehat{X}$*, according to the nonlinear regularity result of Nirenberg [26], would also be* $C^{l-1,\alpha}$*. Then according to the regularity of the linearized problem,* R_τ *would only map into* $C^{l-2,\alpha}$ *and altogether* $P_\tau \circ Q_\tau \circ R_\tau \circ S_\tau$ *would be a bounded linear map* $C^{l,\alpha} \to C^{l-4,\alpha}$*. Thus it is not immediately clear that* DP_τ *would still be Fredholm, as the ellipticity of this now-top-order term* $P_\tau \circ Q_\tau \circ R_\tau \circ S_\tau$ *is by no means clear.*

It may become clearer that this difficulty is one that can be circumvented by noting that (as we have implicitly indicated above) the structure of the operator DP_τ *is such that low-regularity* *variations of high-regularity τ fare strictly better than low-regularity variations of low-regularity τ. This is an artifact of the 'loss of regularity' for the isometric embedding problem.*

4.2.1 Invertibility of DP_τ.

We note that the principal symbol of DP_τ is

$$\xi \mapsto \frac{-|\xi|^4}{\sqrt{|H|^2(1+|\nabla\tau|^2)+(\Delta\tau)^2}},$$

which is evidently homotopic, through symbols which are isomorphisms, to the bilaplacian. So the index of DP_τ is zero.

Given that DP_τ is Fredholm, in order to identify its range we only need to find the kernel of its adjoint DP_τ^*.

Remark. *We recall the general principle that, given a C^2 action functional $F(u)$ and a solution u of its Euler-Lagrange equation, the linearization of the Euler-Lagrange equation of F at u is automatically self-adjoint, as is seen by computing both sides of*

$$\frac{\partial^2}{\partial s\,\partial t}\Big|_{s=t=0} F(u+sv+tw) = \frac{\partial^2}{\partial t\,\partial s}\Big|_{s=t=0} F(u+sv+tw).$$

However, in our case it is not clear that the integrand of the Wang-Yau functional is C^2 as a map between the appropriate Hölder spaces. This is another reason for restricting to smooth data, as it is C^2 as a map from C^∞ to C^∞. The essential difficulty comes from the presence in the integrand of the non-local term $\widehat{H}$.

Now in the case that τ_0 is a critical point, we use the quantiative comparison lemma from the previous section, together with the strong maximum principle, to conclude that v is constant if $v \in \ker DP_{\tau_0}^*$.

Let τ_0 remain a critical point. By making $\tau \in C^{l,\alpha}$ sufficiently close to τ_0 in the Banach norm topology, we use the index theory for Fredholm operators to say

$$\dim\ker DP_\tau = \dim\ker DP_{\tau_0} = 1.$$

Now as DP_τ has index zero, we also have $\dim\operatorname{coker} DP_\tau = 1$. So $\operatorname{coker} DP_\tau$ consists of exactly the constant functions. As $\ker DP_\tau$ also has dimension one, it is also exactly the space of constant functions. Here we recall that the constants are evidently in the kernel since our operator is invariant under time translations, and constants are in the cokernel due to the fact that we have

$$\int (\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd}) \frac{\nabla_a \nabla_b \tau}{\sqrt{1+|\nabla\tau|^2}}\, d\mu = \int (\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd})\widehat{\nabla}_a \widehat{\nabla}_b \tau d\widehat{\mu} = 0$$

due to integration by parts and the Codazzi equations; hence the image of the Euler-Lagrange operator, the other part of which is a divergence term, always has average value zero.

We have shown the following:

Theorem 5. *Given any smooth physical data Σ and any critical point τ_0 of $E(\Sigma, \cdot)$, as long as τ is sufficiently close to τ_0 in the $C^{l,\alpha}$ topology and Σ' is sufficiently close to Σ in the C^∞ topology, the linearization $DP^{\Sigma'}_\tau$ is a bijection $C^{l,\alpha}_- \to C^{l-4,\alpha}_-$.*

Here $C^{p,\alpha}_-$ denotes the $C^{p,\alpha}$ functions with average value zero.

4.3 Local solvability

According to the derivative counts at the beginning of this chapter, the Euler-Lagrange operator is a C^1 map

$$C^{3,\alpha} \times C^{2,\alpha} \times C^{1,\alpha} \times C^{4,\alpha}_- \ni (\sigma, |H|, \alpha_H; \tau) \mapsto P(\sigma, |H|, \alpha_H; \tau) \in C^{\alpha}_-.$$

If we have smooth physical data $(\sigma, |H|, \alpha_H)$ for which there is a smooth critical point τ_0 (such that $|H| < |H_{\tau_0}|$), then in the last section we have shown that the linearization of P at $(\sigma, |H|, \alpha_H; \tau_0)$ is an isomorphism. By the implicit function theorem, we can conclude the following theorem

Theorem 6. *In the above context, if $(\sigma', |H|', \alpha'_H)$ is physical data which is sufficiently close to $(\sigma, |H|, \alpha_H)$ in $C^{3,\alpha} \times C^{2,\alpha} \times C^{1,\alpha}$, then the Euler-Lagrange equation at $(\sigma', |H|', \alpha'_H)$ is solvable with a solution which is close in $C^{4,\alpha}$ to τ_0.*

4.4 Regularity

Here we show the last part of theorem 2, that the solution of Euler-Lagrange equation is smooth if the physical data is smooth. As a result of the implicit function theorem, we have a classical solution τ of the equation which is $C^{4,\alpha}$. Suppose we have shown that τ is $C^{k,\alpha}$ for some $k \geq 4$. We will show that it is then $C^{k+1,\alpha}$.

Since $\tau \in C^{k,\alpha}$ we have

$$\widehat{\sigma} = \sigma + d\tau \otimes d\tau \in C^{k-1,\alpha},$$

so that according to Nirenberg's regularity result for the Weyl problem, $\widehat{\sigma}$ can be isometrically embedded in $\mathbb{R}^3$ by a $C^{k-1,\alpha}$ embedding. The mean curvature and second fundamental form of the embedding is then $C^{k-3,\alpha}$. So

$$(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd})\frac{\nabla_a \nabla_b \tau}{\sqrt{1+|\nabla\tau|^2}} \in C^{k-3,\alpha}.$$

It is clear that $\theta \in C^{k-2,\alpha}$, and so referring back to the Euler-Lagrange equation 2.3 to see that

$$\Delta\theta \in C^{k-3,\alpha}.$$

Then $\theta \in C^{k-1,\alpha}$, which by definition of θ is to say

$$\frac{\Delta\tau}{|H|\sqrt{1+|\nabla\tau|^2}} \in C^{k-1,\alpha}.$$

Since $\nabla\tau \in C^{k-1,\alpha}$, we have $\Delta\tau \in C^{k-1,\alpha}$ and we conclude $\tau \in C^{k+1,\alpha}$.

By induction, we find $\tau \in C^{\infty}$.

Chapter 5

Estimates for the linearized equation

5.1 Introduction

In this chapter, we will derive tame estimates for the inverse of the linearization, so that the Newton-Nash-Moser method will converge to the solution of the perturbed problem [14]. Although we were able to apply the implicit function theorem in Banach spaces, in following the standard Newton method we will lose derivatives, since we had to perturb around points with more regularity than the perturbations we apply to them.

Our estimates will apply in a sufficiently small C^∞ neighborhood of a smooth critical point. We will shrink this neighborhood as necessary without mention, as we will do so quite frequently; most of the following estimates would not hold without doing so.

Although the Newton-Nash-Moser method is not necessary in providing solutions of the Euler-Lagrange equation, there are two reasons to justify the estimates of this chapter. The first is that the Nash-Moser iteration is very rapidly convergent, so for the sake of numerical approximations it may be more useful than the Banach space setup. Secondly, it is worth understanding the nature of the linearized equation even when not at critical points. Such an understanding may very well be useful, for example, in taking the limit of the Wang-Yau energy at the spatial infinity of initial data sets with very general asymptotic data. One may hope, for example, to understand a more general version of the linearization at infinity that takes place within the regime of ADM asymptotics [37].

This chapter is organized as follows: in the third section we will derive estimates for the linearization of the Weyl problem. The key, as in most of our estimates, is to obtain an explicit ('tame') dependence on parameters in order that the Newton-Nash-Moser method will converge. In the fourth section, we derive our result of the third section to estimate the pseudodifferential term in the linearization of the Euler-Lagrange equation. In the last section, we deal with the entire linearized equation.

5.2 Notations

We essentially follow the setup of Nirenberg's work [26]. We work exclusively with functions and tensors on S^2, on which we fix, once and for all, two coordinate (disk) regions R_1 and R_2, compactly contained within $S^2 \setminus \{S\}$ and $S^2 \setminus \{N\}$, which intersect in an annular region around the equator. Further restrict so that both of their coordinate regions in $\mathbb{R}^2$ are the unit disk. Thus any smooth tensor on S^2, when represented through the coordinates associated to R_1 or R_2, will be C^∞ with bounds on all derivatives.

We use these regions to define Hölder spaces of tensors. We prefer to define these using fixed coordinate regions instead of using a choice of Riemannian metric, since in all of the following work we will have many different metrics in use at any given time, so that it is convenient to have norms defined independently of any of them.

We say that a tensor T on S^2 is $C^{k,\alpha}$ if its coordinate representation in both R_1 and R_2 is $C^{k,\alpha}$. We define its corresponding (Banach) Hölder norm

$$|T|_{k,\alpha} = |T|_{C^{k,\alpha}(R_1)} + |T|_{C^{k,\alpha}(R_2)}.$$

The norm of the tensor as defined on the unit disk in $\mathbb{R}^2$, on the right hand side of the above definition, is defined by summing the norms of all of its components. Similarly, we also define the (Fréchet) C^k norms

$$|T|_k = |T|_{C^k(R_1)} + |T|_{C^k(R_2)}$$

with the norms of the tensors on the unit disk in $\mathbb{R}^2$ being defined as before.

We will make repeated use of the following well-known properties of the Hölder norms (see for instance the appendix of [16]).

Lemma. *1.* $|uv|_{k,\alpha} \le C_k\Big(|u|_0|v|_{k,\alpha} + |u|_{k,\alpha}|v|_0\Big)$

2. $|\frac{1}{u}|_{k,\alpha} \le C_k\Big(|u|_{k,\alpha} + 1\Big)$ *where* C_k *depends on* $|u-1|_0$

3. $|A^{ij} - \delta_{ij}|_{k,\alpha} \leq C_k\Big(|A_{ij}|_{k,\alpha} + 1\Big)$ *where* C_k *depends on* $|A_{ij} - \delta_{ij}|_0$ *(*A^{ij} *denotes the inverse of* A_{ij}*)*

4. $|u \circ v|_{k,\alpha} \leq C_k\Big(|f|_{k,\alpha} + |g|_{k,\alpha} + 1\Big)$ *where* C_k *depends on* $|f|_{1,\alpha}$ *and* $|g|_{1,\alpha}$

5. $|u|_{k,\alpha}|v|_{l,\alpha} \leq C\Big(|u|_{k-a,\alpha}|v|_{l+a,\alpha} + |u|_{k+a,\alpha}|v|_{l-a,\alpha}\Big)$.

5.3 Weyl problem estimates

In this section we focus our attention entirely on aspects of the Weyl problem, i.e. the isometric embedding of positively curved metrics on S^2 into $\mathbb{R}^3$. For convenience of notation, we drop the usual hatted notation, i.e. instead of $\widehat{\sigma}$ we say σ. We let σ_0 denote a fixed smooth positively curved metric on S^2, with $\widehat{X}_0 : S^2 \to \mathbb{R}^3$ a (smooth) isometric embedding of σ_0.

5.3.1 Estimate on second fundamental form

According to Nirenberg [26] (bottom of page 353), there is $\overline{K}$ dependent upon $\widehat{X}_0(S^2)$ such that if σ is a positively curved Riemannian metric on S^2 with $|\sigma - \sigma_0|_{2,\alpha} < \frac{1}{4}\overline{K}^{-2}$, then

$$|\widehat{X} - \widehat{X}_0|_{2,\alpha} < 2\overline{K}|\sigma - \sigma_0|_{2,\alpha},$$

where $\widehat{X} : S^2 \to \mathbb{R}^3$ is an isometric embedding of σ (of course, in general a rigid motion of $\mathbb{R}^3$ is implicitly understood here).

Translate both $\widehat{X}$ and $\widehat{X}_0$ in $\mathbb{R}^3$ so that the largest balls contained in $\widehat{X}(S^2)$ and $\widehat{X}_0(S^2)$ are both centered at the origin. Denoting $\rho = \frac{1}{2}|\widehat{X}|^2$, the 'Darboux equation' [8; 26] which is satisfied reads

$$F(\rho, \sigma) \equiv \frac{\det(\mathrm{Hess}\,\rho - \sigma)}{|\sigma|} - K(\sigma)\Big(2\rho - \sigma^{ij}\rho_i\rho_j\Big) = 0.$$

(Of course, also $F(\rho_0, \sigma_0) = 0$ where $\rho_0 = \frac{1}{2}|\widehat{X}_0|^2$.) An immediate consequence of Nirenberg's estimate is the following

Lemma (Nirenberg)**.** *There is* $\overline{K}$ *dependent upon* $\widehat{X}_0(S^2)$ *so that if* σ *is a positively curved Riemannian metric on* S^2 *with* $|\sigma - \sigma_0|_{2,\alpha} < \frac{1}{4}\overline{K}^{-2}$*, then*

$$|\rho - \rho_0|_{2,\alpha} < C|\sigma - \sigma_0|_{2,\alpha}.$$

We require $|\sigma - \sigma_0|_{2,\alpha}$ to be sufficiently small (denote it less than δ) so that the Darboux equation is uniformly elliptic (the ellipticity constant will depend on σ_0).

Lemma. *If $|\sigma - \sigma_0|_{2,\alpha}$ is sufficiently small, as above, then*

$$|\rho|_{k,\alpha} \leq C_k\Big(1 + |\sigma|_{k,\alpha}\Big)$$

for all $k \geq 2$.

Proof. From the triangle inequality, this is an immediate consequence of the previous lemma in the case $k = 2$. We then argue inductively. Fix a k and suppose that the stated estimate of the lemma is true. We differentiate the equation $F(\rho, \sigma) = 0$ to find

$$(\mathrm{Hess}\,\rho - \sigma)^{-1}\,\mathrm{Hess}(\partial\rho) = \frac{\partial\Big(K(\sigma)(2\rho - \sigma^{ij}\rho_i\rho_j)\Big)}{K(\sigma)\Big(2\rho - \sigma^{ij}\rho_i\rho_j\Big)}$$

where ∂ denotes any (first order) coordinate derivative. Fix a neighborhood to make the RHS bounded in the C^α norm, as we will apply the bound (see pg 157 of [14])

$$|\partial\rho|_{k,\alpha} \leq C\Big(|\mathrm{RHS}|_{k-2,\alpha} + |(\mathrm{Hess}\,\rho - \sigma)^{-1}|_{k-2,\alpha}|\mathrm{RHS}|_{0,\alpha}\Big);$$

If we did not fix this neighborhood, the induction would give an estimate of the form $|\rho|_{k,\alpha} \leq C_k(1 + |\sigma|_{3k,\alpha})$, which would not give a tame estimate. At any rate, we use the inductive hypothesis to say

$$|\log K(\sigma)(2\rho - \sigma^{ij}\rho_i\rho_j)|_{k-1,\alpha} \leq C_k\Big(1 + |\sigma_{ij}|_{k+1,\alpha} + |\rho|_{k,\alpha}\Big) \leq C_k\Big(1 + |\sigma_{ij}|_{k+1,\alpha}\Big)$$

Also

$$|(\mathrm{Hess}\,\rho - \sigma)^{-1}|_{k-2,\alpha} \leq C\Big(1 + |\rho|_{k,\alpha} + |\sigma|_{k-2,\alpha}\Big) \leq C\Big(1 + |\sigma|_{k,\alpha}\Big),$$

as we have already fixed a neighborhood to have uniform ellipticity of the differentiated equation. Then appealing to the aforementioned bound, we find

$$|\partial\rho|_{k,\alpha} \leq C\Big(1 + |\sigma|_{k+1,\alpha}\Big),$$

finishing the induction. □

Lemma. *If* $|\sigma - \sigma_0|_{2,\alpha}$ *is sufficiently small (as above) then*

$$|\widehat{h}_{ij}|_{k,\alpha} \leq C_k\Big(1 + |\sigma|_{k+2,\alpha}\Big)$$

for all $k \geq 0$.

Proof. As in Nirenberg [26], we can express the second derivatives of $\widehat{X}$ in terms of the second derivatives of ρ through the linear equations

$$\begin{aligned}
\langle \widehat{X}, \widehat{X}_{uu}\rangle &= \rho_{uu} - \sigma_{uu} \\
\langle \widehat{X}_u, \widehat{X}_{uu}\rangle &= \frac{1}{2}\partial_u \sigma - uu \\
\langle \widehat{X}_v, \widehat{X}_{uu}\rangle &= \partial_u \sigma_{uv} - \frac{1}{2}\partial_v \sigma_{uu}
\end{aligned}$$

with analogous equations for $\widehat{X}_{uv}$ and $\widehat{X}_{vv}$. Via the previous lemma and the estimate for $\widehat{X}$ stated at the beginning of the section, we have control over the coefficients and right hand side of the above system, and we obtain the same estimates for $\widehat{X}$ that we had for ρ in the previous lemma, namely that if $|\sigma - \sigma_0|_{2,\alpha}$ is sufficiently small, then

$$|\widehat{X}|_{k,\alpha} \leq C_k\Big(1 + |\sigma|_{k,\alpha}\Big).$$

From this the current lemma immediately follows. □

5.3.2 Estimate for the linearized problem

Now we take the following context: let $\widehat{X}_0 : S^2 \to \mathbb{R}^3$ be an isometric embedding of the positively curved $C^{k,\alpha}$ metric $\widehat{\sigma}_0$ and we let $\widehat{\sigma}$ be a metric on S^2 which is $C^{k,\alpha}$ close to $\widehat{\sigma}_0$. Let $\widehat{X} : S^2 \to \mathbb{R}^3$ be an isometric embedding of $\widehat{\sigma}$. We want to study the linearized isometric embedding equation

$$2\langle d\widehat{X}, dY\rangle = \delta\widehat{\sigma}$$

for a function $Y : S^2 \to \mathbb{R}^3$. As written, this does not appear to be an elliptic system. To recognize it as such, we decompose Y into a part tangential and normal to $\widehat{X}(S^2)$:

$$Y = A^\dagger + B\widehat{\nu}$$

for A a 1-form on $\widehat{X}(S^2)$ (with $A^\dagger$ the vector field dual to A with respect to $\widehat{\sigma}$) and B a function on S^2 ($\widehat{\nu}$ being the normal vector field of $\widehat{X}(S^2) \subset \mathbb{R}^3$). Rewritten in terms of this decomposition, the linearized isometric embedding equation becomes

$$\widehat{\nabla}_i A_j + \widehat{\nabla}_j A_i + 2B\widehat{h}_{ij} = \delta\widehat{\sigma}_{ij}.$$

With respect to any local coordinate system, we claim that this is an elliptic system for (A_1, A_2, B) in the Douglis-Nirenberg sense. To do so we recall the definition (not stated in its full generality):

Definition (Douglis-Nirenberg ellipticity)**.** *Consider the system of linear partial differential equations*

$$\ell_{i1}(x, D)u_1 + \ell_{i2}(x, D)u_2 + \ell_{i3}(x, D)u_3 = f_i.$$

Choose integers $s_1, s_2, s_3, t_1, t_2, t_3$ so that $\ell_{ij}(x, D)$ has order at most $s_i + t_j$. Denote by ℓ'_{ij} the $s_i + t_j$-order terms of ℓ_{ij}. The indicated system is elliptic if $\det \ell'_{ij}(x, \xi) \neq 0$ *for all* $\xi \neq 0$.

For us, taking $(u_1, u_2, u_3) = (A_1, A_2, B)$, we choose $s_1 = s_2 = s_3 = t_3 = 0$ and $t_1 = t_2 = 1$. Thus

$$\ell'(x, \xi) = \begin{pmatrix} 2\xi_1 & 0 & 2\widehat{h}_{11} \\ \xi_2 & \xi_1 & 2\widehat{h}_{12} \\ 0 & 2\xi_2 & 2\widehat{h}_{22} \end{pmatrix}$$

with determinant

$$\det \ell'(x, \xi) = 4 \det(\widehat{h}_{ij}) \begin{pmatrix} \xi_1 & \xi_2 \end{pmatrix} \begin{pmatrix} \widehat{h}_{11} & \widehat{h}_{12} \\ \widehat{h}_{12} & \widehat{h}_{22} \end{pmatrix}^{-1} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

Since we assumed that $\widehat{\sigma}$ has positive curvature, it is clear that this is a strictly positive expression for $\xi \neq 0$. We now state the Douglis-Nirenberg estimate (essentially directly quoted from [9], in the special case of s_i, t_j as above.

Theorem 7 (Douglis-Nirenberg)**.** *Write*

$$\ell(x, D) = \begin{pmatrix} a_{11,0} + a_{11,x}D_x + a_{11,y}D_y & a_{12,0} + a_{12,x}D_x + a_{12,y}D_y & a_{13,0} \\ a_{21,0} + a_{21,x}D_x + a_{21,y}D_y & a_{22,0} + a_{22,x}D_x + a_{22,y}D_y & a_{23,0} \\ a_{31,0} + a_{31,x}D_x + a_{31,y}D_y & a_{32,0} + a_{32,x}D_x + a_{32,y}D_y & a_{33,0} \end{pmatrix}.$$

We suppose that $a_{i1,0}, a_{i2,0} \in C^{k+1,\alpha}$, that $a_{i1,x}, a_{i1,y}, a_{i2,x}, a_{i2,y} \in C^{k,\alpha}$, and that $a_{i3,0} \in C^{k,\alpha}$ with a bound K_1 on the corresponding norms. We let K_2 denote the constant of ellipticity, i.e. so that

$$\det \ell'(x, \xi) \geq K_2|\xi|^2.$$

Then there are the estimates

$$|u_i|_{k+1,\alpha} \le K\Big(|u_1|_0 + |u_2|_0 + |\ell(x,D)u|_{k+1,\alpha}\Big)$$

fir $i = 1,2,3,$ *where* $K = K(\alpha, K_1, K_2)$.

For us, we have $a_{i1,0}$ and $a_{i2,0}$ given by the Christoffel symbols $\widehat{\Gamma}^k_{ij}$, $a_{i1,k}$ and $a_{i2,k}$ constant, $a_{i3,0}$ given by the components of $\widehat{h}_{ij}$, and f_i given by the components of $\delta\widehat{\sigma}_{ij}$.

From here on, we fix the notation (for an arbitrary positively curved metric $\widehat{\sigma}$ on S^2)

$$\ell_{\widehat{\sigma}} u \equiv \partial_i A_j + \partial_j A_i - 2\widehat{\Gamma}^k_{ij} A_k + 2B\widehat{h}_{ij}, \tag{5.1}$$

where for convenience we use u to denote (A_1, A_2, B). We will also use $\circ\widehat{\Gamma}^k_{ij}$ to denote the Christoffel symbols of $\widehat{\sigma}_0$. (Recall that we always work in a fixed coordinate system as described at the beginning of the chapter.)

Given the infinitesimal rigidity of convex surfaces in $\mathbb{R}^3$, if we normalize our solutions of the linearized isomeric embedding equation so that Y vanishes at a given point, then we recover the estimate

$$|u|_{k+1,\alpha} \le K|\ell_{\widehat{\sigma}} u|_{k+1,\alpha}$$

still for $K = K(\alpha, K_1, K_2)$. For our eventual estimates, we need to understand better the dependence of K upon K_1.

Lemma. $|u|_{k+1,\alpha} \le C_k\Big(|\ell_{\widehat{\sigma}}(x,D)u|_{k+1,\alpha} + (|\widehat{\Gamma}|_{k+1,\alpha} + |\widehat{h}|_{k+1,\alpha})|\ell_{\widehat{\sigma}}(x,D)u|_{1,\alpha}\Big)$ *where* C_k *depends on* k, $\widehat{\sigma}_0$, *and the neighborhood of* $\widehat{\sigma}_0$ *we are restricting* $\widehat{\sigma}$ *to.*

Proof. We have from the Douglis-Nirenberg estimate

$$|u|_{k+1,\alpha} \le C|\ell_{\widehat{\sigma}_0} u|_{k+1,\alpha}.$$

We emphasize that this estimate is at $\widehat{\sigma}_0$, and so we do not care about the dependence of this constant on anything. From observing

$$\ell_{\widehat{\sigma}}(x,D)u - \ell_{\widehat{\sigma}_0}(x,D)u = 2(\widehat{\Gamma}^k_{ij} - \circ\widehat{\Gamma}^k_{ij})A_k + 2B(\widehat{h}_{ij} - \widehat{h}^0_{ij})$$

we have

$$|\ell_{\widehat{\sigma}}(x,D)u - \ell_{\widehat{\sigma}_0}(x,D)u|_{1,\alpha} \le 4\Big(|\widehat{\Gamma}^k_{ij} - \circ\widehat{\Gamma}^k_{ij}|_{1,\alpha} + |\widehat{h}_{ij} - \widehat{h}^0_{ij}|_{1,\alpha}\Big)|u|_{1,\alpha}.$$

So

$$\begin{aligned}|u|_{1,\alpha} &\leq C|\ell_{\widehat{\sigma}_0} u|_{1,\alpha} \\ &\leq C|\ell_{\widehat{\sigma}} u|_{1,\alpha} + 4\Big(|\widehat{\Gamma}^k_{ij} - \circ\widehat{\Gamma}^k_{ij}|_{1,\alpha} + |\widehat{h}_{ij} - \widehat{h}^0_{ij}|_{1,\alpha}\Big)|u|_{1,\alpha}.\end{aligned}$$

So upon supposing that $\widehat{\sigma}$ is sufficiently close to $\widehat{\sigma}_0$, we have

$$|u|_{1,\alpha} \leq C|\ell_{\widehat{\sigma}} u|_{1,\alpha}.$$

This provides the $k = 0$ case of the proposed lemma.

We then proceed by induction: suppose the lemma is true as stated, for some fixed choice of k. We apply the inductive hypothesis for u replaced by one of its first (coordinate) derivatives, which we denote u^d. In so doing, we say

$$\ell_{\widehat{\sigma}}(u^d) = (\ell_{\widehat{\sigma}} u)^d - (\partial\widehat{\Gamma})u + 2(\partial\widehat{h})u$$

with the effect of

$$\begin{aligned}|\ell_{\widehat{\sigma}} u^d|_{k+1,\alpha} &\leq |\ell_{\widehat{\sigma}} u|_{k+2,\alpha} + \Big(|\partial\widehat{\Gamma}|_{k+1,\alpha} + |\partial\widehat{h}|_{k+1,\alpha}\Big)|u|_0 + \Big(|\partial\widehat{\Gamma}|_0 + |\partial\widehat{h}|_0\Big)|u|_{k+1,\alpha} \\ &\leq |\ell_{\widehat{\sigma}} u|_{k+2,\alpha} + C\Big(|\widehat{\Gamma}|_{k+2,\alpha} + |\widehat{h}|_{k+2,\alpha}\Big)|\ell_{\widehat{\sigma}} u|_{1,\alpha} \\ &\qquad + C\Big(|\widehat{\Gamma}|_1 + |\widehat{h}|_1\Big)\Big(|\ell_{\widehat{\sigma}} u|_{k+1,\alpha} + (|\widehat{\Gamma}|_{k+1,\alpha} + |\widehat{h}|_{k+1,\alpha})|\ell_{\widehat{\sigma}} u|_{1,\alpha}\Big) \\ &\leq C|\ell_{\widehat{\sigma}} u|_{k+2,\alpha} + C\Big(|\widehat{\Gamma}|_{k+2,\alpha} + |\widehat{h}|_{k+2,\alpha}\Big)|\ell_{\widehat{\sigma}}|_{1,\alpha}\end{aligned}$$

and in the same way

$$|\ell_{\widehat{\sigma}} u^d|_{1,\alpha} \leq |\ell_{\widehat{\sigma}} u|_{2,\alpha} + \Big(|\widehat{\Gamma}|_{2,\alpha} + |\widehat{h}|_{2,\alpha}\Big)|u|_{1,\alpha}.$$

So we obtain

$$\begin{aligned}|u^d|_{k+1,\alpha} \leq C\Big(&|\ell_{\widehat{\sigma}} u|_{k+2,\alpha} + C(|\widehat{\Gamma}|_{k+2,\alpha} + |\widehat{h}|_{k+2,\alpha})|\ell_{\widehat{\sigma}} u|_{1,\alpha} \\ &+ (|\widehat{\Gamma}|_{k+1,\alpha} + |\widehat{h}|_{k+1,\alpha})\big(|\ell_{\widehat{\sigma}} u|_{2,\alpha} + (|\widehat{\Gamma}|_{2,\alpha} + |\widehat{h}|_{2,\alpha})|u|_{1,\alpha}\big)\Big)\end{aligned}$$

and we finish with an immediate application of the interpolation inequalities:

$$\begin{aligned}
|\widehat{\Gamma}|_{k+1,\alpha}|\ell_{\widehat{\sigma}}u|_{2,\alpha} &\leq C\Big(|\widehat{\Gamma}|_{k+2,\alpha}|\ell_{\widehat{\sigma}}u|_{1,\alpha} + |\widehat{\Gamma}|_{1,\alpha}|\ell_{\widehat{\sigma}}u|_{k+2,\alpha}\Big)\\
|\widehat{h}|_{k+1,\alpha}|\ell_{\widehat{\sigma}}u|_{2,\alpha} &\leq C\Big(|\widehat{h}|_{k+2,\alpha}|\ell_{\widehat{\sigma}}u|_{1,\alpha} + |\widehat{h}|_{1,\alpha}|\ell_{\widehat{\sigma}}u|_{k+2,\alpha}\Big)\\
|\widehat{\Gamma}|_{k+1,\alpha}|\widehat{\Gamma}|_{2,\alpha} &\leq C|\widehat{\Gamma}|_{k+2,\alpha}|\widehat{\Gamma}|_{1,\alpha}\\
|\widehat{h}|_{k+1,\alpha}|\widehat{h}|_{2,\alpha} &\leq C|\widehat{h}|_{k+2,\alpha}|\widehat{h}|_{1,\alpha}\\
|\widehat{\Gamma}|_{k+1,\alpha}|\widehat{h}|_{2,\alpha} &\leq C\Big(|\widehat{\Gamma}|_{k+2,\alpha}|\widehat{h}|_{1,\alpha} + |\widehat{\Gamma}|_{1,\alpha}|\widehat{h}|_{k+2,\alpha}\Big)\\
|\widehat{h}|_{k+1,\alpha}|\widehat{\Gamma}|_{2,\alpha} &\leq C\Big(|\widehat{h}|_{k+2,\alpha}|\widehat{\Gamma}|_{1,\alpha} + |\widehat{h}|_{1,\alpha}|\widehat{\Gamma}|_{k+2,\alpha}\Big)
\end{aligned}$$

so that we can say

$$|u^d|_{k+1,\alpha} \leq C\Big(|\ell_{\widehat{\sigma}}u|_{k+2,\alpha} + \Big(|\widehat{\Gamma}|_{k+2,\alpha} + |\widehat{h}|_{k+2,\alpha}\Big)|\ell_{\widehat{\sigma}}u|_{1,\alpha}\Big)$$

to finish the induction. □

5.4 Estimate for pseudodifferential part of linearization

In this section we are concerned with the map

$$(\sigma, \tau; v) \mapsto \delta(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd})(v)\frac{\nabla_a\nabla_b\tau}{\sqrt{1+|\nabla\tau|^2}} \equiv \Psi_{\sigma,\tau}(v),$$

where, as always, $\widehat{\sigma} = \sigma + d\tau \otimes d\tau$, through which the indicated linearization depends on the infinitesimal perturbation v of τ.

Recall that we have both the physical data and solution we are perturbing around, here the only relevant parts being σ_0 and τ_0, and the perturbed data and 'solution' σ and τ. There is the corresponding $\widehat{\sigma} = \sigma + d\tau \otimes d\tau$ and $\widehat{\sigma}_0 = \sigma_0 + d\tau_0 \otimes d\tau_0$. Let $\widehat{X}$ and $\widehat{X}_0$ denote the isometric embeddings of these metrics into $\mathbb{R}^3$, with corresponding mean curvatures and second fundamental forms $\widehat{H}, \widehat{h}, \widehat{H}_0, \widehat{h}_0$ denoted in the obvious way.

Corresponding to the neighborhoods we needed to choose to have our estimates in the context of the Weyl problem, we need to make τ to be $C^{3,\alpha}$ close to τ_0 and σ to be $C^{2,\alpha}$ close to σ_0.

Lemma. *In this context, we have*

$$|\widehat{h}_{ij}|_{k,\alpha} \leq C\Big(1 + |\sigma_{ij}|_{k+2,\alpha} + |\tau|_{k+3,\alpha}\Big).$$

Proof. This is an immediate consequence of the tame estimate we proved for the second fundamental form of the Weyl problem, giving

$$|\widehat{h}_{ij}|_{k,\alpha} \leq C\Big(1 + |\widehat{\sigma}_{ij}|_{k+2,\alpha}\Big).$$

□

Lemma. *There are the formulas*

$$\delta\widehat{h}_{ij} = \widehat{h}_{ik}\widehat{\nabla}_j A^k + \widehat{h}_{jk}\widehat{\nabla}_i A^k + A^k\widehat{\nabla}_i\widehat{h}_{jk} - \widehat{\nabla}_i\widehat{\nabla}_j B + B\widehat{h}_{ik}\widehat{h}^k_j$$
$$\delta\widehat{H} = 2\widehat{h}_{ik}\widehat{\nabla}^i A^k + A^k\widehat{\nabla}_k\widehat{H} - \widehat{\Delta}B + B|\widehat{h}|^2 - \widehat{h}^{ij}\delta\widehat{\sigma}_{ij}$$

Proof. This is a direct calculation. The calculation for $\delta\widehat{H}$ is included at the end of [36]. □

Lemma. *There are the estimates*

$$|\widehat{\sigma}_{ij}|_{k,\alpha} \leq |\sigma_{ij}|_{k,\alpha} + C|\tau|_{k+1,\alpha}$$

Proof. Use $\widehat{\sigma}_{ij} = \sigma_{ij} + \tau_i\tau_j$ and restrict to a C^1 neighborhood of τ_0 for τ; the constant C will depend on this neighborhood.. □

Lemma. *There is*

$$|\widehat{\Gamma}^k_{ij}|_{k,\alpha} \leq C\Big(|\sigma_{ij}|_{k+1,\alpha} + |\tau|_{k+2,\alpha} + 1\Big)$$

Proof. Using $\widehat{\sigma}_{ij} = \sigma_{ij} + \tau_i\tau_j$ we calculate

$$\widehat{\sigma}^{ij} = \sigma^{ij} - \frac{\sigma^{ik}\sigma^{jl}\tau_k\tau_l}{1 + \sigma^{kl}\tau_k\tau_l}$$

and then

$$\widehat{\Gamma}^k_{ij} = \Gamma^k_{ij} + \widehat{\sigma}^{kl}\partial_i\partial_j\tau\partial_l\tau$$

by a direct computation. In the following we will use the C^1 bound on σ and the C^2 bound on τ. Using the Hölder estimate for inverses, we say

$$|\sigma^{ij}|_{k,\alpha} \leq C_k\Big(|\sigma_{ij}|_{k,\alpha} + 1\Big).$$

Using this and the Hölder estimate for compositions, we say

$$\begin{aligned}\left|\frac{1}{1+\sigma^{kl}\tau_k\tau_l}\right|_{k,\alpha} &\le C\Big(|\sigma^{kl}\tau_k\tau_l|_{k,\alpha}+1\Big)\\ &\le C\Big(|\sigma^{kl}|_{k,\alpha}|\tau|_1^2+|\tau|_1|\tau|_{k+1,\alpha}+1\Big)\\ &\le C\Big(|\sigma_{ij}|_{k,\alpha}+|\tau|_{k+1,\alpha}+1\Big).\end{aligned}$$

Putting both of these together, from the above computation of $\widehat{\sigma}^{ij}$, we find

$$|\widehat{\sigma}^{ij}|_{k,\alpha} \le C\Big(|\sigma_{ij}|_{k,\alpha}+|\tau|_{k+1,\alpha}+1\Big).$$

Given our estimate of σ^{ij}, it is rather easy to see from the definition of the Christoffel symbols that we have

$$|\Gamma^k_{ij}|_{k,\alpha} \le C\Big(|\sigma_{ij}|_{k+1,\alpha}+1\Big).$$

Putting the last two estimates together into the computation of $\widehat{\Gamma}^k_{ij}$, the lemma statement follows.

□

Lemma. *There is*

$$\begin{aligned}|\widehat{h}_{ij}|_{k+1,\alpha} &\le C\Big(1+|\sigma_{ij}|_{k+3,\alpha}+|\tau|_{k+4,\alpha}\Big)\\ |\widehat{\Gamma}^k_{ij}|_{k+1,\alpha} &\le C\Big(1+|\sigma_{ij}|_{k+2,\alpha}+|\tau|_{k+3,\alpha}\Big)\\ |\delta\widehat{\sigma}_{ij}|_{1,\alpha} &\le C|v|_{2,\alpha}\\ |\delta\widehat{\sigma}_{ij}|_{k+1,\alpha} &\le C\Big(|v|_{k+2,\alpha}+|\tau|_{k+2,\alpha}|v|_{1,\alpha}\Big)\end{aligned}$$

Proof. Restating the above. □

Lemma. *There is*

$$|A,B|_{k+1,\alpha} \le C\Big(|v|_{k+2,\alpha}+(1+|\sigma_{ij}|_{k+3,\alpha}+|\tau|_{k+4,\alpha})|v|_{2,\alpha}\Big)$$

Proof. Combination of the previous lemma and the lemma from the previous section. □

Lemma. *There is*

$$\begin{aligned}|\delta\widehat{h}|_{k,\alpha} &\le C\Big(1+|\sigma|_{k+4,\alpha}+|\tau|_{k+5,\alpha}\Big)|v|_{2,\alpha}+C|v|_{k+3,\alpha}\\ |\delta\widehat{H}|_{k,\alpha} &\le C\Big(1+|\sigma|_{k+4,\alpha}+|\tau|_{k+5,\alpha}\Big)|v|_{3,\alpha}+C|v|_{k+3,\alpha}.\end{aligned}$$

Proof. Schematically, from the previous calculation, we have

$$\delta\widehat{h} = \widehat{\sigma}^{-1}\Big(\widehat{h} * \partial A + A * \partial\widehat{h} + \widehat{\Gamma} * \widehat{h} * A\Big) + \partial^2 B + \widehat{\Gamma} * \partial B + \widehat{\sigma}^{-1} * B * \widehat{h} * \widehat{h}.$$

First we consider the A terms (the first three terms):

$$\begin{aligned}
|\delta\widehat{h}|^A_{k,\alpha} &\le C|\widehat{\sigma}^{-1}|_{k,\alpha}\Big(1 + |\sigma|_{3,\alpha} + |\tau|_{4,\alpha}\Big)|v|_{2,\alpha} + C\Big(|\widehat{h}|_{k+1,\alpha} + |\widehat{\Gamma}|_{k,\alpha}\Big)\Big(1 + |\sigma|_{3,\alpha} + |\tau|_{4,\alpha}\Big)|v|_{2,\alpha} \\
&\quad + C\Big(|v|_{k+2,\alpha} + (1 + |\sigma|_{k+3,\alpha} + |\tau|_{k+4,\alpha})|v|_{2,\alpha}\Big) \\
&\le C\Big(|\sigma|_{k,\alpha} + |\tau|_{k+1,\alpha} + 1\Big)|v|_{2,\alpha} + C\Big(1 + |\sigma|_{k+3,\alpha} + |\tau|_{k+4,\alpha}\Big)|v|_{2,\alpha} \\
&\quad + C\Big|v|_{k+2,\alpha} + C\Big(1 + |\sigma|_{k+3,\alpha} + |\tau|_{k+4,\alpha}\Big)|v|_{2,\alpha} \\
&\le C\Big(1 + |\sigma|_{k+3,\alpha} + |\tau|_{k+4,\alpha}\Big)|v|_{2,\alpha} + C|v|_{k+2,\alpha}.
\end{aligned}$$

Similarly we estimate the B terms (the remaining three terms):

$$\begin{aligned}
|\delta\widehat{h}|^B_{k,\alpha} &\le |B|_{k+2,\alpha} + \Big(|\widehat{\Gamma}|_{k,\alpha} + |\widehat{\sigma}^{-1}|_{k,\alpha} + |\widehat{h}|_{k,\alpha}\Big)|v|_{2,\alpha} \\
&\le C\Big(|v|_{k+3,\alpha} + (1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha})|v|_{2,\alpha}\Big) + C\Big(1 + |\sigma|_{k+2,\alpha} + |\tau|_{k+3,\alpha}\Big)|v|_{2,\alpha} \\
&\le C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{2,\alpha} + C|v|_{k+3,\alpha}.
\end{aligned}$$

So overall

$$|\delta\widehat{h}|_{k,\alpha} \le C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{2,\alpha} + C|v|_{k+3,\alpha}.$$

Similarly, schematically

$$\delta\widehat{H} = \widehat{\sigma}^{-1} * \delta\widehat{h} + \widehat{\sigma}^{-1} * \widehat{\sigma}^{-1} * \widehat{h} * (d\tau \otimes dv).$$

So we find

$$\begin{aligned}
|\delta\widehat{H}|_{k,\alpha} &\le C\Big(|\sigma|_{k,\alpha} + |\tau|_{k+1,\alpha} + 1\Big)|v|_{3,\alpha} + C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{2,\alpha} + C|v|_{k+3,\alpha} \\
&\quad + C\Big(1 + |\sigma|_{k+2,\alpha} + |\tau|_{k+3,\alpha}\Big)|v|_{1,\alpha} + C|v|_{k+1,\alpha} \\
&\le C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{3,\alpha} + C|v|_{k+3,\alpha}
\end{aligned}$$

as stated in the lemma. □

Lemma. *There is*

$$|\delta(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd})(v)|_{k,\alpha} \le C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{3,\alpha} + C|v|_{k+3,\alpha}$$

Proof. We have

$$\delta\widehat{\sigma}^{ab} = -\widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\delta\widehat{\sigma}_{cd}$$

with $\delta\widehat{\sigma}_{ij} = \tau_i v_j + \tau_j v_i$, so that

$$\begin{aligned}|\delta\widehat{\sigma}^{ab}|_{k,\alpha} &\leq C\Big(|\sigma|_{k,\alpha} + |\tau|_{k+1,\alpha}\Big)|v|_{1,\alpha} + C\Big(|v|_{k+1,\alpha} + |\tau|_{k+1,\alpha}|v|_{1,\alpha}\Big)\\ &\leq C\Big(|\sigma|_{k,\alpha} + |\tau|_{k+1,\alpha}\Big)|v|_{1,\alpha} + C|v|_{k+1,\alpha}.\end{aligned}$$

Putting this together with the previous lemma, we obtain

$$|\delta(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd})(v)|_{k,\alpha} \leq C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{3,\alpha} + C|v|_{k+3,\alpha}$$

as stated. □

Lemma. *There is*

$$|\Psi_{\sigma,\tau}(v)|_{k,\alpha} \leq C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{3,\alpha} + C|v|_{k+3,\alpha}.$$

Proof. From the previous lemma, it is immediate that we have

$$|\Psi_{\sigma,\tau}(v)|_{k,\alpha} \leq C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{3,\alpha} + C|v|_{k+3,\alpha} + C|v|_{3,\alpha}\left|\frac{\nabla\nabla\tau}{\sqrt{1+|\nabla\tau|^2}}\right|_{k,\alpha}.$$

So we just need to say

$$\left|\frac{\nabla\nabla\tau}{\sqrt{1+|\nabla\tau|^2}}\right|_{k,\alpha} \leq C|\tau|_{k+2,\alpha}$$

to finish. □

5.5 Estimate for linearized operator

To estimate the linearized operator we will follow the same outline of argument as for our estimate of the linearized Weyl problem. In particular, we begin by saying that, having identified the kernel and cokernel, there is

$$|v|_{k+4,\alpha} \leq C|D_\tau P_{(\sigma_0,|H|_0,\alpha_H^0;\tau_0)}(v)|_{k,\alpha}$$

where C depends upon the center point $(\sigma_0, |H|_0, \alpha_H^0; \tau_0)$. Now we need to estimate the dependence of $P_{(\sigma,|H|,\alpha_H;\tau)}(v)$ on a perturbation of $(\sigma, |H|, \alpha_H; \tau)$. As before, this is naturally split into two parts, that of the elliptic partial differential operator part and that of the pseudodifferential part.

5.5.1 Elliptic part

Recall that this part refers to the operator

$$E_{(\sigma,|H|,\alpha_H;\tau)}(v) = \delta\left(\operatorname{div}\left(\frac{\cosh\theta|H|\nabla\tau}{\sqrt{1+|\nabla\tau|^2}} - \nabla\theta - \alpha_H\right)(v)\right)$$

where θ denotes the expression

$$\sinh^{-1}\frac{-\Delta\tau}{|H|\sqrt{1+|\nabla\tau|^2}}.$$

This is a fourth order partial differential operator in v. We go through the coefficients carefully:

- in the fourth order coefficients, there is one derivative of σ, zero derivatives in $|H|$, no appearance of α_H, and two derivatives in τ
- in the third order coefficients, there are two derivatives of σ, one derivative of $|H|$, no appearance of α_H, and three derivatives of τ
- in the second order coefficients, there are three derivatives of σ, two derivatives of $|H|$, no appearance of α_H, and four derivatives of τ
- in the first order coefficients, there are three derivatives of σ, two derivatives of $|H|$, no appearance of α_H, and four derivatives of τ
- the zeroth order coefficient is zero.

Given that the dependence on these derivatives is all through elementary functions, compactly contained within their domains of dependence, we can then say from the estimate in [14] (see lemma II.3.3.2)

Lemma. *There is*

$$|v|_{k+4,\alpha} \leq C\Big(|E_{(\sigma,|H|,\alpha_H;\tau)}(v)|_{k,\alpha} + \big(1 + |\sigma|_{k+3,\alpha} + ||H||_{k+2,\alpha} + |\tau|_{k+4,\alpha}\big)|E_{(\sigma,|H|,\alpha_H;\tau)}(v)|_{0,\alpha}\Big).$$

5.5.2 Pseudodifferential part

Here we want to estimate

$$\Psi_{\sigma,\tau}(v) - \Psi_{\sigma_0,\tau_0}(v)$$

for nearby σ and σ_0 and nearby τ and τ_0. Denoting $\Psi_{\sigma,\tau}(v)$ by A_i, B and $\Psi_{\sigma_0,\tau_0}(v)$ by A_i^0, B^0, by definition we have

$$\partial_i A_j + \partial_j A_i - 2\widehat{\Gamma}_{ij}^k A_k + 2B\widehat{h}_{ij} = \tau_i v_j + \tau_j v_i$$
$$\partial_i A_j^0 + \partial_j A_i^0 - 2\circ\widehat{\Gamma}_{ij}^k A_k^0 + 2B^0\widehat{h}_{ij}^0 = \tau_i^0 v_j + \tau_j^0 v_i.$$

Subtracting these equations, we have

$$\partial_i(A_j - A_j^0) + \partial_j(A_i - A_i^0) - 2\circ\widehat{\Gamma}_{ij}^k(A_k - A_k^0) + 2(B - B^0)\widehat{h}^0 = 2(\widehat{\Gamma}_{ij}^k - \circ\widehat{\Gamma}_{ij}^k)A_k + 2B(\widehat{h}^0 - \widehat{h}) + (\tau_i - \tau_i^0)v_j + (\tau_j - \tau_j^0)v_i$$

That is to say,

$$\Psi_{\sigma,\tau}(v) - \Psi_{\sigma_0,\tau_0}(v) = \Psi_{\sigma_0,\tau_0}\Big(2(\widehat{\Gamma}_{ij}^k - \circ\widehat{\Gamma}_{ij}^k)A_k + 2B(\widehat{h}^0 - \widehat{h}) + (\tau_i - \tau_i^0)v_j + (\tau_j - \tau_j^0)v_i\Big).$$

According to our estimate for the linearized Weyl problem, we then have

$$\begin{aligned}
|\Psi_{\sigma,\tau}(v) - \Psi_{\sigma_0,\tau_0}(v)|_{k+1,\alpha} &\le C\Big|2(\widehat{\Gamma}_{ij}^k - \circ\widehat{\Gamma}_{ij}^k)A_k + 2B(\widehat{h}^0 - \widehat{h}) + (\tau_i - \tau_i^0)v_j + (\tau_j - \tau_j^0)v_i\Big|_{k+1,\alpha} \\
&\le C\Big(|\widehat{\sigma} - \widehat{\sigma}_0|_{k+2,\alpha}|\Psi_{\sigma,\tau}(v)|_{1,\alpha} + |\widehat{\sigma} - \widehat{\sigma}_0|_{2,\alpha}|\Psi_{\sigma,\tau}(v)|_{k+1,\alpha}\Big) \\
&\quad + C|\tau - \tau_0|_{k+2,\alpha}|v|_{1,\alpha} + C|\tau - \tau_0|_{1,\alpha}|v|_{k+2,\alpha}
\end{aligned}$$

We say

$$|\widehat{\sigma} - \widehat{\sigma}_0|_{k,\alpha} \le |\sigma - \sigma_0|_{k,\alpha} + |d\tau^2 - d\tau_0^2|_{k,\alpha} \le |\sigma - \sigma_0|_{k,\alpha} + C|\tau - \tau_0|_{k+1,\alpha}$$

and also recall our linearized Weyl estimate

$$\begin{aligned}
|\Psi_{\sigma,\tau}(v)|_{k+1,\alpha} &\le C\Big(|\tau_i v_j|_{k+1,\alpha} + (|\widehat{\Gamma}|_{k+1,\alpha} + |\widehat{h}|_{k+1,\alpha})|\tau_i v_j|_{1,\alpha}\Big) \\
&\le C\Big(|v|_{k+2,\alpha} + |\tau|_{k+2,\alpha}|v|_{1,\alpha} + (1 + |\sigma|_{k+3,\alpha} + |\tau|_{k+4,\alpha})|v|_{2,\alpha}\Big) \\
&\le C\Big(1 + |\sigma|_{k+3,\alpha} + |\tau|_{k+4,\alpha}\Big)|v|_{2,\alpha} + C|v|_{k+2,\alpha}
\end{aligned}$$

to conclude

$$\begin{aligned}
|\Psi_{\sigma,\tau}(v) - \Psi_{\sigma_0,\tau_0}(v)|_{k+1,\alpha} &\le C\Big(|\sigma - \sigma_0|_{k+2,\alpha} + |\tau - \tau_0|_{k+3,\alpha}\Big)|v|_{2,\alpha} \\
&\quad + C\Big(|\sigma - \sigma_0|_{2,\alpha} + |\tau - \tau_0|_{3,\alpha}\Big)\Big((1 + |\sigma|_{k+3,\alpha} + |\tau|_{k+4,\alpha})|v|_{2,\alpha} + |v|_{k+2,\alpha})\Big) \\
&\quad + C|\tau - \tau_0|_{k+2,\alpha}|v|_{1,\alpha} + C|\tau - \tau_0|_{1,\alpha}|v|_{k+2,\alpha} \\
&\le C\Big(|\sigma - \sigma_0|_{k+2,\alpha} + |\tau - \tau_0|_{k+3,\alpha}\Big)|v|_{2,\alpha} \\
&\quad + C\Big(|\sigma - \sigma_0|_{2,\alpha} + |\tau - \tau_0|_{3,\alpha}\Big)\Big((1 + |\sigma|_{k+3,\alpha} + |\tau|_{k+4,\alpha})|v|_{2,\alpha} + |v|_{k+2,\alpha})\Big).
\end{aligned}$$

Lemma. *We have*

$$\Phi_{\widehat{\sigma}}(\tau_i v_j + \tau_j v_i)^d - \Phi_{\widehat{\sigma}}(\tau_i v_j^d + \tau_j v_i^d) = \Phi_{\widehat{\sigma}}\Big(\tau_i^d v_j + \tau_j^d v_i - 2B\widehat{h}_{ij}^d - 2A_k(\widehat{\Gamma}_{ij}^k)^d\Big)$$

where $(A_i, B) = \Phi_{\widehat{\sigma}}(\tau_i v_j + \tau_j v_i)$.

Proof. This is a direct computation as follows. Denote $(\widetilde{A}_i, \widetilde{B}) = \Phi_{\widehat{\sigma}}(\tau_i v_j^d + \tau_j v_i^d)$. By definition

$$\partial_i A_j + \partial_j A_i + 2\widehat{\Gamma}_{ij}^k A_k + 2B\widehat{h}_{ij} = \tau_i v_j + \tau_j v_i$$
$$\partial_i \widetilde{A}_j + \partial_j \widetilde{A}_i + 2\widehat{\Gamma}_{ij}^k \widetilde{A}_k + 2\widetilde{B}\widehat{h}_{ij} = \tau_i v_j^d + \tau_j v_i^d.$$

Differentiate the first equation and subtract the second equation to get

$$\partial_i(A_j^d - \widetilde{A}_j) + \partial_j(A_i^d - \widetilde{A}_i) + 2\widehat{\Gamma}_{ij}^k(A_k^d - \widetilde{A}_k) + 2(B^d - \widetilde{B})\widehat{h}_{ij} = \tau_i^d v_j + \tau_j^d v_i - 2B\widehat{h}_{ij}^d - 2(\widehat{\Gamma}_{ij}^k)^d A_k$$

which proves the lemma. □

Lemma. *There is*

$$\Big|\Phi_{\widehat{\sigma}}(\tau_i v_j + \tau_j v_i)^d - \Phi_{\widehat{\sigma}}(\tau_i v_j^d + \tau_j v_i^d)\Big|_{k+1,\alpha} \le C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{2,\alpha} + C|v|_{k+2,\alpha}.$$

Proof. We apply our estimate for the linearized Weyl problem, noting

$$\begin{aligned}
|\tau_i^d v_j + \tau_j^d v_i - 2B\widehat{h}_{ij}^d - 2A_k(\widehat{\Gamma}_{ij}^k)^d|_{k+1,\alpha} &\le C\Big(|\tau|_{k+3,\alpha}|v|_{1,\alpha} + |v|_{k+2,\alpha}\Big) \\
&\quad + C(|\widehat{h}|_{1,\alpha} + |\widehat{\Gamma}_{ij}^k|_{1,\alpha})|\Phi_{\widehat{\sigma}}(\tau_i v_j + \tau_j v_i)|_{k+1,\alpha} \\
&\quad + C(|\widehat{h}|_{k+2,\alpha} + |\widehat{\Gamma}_{ij}^k|_{k+2,\alpha})|\Phi_{\widehat{\sigma}}(\tau_i v_j + \tau_j v_i)|_{1,\alpha} \\
&\le C\Big(|\tau|_{k+3,\alpha}|v|_{1,\alpha} + |v|_{k+2,\alpha}\Big) \\
&\quad + C\Big(|v|_{k+2,\alpha} + (1 + |\sigma|_{k+3,\alpha} + |\tau|_{k+4,\alpha})|v|_{2,\alpha}\Big) \\
&\quad + C\Big(|\widehat{h}|_{k+2,\alpha} + |\widehat{\Gamma}_{ij}^k|_{k+2,\alpha}\Big)|v|_{2,\alpha} \\
&\le C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{2,\alpha} + C|v|_{k+2,\alpha}
\end{aligned}$$

and

$$\begin{aligned}
|\tau_i^d v_j + \tau_j^d v_i - 2B\widehat{h}_{ij}^d - 2A_k(\widehat{\Gamma}_{ij}^k)^d|_{1,\alpha} &\le C\Big(|v|_{2,\alpha} + |\Phi_{\widehat{\sigma}}(\tau_i v_j + \tau_j v_i)|_{1,\alpha}\Big) \\
&\le C|v|_{2,\alpha}
\end{aligned}$$

so that according to our Weyl estimate,

$$\Big|\Phi_{\widehat{\sigma}}(\tau_i v_j + \tau_j v_i)^d - \Phi_{\widehat{\sigma}}(\tau_i v_j^d + \tau_j v_i^d)\Big|_{k+1,\alpha} \leq C\Big(1 + |\sigma|_{k+4,\alpha} + |\tau|_{k+5,\alpha}\Big)|v|_{2,\alpha} + C|v|_{k+2,\alpha}.$$

□

Lemma. *There is*

$$|\Psi_{\sigma,\tau}(v)^d - \Psi_{\sigma,\tau}(v^d)|_{k,\alpha} \leq C\Big(1 + |\sigma|_{k+5,\alpha} + |\tau|_{k+6,\alpha}\Big)|v|_{3,\alpha} + C|v|_{k+3,\alpha}$$

Proof. As in the previous section. □

5.5.3 Estimate

As indicated in the beginning of the section, we begin with the estimate

$$|v|_{4,\alpha} \leq C|D_\tau P_{(\sigma_0,|H|_0,\alpha_H^0;\tau_0)}(v)|_{0,\alpha}.$$

By the triangle inequality, this implies

$$|v|_{4,\alpha} \leq C|D_\tau P_{(\sigma,|H|,\alpha_H;\tau)}(v)|_{0,\alpha} + C\Big|E_{(\sigma,|H|,\alpha_H;\tau)}(v) - E_{(\sigma_0,|H|_0,\alpha_H^0;\tau_0)}(v)\Big|_{0,\alpha} + C\Big|\Psi_{\sigma,\tau}(v) - \Psi_{\sigma_0,\tau_0}(v)\Big|_{0,\alpha}.$$

According to the previous section, if we make σ sufficiently close to σ_0 in $C^{2,\alpha}$ and τ sufficiently close to τ_0 in $C^{3,\alpha}$, then the last term on the right can be absorbed on the left. According to the derivative counts in the previous previous section, if we make σ sufficiently close to σ_0 in $C^{3,\alpha}$, $|H|$ sufficiently close to $|H|$ in $C^{2,\alpha}$, and τ sufficiently close to τ_0 in $C^{4,\alpha}$, then the other term on the right can also be absorbed on the left, and we have

$$|v|_{4,\alpha} \leq C|D_\tau P_{(\sigma,|H|,\alpha_H;\tau)}(v)|_{0,\alpha}.$$

Upon applying this to a derivative v^d of v, we find

$$\begin{aligned}|v^d|_{4,\alpha} &\leq C|D_\tau P_{(\sigma,|H|,\alpha_H;\tau)}(v^d)|_{0,\alpha} \\ &\leq C|D_\tau P_{(\sigma,|H|,\alpha_H;\tau)}(v)^d|_{0,\alpha} + C|E_{(\sigma,|H|;\tau)}(v)^d - E_{(\sigma,|H|;\tau)}(v^d)|_{0,\alpha} + C|\Psi_{\sigma,\tau}(v)^d - \Psi_{\sigma,\tau}(v^d)|_{0,\alpha} \\ &\leq C|D_\tau P_{(\sigma,|H|,\alpha_H;\tau)}(v)^d|_{0,\alpha} + C(1 + |\sigma|_{3,\alpha} + ||H||_{2,\alpha} + |\tau|_{4,\alpha})|v|_{4,\alpha} + C\Big(1 + |\sigma|_{5,\alpha} + |\tau|_{6,\alpha}\Big)|v|_{3,\alpha}.\end{aligned}$$

Using the previous estimate we obtain

$$|v|_{5,\alpha} \leq C|D_\tau P_{(\sigma,|H|,\alpha_H;\tau)}(v)|_{1,\alpha} + C\Big(1 + |\sigma|_{5,\alpha} + ||H||_{2,\alpha} + |\tau|_{6,\alpha}\Big)|D_\tau P_{(\sigma,|H|,\alpha_H;\tau)}(v)|_{0,\alpha}.$$

Continuing inductively we get the following estimate

Lemma. *There is*

$$|v|_{k+4,\alpha} \leq C|D_\tau P_{(\sigma,|H|,\alpha_H;\tau)}(v)|_{k,\alpha} + C\Big(1 + |\sigma|_{k+4,\alpha} + ||H||_{k+1,\alpha} + |\tau|_{k+5,\alpha}\Big)|D_\tau P_{(\sigma,|H|,\alpha_H;\tau)}(v)|_{0,\alpha}.$$

This provides a tame estimate for the inverse of $D_\tau P_{(\sigma,|H|,\alpha_H;\tau)}$, considered as a map $C^\infty \times C^\infty \times C^\infty \times C^\infty_- \to C^\infty_-$. According to the Nash-Moser implicit function theorem [14], this provides the convergence of the Newton-Nash-Moser iteration.

Table 5.1: Notations

Notation	**Meaning**
σ_0	a fixed Riemannian metric on S^2
σ	a Riemannian metric on S^2 which is close to σ_0
τ_0	a fixed function on S^2
τ	a function on S^2 which is close to τ_0
$\widehat{\sigma}_0$	$\sigma_0 + d\tau_0 \otimes d\tau_0$
$\widehat{\sigma}$	$\sigma + d\tau \otimes d\tau$
$\widehat{X}_0$	an isometric embedding of $\widehat{\sigma}_0$ into $\mathbb{R}^3$
$\widehat{X}$	an isometric embedding of $\widehat{\sigma}$ into $\mathbb{R}^3$
$\widehat{h}_0, \widehat{h}$	the second fundamental forms of $\widehat{X}_0$ and $\widehat{X}$
$\widehat{H}_0, \widehat{H}$	the mean curvatures of $\widehat{X}_0$ and $\widehat{X}$
A_j, B	the decomposition of a function $Y = \widehat{\sigma}^{ab} A_a \frac{\partial \widehat{X}}{\partial u^b} + B\widehat{\nu}$
$\Psi_{\sigma,\tau}(v)$	$\delta(\widehat{H}\widehat{\sigma}^{ab} - \widehat{\sigma}^{ac}\widehat{\sigma}^{bd}\widehat{h}_{cd})(v)\dfrac{\nabla_a\nabla_b\tau}{\sqrt{1+\|\nabla\tau\|^2}}$
$\Phi_{\widehat{\sigma}}(\delta\widehat{\sigma})$	solution of the linearized Weyl problem at $\widehat{\sigma}$ in direction $\delta\widehat{\sigma}$
$P(\sigma, \|H\|, \alpha_H; \tau)$	Euler-Lagrange operator for Wang-Yau energy
$E_{(\sigma,\|H\|;\tau)}(v)$	$DP_{(\sigma,\|H\|;\tau)}(v) + \Psi_{\sigma,\tau}(v)$
θ	$\sinh^{-1} \dfrac{-\Delta\tau}{\|H\|\sqrt{1+\|\nabla\tau\|^2}}$
$\Big(\Big)^d$	some coordinate derivative of the expression $\Big(\Big)$

Bibliography

[1] Richard Arnowitt, Stanley Deser and Charles Misner, *Dynamical structure and definition of energy in general relativity.* Physical Review 116:1322-1330 (1959).

[2] H. Bondi, M.G.J. van der Burg and A.W.K. Metzner, *Gravitational waves in general relativity. VII. Waves from axi-symmetric isolated systems.* Proceedings of the Royal Society, Series A 269:21-52 (1962).

[3] J. David Brown and James York Jr., *Quasilocal energy and conserved charges derived from the gravitational action.* Physical Review D 47(3):1407-1419 (1993).

[4] Po-Ning Chen, Mu-Tao Wang and Shing-Tung Yau, *Evaluating quasilocal energy and solving optimal embedding equation at null infinity.* Communications in Mathematical Physics 308(3):845-863 (2011).

[5] Po-Ning Chen, Mu-Tao Wang and Shing-Tung Yau, *Minimizing properties of critical points of quasi-local energy.* Communications in Mathematical Physics 329(3):919-935 (2014).

[6] Po-Ning Chen, Mu-Tao Wang and Shing-Tung Yau, *Conserved quantities in general relativity: from the quasi-local level to spatial infinity.* Communications in Mathematical Physics 338(1):31-80 (2015).

[7] Demetrios Christodoulou and Sergiu Klainerman, *The Global Nonlinear Stability of the Minkowski Space.* Princeton University Press, Princeton, NJ (1993).

[8] Gaston Darboux, *Théorie des Surfaces, Volume 3.* Paris (1894).

[9] Avron Douglis and Louis Nirenberg, *Interior estimates for elliptic systems of partial differential equations.* Communications on Pure and Applied Mathematics 8:503-538 (1955).

[10] David Gilbarg and Neil Trudinger, *Elliptic partial differential equations of second order*. Springer-Verlag, Berlin (1983).

[11] Mikhael Gromov, *Partial Differential Relations*. Springer-Verlag, Berlin (1986).

[12] Mikhael Gromov and Vladimir Rokhlin, *Imbeddings and immersions in Riemannian geometry*. Uspehi Mat. Nauk 25(5):3-62 (1970).

[13] Pengfei Guan and Yanyan Li, *The Weyl problem with nonnegative Gauss curvature*. Journal of Differential Geometry 39:331-342 (1994).

[14] Richard Hamilton, *The inverse function theorem of Nash and Moser*. Bulletin of the AMS 7(1):65-222 (1982).

[15] Gustav Herglotz, *Über die Starrheit der Eiflächen*. Abhandlungen des mathematischen Seminars der Hansischen Universität, 1943.

[16] Lars Hörmander, *The boundary problems of physical geodesy*. Archive for Rational Mechanics and Analysis 62(1):1-52 (1976).

[17] Lars Hörmander, *The Analysis of Linear Partial Differential Operators III: Pseudo-Differential Operators*. Springer-Verlag, Berlin (1994).

[18] Gerhard Huisken and Tom Ilmanen, *The inverse mean curvature flow and the Riemannian Penrose inequality*. Journal of Differential Geometry 59(3):353-437 (2001).

[19] H. Blaine Lawson and Marie-Louise Michelsohn, *Spin Geometry*. Princeton University Press, Princeton, NJ (1989).

[20] Chiu-Chu Melissa Liu and Shing-Tung Yau, *Positivity of quasi-local mass II*. Journal of the AMS 19:181-204 (2006).

[21] Michele Maggiore, *Gravitational Waves, Volume 1: Theory and Experiments*. Oxford University Press, New York (2008).

[22] Pengzi Miao, Luen-Fai Tam and Naqing Xie, *Critical points of Wang-Yau quasi-local energy*. Annales Henri Poincaré 12:987-1017 (2011).

[23] Charles Morrey, *Multiple Integrals in the Calculus of Variations*. Springer-Verlag, New York (1966).

[24] Jürgen Moser, *A rapidly convergent iteration method and non-linear partial differential equations I*. Annali della Scuola Normale Superiore di Pisa - Classe de Scienze 20(2):265-315 (1966).

[25] John Nash, *The imbedding problem for Riemannian manifolds*. Annals of Mathematics 63(1):20-63 (1956).

[26] Louis Nirenberg, *The Weyl and Minkowski problems in differential geometry in the large*. Communications on Pure and Applied Mathematics 6:337-394 (1953).

[27] Thomas Parker and Clifford Taubes, *On Witten's proof of the positive energy theorem*. Communications in Mathematical Physics 80(2):223-238 (1982).

[28] Roger Penrose. *Some unsolved problems in classical general relativity*. Seminar on Differential Geometry, Ann. of Math. Stud. 102, Princeton, NJ. Princeton University Press (1982).

[29] Aleksei Pogorelov, *Extrinsic Geometry of Convex Surfaces*. American Mathematical Society, Providence, R.I. (1973).

[30] Richard Schoen and Shing-Tung Yau, *On the proof of the positive mass conjecture in general relativity*. Communications in Mathematical Physics 65(1):45-76 (1979).

[31] Richard Schoen and Shing-Tung Yau, *Proof of the positive mass theorem. II*. Communications in Mathematical Physics 79(2):231-260 (1981).

[32] Jacob Schwartz, *On Nash's implicit function theorem*. Communications on Pure and Applied Mathematics 13:509-530 (1960).

[33] Yuguang Shi and Luen-Fai Tam, *Positive mass theorem and the boundary behaviors of compact manifolds with nonnegative scalar curvature*. Journal of Differential Geometry 62(1):79-125 (2002).

[34] Robert Wald, *General Relativity*. Chicago University Press, Chicago (1984).

[35] Mu-Tao Wang and Shing-Tung Yau, *Quasilocal mass in general relativity*. Physical Review Letters 102(2), 021101 (2009).

[36] Mu-Tao Wang and Shing-Tung Yau, *Isometric embeddings into the Minkowski space and new quasi-local mass*. Communications in Mathematical Physics 288(3):919-942 (2009).

[37] Mu-Tao Wang and Shing-Tung Yau, *Limit of quasilocal mass at spatial infinity*. Communications in Mathematical Physics 296(1):271-283 (2010).

[38] Edward Witten, *A new proof of the positive energy theorem*. Communications in Mathematical Physics 80(3):381-402 (1981).

[39] Eduard Zehnder, *Generalized implicit function theorems with applications to some small divisor problems I*. Communications on Pure and Applied Mathematics 28:91-141 (1975).

[40] Eduard Zehnder, *Generalized implicit function theorems with applications to some small divisor problems II*. Communications on Pure and Applied Mathematics 29:49-113 (1976).