

# Columbia Journal *of* Undergraduate Mathematics

ISSUE 1

June 29, 2026

## About the Journal

*Aim and Scope.* The primary goal of the *Columbia Journal of Undergraduate Mathematics* is to provide undergraduate readers with high-quality, accessible articles on challenging topics, or novel approaches to teaching more familiar concepts. Articles published are purely expository; we do not accept research papers. Most are under 20 pages in length, with the primary exceptions being senior theses written by students at Columbia and other universities alike. The journal also accepts and publishes mathematical artwork with clear pedagogical value.

*Copyrights and Permissions.* The *Columbia Journal of Undergraduate Mathematics* is an open access journal, which means that all content is freely available without charge to the user or their institution. Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles in this journal without asking prior permission from the publisher or the author. Authors retain their copyright and agree to license their articles with a CC BY-NC-ND 4.0 license. You can read more about Creative Commons licenses at [creativecommons.org](http://creativecommons.org).

*Contact.* Email [columbiajournalofundergradmath@gmail.com](mailto:columbiajournalofundergradmath@gmail.com) with any questions, or visit <https://journals.library.columbia.edu/index.php/cjum> for more information.

## Letter from the Editors

The *Columbia Journal of Undergraduate Mathematics*' second annual thesis issue continues to, as we had originally hoped for and planned, provide a record for the excellent work done by graduating seniors for their capstone projects. As in the first iteration of our thesis issue, the works showcased here are less accessible to the wider undergraduate audience than those of our general issue. All the same, they are meant to serve as a touchpoint for undergraduates intending to write expository senior theses, in that they can gain a good sense of the scope of such a project. In pursuit of the goal of showcasing the various interests of the graduating student body, we often anticipate submissions drawing from a wide range of mathematical fields. However, this does not preclude the possibility that we receive many idiosyncratic works originating from the same overarching field. Case in point, all of our submissions for this year's senior thesis edition were in some way interwoven with the study of geometry.

The two submissions we chose for publication are specifically focused on the field of geometric analysis, wherein the behavior of partial differential equations (PDEs) is used to deduce geometric properties of the space they live in, and vice versa. Both address questions that are either physically or intuitively well-motivated, but whose answers shed light on the nature of more advanced geometrical and analytical techniques used within the field. "One Small Step for Stability, One Giant Leap for Schwarzschild: Boundedness of Scalar Waves on Schwarzschild Spacetimes" studies how a particularly important example of a PDE – the wave equation – behaves over a Lorentzian manifold, which is the natural setting for Einstein's theory of general relativity. The thesis draws connections between physical phenomena, like gravitational redshift, to the geometrical intricacies of a Lorentzian manifold, such as curvature and causality, in order to study pointwise bounds on the solutions to the wave equation in this space. "On Gunther's Perturbation Results for Nash's Isometric Embedding Theorems" asks, given an  $n$ -dimensional manifold of arbitrary geometry, can we always isometrically embed it within a higher-dimensional Euclidean space, and if so, can we place a bound on the dimension of this space? This question is concerned only with the global geometry of the manifold and appears, on the outset, to be entirely divorced from the study of PDEs. Despite this, this thesis demonstrates how the solution to this geometrical problem is, in fact, found in the analysis of a system of non-linear PDEs, illuminating the uncanny connection that lies at the heart of geometric analysis.

Finally, we want to highlight our wonderful team of content editors, copy editors, and graduate student editors for all the time and effort they have put into making this first issue. We would also like to acknowledge the support

from the Columbia Libraries and the Department of Mathematics, who were the technical and financial backbone to make this project come to life.

Sincerely,

Analisa Faulkner, Jon Kwong, and Cassidy Kao

## Masthead

<b>Analisa Faulkner Valiente</b>	Editor-in-Chief
<b>Jon Kwong</b>	Head of Content
<b>Cassidy Kao</b>	Head of Copy
<b>Elizabeth Frost</b>	Managing Editor
<b>Sabrina Yu</b>	Assistant Head of Content
<b>Roger Yang</b>	Assistant Head of Copy
<b>Harry Zhang</b>	Assistant Managing Editor
<b>Joseph McGill</b>	Events Coordinator
<b>Michael Ma</b>	Outreach Coordinator
<b>Cathy Deng</b>	Social Media Co-chair
<b>Jia Hao Liang (Frank)</b>	Social Media Co-chair
<b>Sophie He</b>	Treasurer
<b>Emma Jia</b>	Merch Chair
<b>Amanda Lin</b>	Copy Editor
<b>Andrew Yoo</b>	Content Editor
<b>Aruzhan Abil</b>	Content Editor
<b>Carmen Alves Sabin</b>	Content Editor
<b>Emory Sun</b>	Content Editor
<b>Eshan Kabir</b>	Content Editor
<b>Felix Liu</b>	Content Editor
<b>Jackson Bramlette</b>	Content Editor

<b>Jasper Kra-Caskey</b>	Content Editor
<b>Keir Dorchen</b>	Content Editor
<b>Luca Nashabeh</b>	Content Editor
<b>Shilpa Kesavan</b>	Content Editor
<b>Shreyas Singh</b>	Content Editor
<b>Shrina Dong</b>	Content Editor
<b>Yanyi Foong</b>	Copy Editor
<b>Azelie Picot</b>	Graduate Student Editor
<b>Carlos Alvarado</b>	Graduate Student Editor
<b>Ethan Hall</b>	Graduate Student Editor
<b>Felix Roz</b>	Graduate Student Editor
<b>George Dragomir</b>	Faculty Advisor
<b>Francesco Lin</b>	Faculty Advisor

# Contents

Page No.

**About The Journal**

2

**Letter from the Editors**

3

**Masthead**

5

.....

**One Small Step for Stability, One Giant Leap for Schwarzschild:  
Boundedness of Scalar Waves on Schwarzschild Spacetime**

Kate Mekechuk

**On Gunther's Perturbation Results for Nash's Isometric  
Embedding Theorems**

Shiv Yajnik

# One small step for stability, one giant leap for Schwarzschild: Boundedness of scalar waves on Schwarzschild spacetimes

By KATHERINE MEKECHUK

## Abstract

This paper gives a detailed proof of the boundedness of the scalar wave equation on a Schwarzschild spacetime using modern energy estimates. The first proof of this statement was given by Kay and Wald in [10] by using Killing vector fields and discrete isometries of the spacetime. However, since then, new techniques to analyze vector fields near the event horizon have been developed by Dafermos and Rodnianski in [4] which has given rise to a new proof strategy. This new proof, using the red shift effect, has been sketched in many lecture notes (see [5], [3], and [9]), but many details are left out. This paper fills in those details and presents the theorem in a self-contained manner.

## Contents

1. Introduction	2
2. The Homogeneous Scalar Wave Equation	4
2.1. Associated Energy Identity	6
2.2. Well Posedness and Uniqueness	11
3. The Redshift Effect	12
3.1. Global and Local Redshift	12
3.2. Redshift in Geometric Optics	13
4. Proof of Theorem 1.1	14
4.1. Schwarzschild Geometry	14
4.2. Pointwise Bounds away from the Horizon	18
4.3. Pointwise Bounds including the Horizon	20
5. Conclusion	24

---

Received by the editors May 2025.

© 2026 Mekechuk, Katherine. This is an open access article distributed under the terms of the Creative Commons BY-NC-ND 4.0 license.

## 1. Introduction

One of the major types of open problems in general relativity is the nonlinear stability of various black hole spacetimes. A *black hole spacetime* is one that has a complete future null infinity for which the past includes a regular future boundary. This regular future boundary is called the *event horizon*, and the future of the event horizon is the *black hole*. We can further define the entire past of null infinity to be the *black hole exterior*.

The simplest exact solution to the Einstein vacuum equation

$$(1) \quad \text{Ric}(g) = 0$$

that contains a black hole is the one parameter family of Schwarzschild metrics. These black hole spacetimes are spherically symmetric and static, and each metric depends on the parameter  $M$  which is interpreted as the mass of the black hole. In local coordinates  $(t, r, \theta, \phi)$ , the metric take the form

$$(2) \quad g_M = - \left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2).$$

However in these coordinates, a Schwarzschild spacetime has two singularities: a metric singularity at  $r = 0$  and a coordinate singularity at  $r = 2M$ . For this reason, only when Kruskal showed that the event horizon could be covered by regular coordinates, did the notion of a “black hole spacetime” become a formal definition [11]. This allowed for a geometric formulation of the question of stability. In his paper, Kruskal showed that we can consider new coordinates that take the form  $(T, R, \theta, \phi)$  where

$$T^2 - R^2 = \left(\frac{r}{2M} - 1\right) e^{r/2M} \quad \text{and} \quad 2M \log \left| \frac{T + R}{T - R} \right| = t.$$

In these coordinates and when  $T^2 - R^2 < 1$ , a Schwarzschild metric takes the form

$$(3) \quad g = \frac{32M^3}{r} e^{-r/2M} (-dT^2 + dR^2) + r^2 (d\theta^2 + \sin^2 \theta d\phi^2).$$

As it turns out, the one parameter family of Schwarzschild metrics is a subfamily of the two-parameter family of Kerr metrics. These metrics are stationary and axisymmetric and describe stationary rotating black holes. When the

metrics are also spherically symmetric and static, they are called the Reissner-Nordström metrics, and take the form

$$(4) \quad g_{M,Q} = - \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right) dt^2 + \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)^{-1} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2).$$

Here,  $M$  is interpreted as the black hole mass and  $Q$  is interpreted as the electric charge of the black hole. For more information about the linear stability of Reissner-Nordström metrics, please see [8].

While the nonlinear stability of Kerr metrics is still an open problem, techniques have been developed through the study of simplified versions. This is exactly the case for uniform boundedness of scalar waves on Schwarzschild spacetimes—the main theorem of the present paper. The theorem was first proven in 1986 by Kay and Wald in [10].

**THEOREM 1.1.** *Let  $(M, g)$  be a Schwarzschild spacetime and let  $\varphi$  be the smooth solution to the wave equation that satisfies some initial data  $f \in H_{\text{loc}}^m(\Sigma_0)$  given on an initial data set  $(\Sigma_0, \mathfrak{g}, \kappa)$ . Then, there exists a  $C > 0$  such that for all events  $p$  in exterior Schwarzschild region (including the horizons)*

$$|\varphi(p)| \leq C.$$

Their proof uses Killing vector fields and derivative estimates as outlined in Section 2 of this paper. However, it also relies heavily on discrete isometries of a Schwarzschild spacetime, making it difficult to generalize their methods in the nonlinear case. In 2005, Dafermos and Rodnianski developed methods to study non-Killing vector fields defined by the *local redshift effect* in Schwarzschild spacetimes. The redshift effect and surface gravity are discussed in Section 3. In fact, this construction generalizes Theorem 1.1 and has been useful for solving other problems in mathematical general relativity. This is discussed in more detail in Section 5.

The present paper uses the techniques developed by Dafermos and Rodnianski [4] to prove Kay and Wald’s Theorem 1.1 in Section 4. The proof of this is outlined in many lecture notes (see [5], [3], and [9]), all of which leave out important details as exercises to the reader. This paper delves into those details, using rigorous proof techniques, to make the details accessible to newer students of the subject.

This paper is meant to be read by an advanced undergraduate student interested in learning more about general relativity. The ideal student has been introduced to the theory and has an advanced background in the analysis of

partial differential equations. More specifically, she should have have knowledge about the following.

- (1) With respect to general relativity, the reader is assumed to have studied Lorentzian geometry. She should know about causal properties and be able to analyze Penrose diagrams. Furthermore, she should know about trapped surfaces and black holes, the Penrose incompleteness theorem, and the Schwarzschild metric. Please refer to the first seven chapters of [1] for a full introduction to Lorentzian geometry.
- (2) With respect to the analysis of partial differential equations, the reader should have a first year graduate student level background in theory. This means that she should have already seen the material outlined in the classic partial differential equations textbook by Evans [6].

## 2. The Homogeneous Scalar Wave Equation

Theorem 1.1 gives a pointwise bound for solutions of the homogeneous wave equation on Schwarzschild spacetimes. To understand what this entails, we will first define this equation on general Lorentzian manifolds. Let  $(\mathcal{M}, g)$  be an  $(n + 1)$  dimensional Lorentzian manifold. For a scalar function  $\varphi : \mathcal{M} \rightarrow \mathbb{R}$ , the *covariant wave operator*  $\square_g$  is defined as follows.

$$\square_g \varphi := \nabla^a \nabla_a \varphi = g^{\mu\nu} \partial_\mu \partial_\nu \varphi - g^{\mu\nu} \Gamma_{\mu\nu}^\alpha \partial_\alpha \varphi$$

Moreover, the *homogeneous wave equation* is

$$(5) \quad \square_g \varphi = 0.$$

With respect to Minkowski spacetime  $(\mathbb{R}^4, m)$  and in coordinates  $(t, x, y, z)$ , this looks like

$$\square_m \varphi = \partial_t^2 \varphi - \Delta \varphi = 0$$

where  $\Delta \varphi$  represents the induced Laplacian of  $\varphi$  on  $\mathbb{R}^3$ .

We can view (5) as the Euler-Lagrange equation corresponding to the Lagrangian

$$\mathcal{L}[\varphi] = -\frac{1}{2} \sqrt{-g} g^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi$$

where  $g = \det(g_{\mu\nu})$ . Furthermore, the corresponding action to this density is

$$S[\varphi] = \int_{\mathcal{M}} \mathcal{L}[\varphi] d^n x = -\frac{1}{2} \int_{\mathcal{M}} g^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi \sqrt{-g} d^n x.$$

On the other hand, we can define the stress energy tensor  $\mathbb{T}^\mu{}_\nu$  in terms of the same Lagrangian:

$$\mathbb{T}^\mu{}_\nu := \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \partial_\nu \varphi - \delta_\nu^\mu \mathcal{L}.$$

*Definition 2.1.* For a solution to the homogeneous wave equation  $\varphi$ , the *energy-momentum tensor*  $\mathbb{T}_{\mu\nu}$  is

$$\mathbb{T}_{\mu\nu}[\varphi] = g_{\mu\alpha}\mathbb{T}^{\alpha}_{\nu}[\varphi] = g_{\mu\alpha}g_{\nu\beta}\mathbb{T}^{\alpha\beta}[\varphi] = \partial_{\mu}\varphi\partial_{\nu}\varphi - \frac{1}{2}g_{\mu\nu}(g^{\alpha\beta}\partial_{\alpha}\varphi\partial_{\beta}\varphi).$$

In fact, for a solution to (5), the energy-momentum tensor is divergence free. To see this, one can easily derive that

$$\nabla^{\mu}\mathbb{T}_{\mu\nu} = (\square_g\varphi)\nabla_{\nu}\varphi.$$

Thus, if  $\square_g\varphi = 0$ , then  $\nabla^{\mu}\mathbb{T}_{\mu\nu} = 0$ . We call this property the *divergence identity* and it will play a very important roll in later calculations. Before that discussion, we need to understand another important property of the energy momentum tensor: the *positivity property*.

**THEOREM 2.2 (Positivity Property).** *Let  $X$  and  $Y$  be two future directed causal vectors. Then,*

$$(6) \quad \mathbb{T}_{\mu\nu}X^{\mu}Y^{\nu} \geq 0.$$

*Proof.* First, note that the coordinate free formula for the energy momentum tensor is

$$\mathbb{T}(A, B) = (A\varphi)(B\varphi) - \frac{1}{2}g(A, B)|\nabla\varphi|^2$$

where  $A$  and  $B$  are two vectors. Now, define a null frame  $(L, \underline{L}, e_1, e_2)$  such that

- $L$  and  $\underline{L}$  are future directed non-collinear null vectors normalizes so that  $g(L, \underline{L}) = -2$ ,
- $e_1$  and  $e_2$  are spacelike unit vectors orthonormal to  $L, \underline{L}$ , and to each other.

In this frame,

$$|\nabla\varphi|^2 = g^{\alpha\beta}\nabla_{\alpha}\varphi\nabla_{\beta}\varphi = -(L\varphi)(\underline{L}\varphi) + (e_1\varphi)^2 + (e_2\varphi)^2.$$

Since  $g(L, L) = 0 = g(\underline{L}, \underline{L})$  it follows that

$$\mathbb{T}(L, L) = (L\varphi)^2 \geq 0 \quad \text{and} \quad \mathbb{T}(\underline{L}, \underline{L}) = (\underline{L}\varphi)^2 \geq 0.$$

On the other hand, the normalization  $g(L, \underline{L}) = -2$  allows us to calculate

$$\begin{aligned} \mathbb{T}(L, \underline{L}) &= (L\varphi)(\underline{L}\varphi) - \frac{1}{2}(-2) \left( -(L\varphi)(\underline{L}\varphi) + (e_1\varphi)^2 + (e_2\varphi)^2 \right) \\ &= (L\varphi)(\underline{L}\varphi) - (L\varphi)(\underline{L}\varphi) + (e_1\varphi)^2 + (e_2\varphi)^2 \\ &= (e_1\varphi)^2 + (e_2\varphi)^2 \geq 0 \\ &= \mathbb{T}(\underline{L}, L). \end{aligned}$$

We can use this to calculate  $\mathbb{T}(X, Y)$  when  $X$  and  $Y$  future directed causal vectors. If this is the case, then there exists some constants  $a, b, c, d \in (0, \infty)$

such that

$$X = aL + b\underline{L} \quad \text{and} \quad Y = cL + d\underline{L}.$$

This means that

$$\begin{aligned} \mathbb{T}(X, Y) &= \mathbb{T}(aL + b\underline{L}, cL + d\underline{L}) \\ &= ac\mathbb{T}(L, L) + bc\mathbb{T}(\underline{L}, L) + ad\mathbb{T}(L, \underline{L}) + bd\mathbb{T}(\underline{L}, \underline{L}) \\ &= ac(L\varphi)^2 + bd(\underline{L}\varphi)^2 + (bc + ad)[(e_1\varphi)^2 + (e_2\varphi)^2] \geq 0. \end{aligned}$$

Therefore,  $\mathbb{T}_{\mu\nu}X^\mu Y^\nu \geq 0$ , which proves the positivity property.  $\square$

*Remark 2.3.* Note that this proof shows us that if  $X$  and  $Y$  are future directed timelike vectors, then

$$(7) \quad \mathbb{T}(X, Y) \geq c \sum_{i=0}^n (\partial_i \varphi)^2.$$

*2.1. Associated Energy Identity.* In the theory of hyperbolic partial differential equations, many times one will need to obtain estimates on solutions to such equations. This is sometimes done by integral identities. These are equations of the form

$$\int_{\partial\mathcal{R}} J \circ d\varphi = \int_{\mathcal{R}} K \circ d\varphi$$

where  $\mathcal{R} \subset \mathcal{M}$  is some region and  $\varphi : \mathcal{M} \rightarrow \mathbb{R}$  is a solution to the Euler-Lagrange equations associated to some Lagrangian  $\mathcal{L}$ . In general, we call sections  $J$  and  $K$  the *energy current* and *scalar current*, respectively. If this integral identity holds for all regions  $\mathcal{R} \subseteq \mathcal{M}$  and for all solutions  $\varphi$  to the Euler-Lagrange equations, then we say that  $J$  is *compatible* with the Lagrangian  $\mathcal{L}$ . For more information about the theory of energy currents and hyperbolic partial differential equations, see [2].

Now in our setting,  $\varphi$  is a solution to (5) and we can precisely define the *energy current*  $J_\mu^X$  and the *scalar current*  $K^X$  of the energy-momentum tensor such that

$$J_\mu^X[\varphi] := \mathbb{T}_{\mu\nu}[\varphi]X^\nu \quad \text{and} \quad K^X[\varphi] := {}^{(X)}\pi_{\mu\nu}\mathbb{T}^{\mu\nu}[\varphi]$$

where  $X$  is a given vector field on  $\mathcal{M}$ . Here,  ${}^{(X)}\pi^{\mu\nu}$  is the *deformation tensor* of  $X$ . Specifically we have that  ${}^{(X)}\pi^{\mu\nu} = \frac{1}{2}(\mathcal{L}_X g)_{\mu\nu}$ . In fact, we can define  $K^X$  in terms of  $J_\mu^X$ . This is called the *divergence relation*:

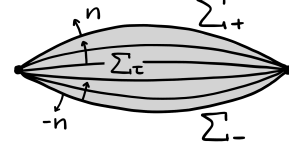
$$(8) \quad K^X[\varphi] = \nabla^\mu J_\mu^X[\varphi].$$

Note that since  ${}^{(X)}\pi_{\mu\nu} = 0$  when  $X$  is Killing, it follows that

$$\nabla^\mu J_\mu^X[\varphi] = K^X[\varphi] = 0.$$

For this reason, when  $X$  is Killing, we call  $J_\mu^X$  a *conserved current*.

Let us now suppose further that  $(\mathcal{M}, g)$  is time oriented (meaning that it is a *spacetime*). Additionally, let  $\mathcal{R} \subset \mathcal{M}$  be bounded by two homologous compact spacelike hypersurfaces  $\Sigma_+$  and  $\Sigma_-$  that have a common boundary. Using Stokes' theorem and the divergence relation, we can define the following integral identity:



$$\int_{\partial\mathcal{R}} J_\mu^X[\varphi] n_{\partial\mathcal{R}}^\mu dVol_{g|_{\partial\mathcal{R}}} = \int_{\mathcal{R}} K^X[\varphi] dVol_g$$

where  $n_{\partial\mathcal{R}}^\mu$  is the future directed unit normal on the boundary of  $\mathcal{R}$ . Since  $n_{\partial\mathcal{R}}^\mu$  is timelike (ie.  $n_\mu n^\mu = -1$ ) and since the unit normal of  $n_{\Sigma_-}^\mu$  is past directed, we get

$$\int_{\partial\mathcal{R}} J_\mu^X[\varphi] n_{\partial\mathcal{R}}^\mu = - \int_{\Sigma_+} J_\mu^X[\varphi] n_{\Sigma_+}^\mu + \int_{\Sigma_-} J_\mu^X[\varphi] n_{\Sigma_-}^\mu.$$

This means we have

$$(9) \quad \int_{\Sigma_+} J_\mu^X[\varphi] n_{\Sigma_+}^\mu = \int_{\Sigma_-} J_\mu^X[\varphi] n_{\Sigma_-}^\mu - \int_{\mathcal{R}} K^X[\varphi].$$

Thus if  $X$  is a Killing vector field, then

$$\int_{\Sigma_+} J_\mu^X[\varphi] n^\mu = \int_{\Sigma_-} J_\mu^X[\varphi] n^\mu.$$

LEMMA 2.4. *Let  $\mathcal{R} \subset \mathcal{M}$  such that its boundary is  $\Sigma_-$  and  $\Sigma_+$ , homologous spacelike hypersurfaces with common boundary. For a timelike vector field  $X$ , there exists a constant  $C_{\mathcal{R}} > 0$  that only depends on  $\mathcal{R}$  such that*

$$(10) \quad \|\varphi\|_{\dot{H}^1(\Sigma_+)} + \|n\varphi\|_{L^2(\Sigma_+)} \leq C_{\mathcal{R}} \left( \|\varphi\|_{\dot{H}^1(\Sigma_-)} + \|n\varphi\|_{L^2(\Sigma_-)} \right).$$

*Remark 2.5.* In this lemma,  $\dot{H}^1$  denotes the *homogeneous Sobolev space*. In the case of  $p = 2$ , we denote  $D^{k,2}(\Omega)$  as  $\dot{H}^k(\Omega)$ . This means that

$$\dot{H}^k(\Omega) = \{u \in L_{\text{loc}}^1(\Omega) \mid D^\ell u \in L^2(\Omega) \text{ for } |\ell| = k\} \quad \text{and} \quad \|u\|_{\dot{H}^k(\Omega)} = \left( \sum_{|\ell|=k} \int_{\Omega} |D^\ell u|^2 \right)^{\frac{1}{2}}.$$

For more information about this space and others like it, please see [7].

*Proof.* Recall that for a spacelike hypersurface  $\Sigma \subset \mathcal{M}$ , we can define the following norms for functions  $f$  defined on  $\Sigma$ :

$$\|f\|_{\dot{H}^1(\Sigma)} = \left( \int_{\Sigma} |\nabla f|^2 \right)^{1/2} \quad \text{and} \quad \|nf\|_{L^2(\Sigma)} = \left( \int_{\Sigma} |nf|^2 \right)^{1/2}.$$

Also, note that for a future directed timelike unit vector  $n$ , we can calculate

$$\begin{aligned}\mathbb{T}(n, n) &= (n\varphi)^2 - \frac{1}{2}g(n, n)|\nabla\varphi|^2 \\ &= (n\varphi)^2 - \frac{1}{2}(-1)(-(n\varphi)^2 + |\nabla\varphi|^2) \\ &= \frac{1}{2}(n\varphi)^2 + \frac{1}{2}|\nabla\varphi|^2.\end{aligned}$$

This shows us that we can study  $|\nabla\varphi|^2$  and  $|n\varphi|^2$  by looking at its associated energy momentum tensor. Luckily, (9) will help us do just that.

To formalize this idea, we will start by defining a foliation of  $\mathcal{R}$  so that

$$\mathcal{R} = \bigcup_{\tau \in [0,1]} \Sigma_\tau$$

where  $\Sigma_0 = \Sigma_-$  and  $\Sigma_1 = \Sigma_+$ . Furthermore, denote  $\mathcal{R}_\tau = \cup \Sigma_{\tau'}$  where  $\tau' \leq \tau$ . Assume that the interiors of  $\Sigma_\tau$  are level sets of smooth functions  $\phi = \tau$  on the interior of  $\mathcal{R}$  such that for some constant  $c > 0$ ,  $-g(\nabla\phi, \nabla\phi) \geq c$ . Assume that the interiors of  $\Sigma_\tau$  are level sets of smooth functions  $\phi = \tau$  on the interior of  $\mathcal{R}$  such that for some constant  $c > 0$ ,  $-g(\nabla\phi, \nabla\phi) \geq c$ .

For a timelike vector field  $X$  (to be specified later), applying (9) to the region  $\mathcal{R}_\tau$  gives us

$$\int_{\Sigma_\tau} J_\mu^X[\varphi]n^\mu = \int_{\Sigma_-} J_\mu^X[\varphi]n^\mu - \int_{\mathcal{R}_\tau} K^X[\varphi].$$

Now, since both  $X$  and  $n$  are future directed timelike vector fields, the positivity property tells us that

$$c(x) \sum_{i=0}^n (\partial_i \varphi)^2 \leq \mathbb{T}(X, n) = J_\mu^X[\varphi]n^\mu$$

for some function  $c(x) > 0$  that depends continuously on  $x \in \Sigma_\tau$ . On the other hand, since  $K^X[\varphi]$  is defined to be the contraction of the energy momentum tensor with the deformation tensor, it is a linear combination of the squares of the first derivatives of  $\varphi$ . This means that there exists a continuous function  $C(x)$  such that

$$|K^X[\varphi]| \leq C(x) \sum_{i=0}^n (\partial_i \varphi)^2.$$

But recall that  $\mathcal{R}$  is compact. So by the Extreme Value Theorem, the continuous functions  $c(X)$  and  $C(X)$  on  $\mathcal{R}$  attain maximums and minimums. Thus, there exists a constant  $c_m > 0$  dependent on  $\mathcal{R}$  such that  $c_m \leq c(x)$ , and there

exists a constant  $C_M > 0$  dependent on  $\mathcal{R}$  such that  $C(x) \leq C_M$ . This shows us that

$$|K^X[\varphi]| \leq C_M \sum_{i=0}^n (\partial_i \varphi)^2 \leq \frac{C_M}{c_m} J_\mu^X[\varphi] n^\mu =: C J_\mu^X[\varphi] n^\mu.$$

We can use this inequality to bound  $\int_{\Sigma_\tau} J_\mu^X[\varphi] n^\mu$ . Specifically,

$$- \int_{\mathcal{R}_\tau} K^X[\varphi] \leq \int_{\mathcal{R}_\tau} |K^X[\varphi]| \leq C \int_{\mathcal{R}_\tau} J_\mu^X[\varphi] n^\mu.$$

Additionally, by the coarea formula, there exists some constant  $c > 0$  dependent on  $\mathcal{R}$  such that

$$dVol_{\mathcal{R}} \leq \frac{1}{c} d\tau dVol_{\Sigma_\tau}.$$

Putting these together,

$$C \int_{\mathcal{R}_\tau} J_\mu^X[\varphi] n^\mu \leq \frac{C}{c} \int_0^\tau \int_{\Sigma_{\tau^*}} J_\mu^X[\varphi] n^\mu.$$

This means that we have

$$(11) \quad \int_{\Sigma_\tau} J_\mu^X[\varphi] n^\mu \leq \int_{\Sigma_-} J_\mu^X[\varphi] n^\mu + \tilde{C} \int_0^\tau \int_{\Sigma_{\tau^*}} J_\mu^X[\varphi] n^\mu$$

where  $\tilde{C}$  is a positive constant dependent on  $\mathcal{R}$ . Define a function  $f$  such that

$$f_X(\tau) := \int_{\Sigma_\tau} J_\mu^X[\varphi] n^\mu.$$

Rewriting (11) and remembering that  $\Sigma_- = \Sigma_0$ , we see

$$f_X(\tau) \leq f_X(0) + \tilde{C} \int_0^\tau f_X(\tau^*) d\tau^*.$$

The integral form of Grönwall's inequality tells us that the above inequality implies that

$$f_X(\tau) \leq f_X(0) e^{\tilde{C}\tau}$$

for every  $\tau \in [0, 1]$ . Let  $\tau = 1$ , and define  $\tilde{C}_{\mathcal{R}} = e^{\tilde{C}}$ . Then

$$(12) \quad f_X(1) \leq \tilde{C}_{\mathcal{R}} f_X(0).$$

Now, observe that when  $X = n$

$$f_n(\tau) = \int_{\Sigma_\tau} J_\mu^n[\varphi] n^\mu = \int_{\Sigma_\tau} T_{\mu\nu}[\varphi] n^\nu n^\mu = \int_{\Sigma_\tau} \frac{1}{2} (n\varphi)^2 + \int_{\Sigma_\tau} \frac{1}{2} |\nabla\varphi|^2 = \frac{1}{2} \|n\varphi\|_{L^2(\Sigma_\tau)}^2 + \frac{1}{2} \|\varphi\|_{\dot{H}^1(\Sigma_\tau)}^2.$$

This means that (12) states

$$\|n\varphi\|_{L^2(\Sigma_1)}^2 + \|\varphi\|_{\dot{H}^1(\Sigma_1)}^2 \leq \tilde{C}_{\mathcal{R}} \left( \|n\varphi\|_{L^2(\Sigma_0)}^2 + \|\varphi\|_{\dot{H}^1(\Sigma_0)}^2 \right).$$

To remove the squares, recall the standard algebraic inequality  $\frac{1}{\sqrt{2}}(a+b) \leq \sqrt{a^2+b^2} \leq a+b$  where  $a$  and  $b$  are two positive constants. This means that

$$\frac{1}{\sqrt{2}} \left( \|\varphi\|_{\dot{H}^1(\Sigma_1)} + \|n\varphi\|_{L^2(\Sigma_1)} \right) \leq \sqrt{\tilde{C}_{\mathcal{R}}} \left( \|\varphi\|_{\dot{H}^1(\Sigma_0)} + \|n\varphi\|_{L^2(\Sigma_0)} \right).$$

Finally, define  $C_{\mathcal{R}} = \sqrt{2\tilde{C}_{\mathcal{R}}}$ . Then,

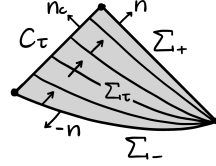
$$\|\varphi\|_{\dot{H}^1(\Sigma_+)} + \|n\varphi\|_{L^2(\Sigma_+)} \leq C_{\mathcal{R}} \left( \|\varphi\|_{\dot{H}^1(\Sigma_-)} + \|n\varphi\|_{L^2(\Sigma_-)} \right)$$

where we remember  $\Sigma_+ = \Sigma_1$  and  $\Sigma_- = \Sigma_0$ .  $\square$

Now let us return to the situation when  $(\mathcal{M}, g)$  is a Lorentzian manifold of dimension  $n+1$ , and  $\mathcal{R}$  is a compact region. In this consideration, suppose that  $\mathcal{R}$  is foliated by level sets  $\Sigma_\tau$  of a smooth function  $\phi = \tau$  such that  $0 < c < |\nabla\phi| \leq C$ . Suppose further that the foliation has a smooth null boundary  $C_\tau$ . Geometrically, consider the figure below.

A difficulty in this scenario is that null cones do not come with a natural volume form nor a natural normal vector. However, if we choose the normal to be

$$n_C = -\partial_\tau$$



then one can calculate  $g|_{C_\tau}$  to determine the induced volume form on  $C_\tau$ . Stokes theorem then states

$$(13) \quad \int_{\Sigma_-} J_\mu^X[\varphi]n^\mu - \int_{\Sigma_+} J_\mu^X[\varphi]n^\mu - \int_{C_+} J_\mu^X[\varphi]n_C^\mu = \int_{\mathcal{R}} K^X[\varphi].$$

The previous equality clearly implies the following inequality.

$$\int_{\Sigma_+} J_\mu^X[\varphi]n^\mu \leq \int_{\Sigma_-} J_\mu^X[\varphi]n^\mu - \int_{\mathcal{R}} K^X[\varphi]$$

So, we can define  $f_X$  so that

$$f_X(\tau) := \int_{\Sigma_\tau} J_\mu^X[\varphi]n^\mu.$$

Since we can still show that

$$\left| \int_{\mathcal{R}} K^X[\varphi] \right| \leq \int_0^\tau \int_{\Sigma_t} |J_\mu^X[\varphi]| d\mu_{\Sigma_t} dt \leq C \int_0^\tau f(t) dt$$

it follows from Grönwall's inequality that

$$f(\tau) \leq f(0)e^{C\tau}.$$

Using the same arguments as before, it's clear that the following theorem holds.

**THEOREM 2.6.** *Let  $\mathcal{R} \subset \mathcal{M}$  such that its boundary is  $\Sigma_-$  and  $\Sigma_+$  (homologous spacelike hypersurfaces) and  $C_+$ , a null hypersurface. For a timelike vector field  $X$ , there exists a constant  $C_{\mathcal{R}} > 0$  that only depends on  $\mathcal{R}$  such that*

$$(14) \quad \|\varphi\|_{\dot{H}^1(\Sigma_+)} + \|n\varphi\|_{L^2(\Sigma_+)} \leq C_{\mathcal{R}} \left( \|\varphi\|_{\dot{H}^1(\Sigma_-)} + \|n\varphi\|_{L^2(\Sigma_-)} \right).$$

**2.2. Well Posedness and Uniqueness.** While studying partial differential equations, questions about uniqueness and well posedness are natural to ask. It turns out, we can define uniqueness locally and within the domain of dependence for a given spacelike hypersurface. Furthermore, the wave equation on Lorentzian manifolds is well defined, resulting in well posedness.

**COROLLARY 2.7 (Local Uniqueness).** *Suppose that  $\varphi$  and  $\tilde{\varphi}$  are two sufficiently regular solutions to (5) such that  $\varphi = \tilde{\varphi}$  and  $n\varphi = n\tilde{\varphi}$  on  $\Sigma_-$ . Then,  $\varphi = \tilde{\varphi}$  in  $\mathcal{R}$ .*

*Proof.* Define  $\psi = \varphi - \tilde{\varphi}$ . Then,  $\psi$  is a sufficiently regular solution to (5) such that  $\psi = 0$  and  $n\psi = 0$  on  $\Sigma_-$ . Using the same foliation of  $\mathcal{R}$  as given in the proof above, the lemma states that for each  $\tau \in [0, 1]$ ,

$$\|\psi\|_{\dot{H}^1(\Sigma_\tau)} + \|n\psi\|_{L^2(\Sigma_\tau)} \leq C_{\mathcal{R}} \left( \|\psi\|_{\dot{H}^1(\Sigma_-)} + \|n\psi\|_{L^2(\Sigma_-)} \right) = C_{\mathcal{R}}(0) = 0.$$

Since  $\|\psi\|_{\dot{H}^1(\Sigma_\tau)}$  and  $\|n\psi\|_{L^2(\Sigma_\tau)}$  are non-negative numbers, this inequality implies that

$$\|\psi\|_{\dot{H}^1(\Sigma_\tau)} = 0 \quad \text{and} \quad \|n\psi\|_{L^2(\Sigma_\tau)} = 0$$

on  $\Sigma_\tau$ . Now using a Poincaré type inequality, we see that there is some constant  $C > 0$  such that

$$(15) \quad \|\psi\|_{H^1(\Sigma_\tau)} \leq C \left( \|\psi\|_{\dot{H}^1(\Sigma_\tau)} + \|\psi\|_{L^2(\partial\Sigma_-)} \right) = C(0) = 0.$$

Recall that  $\partial\Sigma_- = \partial\Sigma_\tau$ . Again, since  $\|\psi\|_{H^1(\Sigma_\tau)}$  is non-negative, it follows that  $\varphi = \tilde{\varphi}$  and  $n\varphi = n\tilde{\varphi}$  on  $\Sigma_\tau$ . Since this holds for every  $\tau \in [0, 1]$ , it holds in  $\mathcal{R}$ .  $\square$

**THEOREM 2.8 (Uniqueness in the Domain of Dependence).** *Let  $(\mathcal{M}, g)$  be a 4-dimensional spacetime. Suppose that  $M$  has an acausal spacelike hypersurface  $\Sigma \subset \mathcal{M}$  without a boundary. Let  $n$  be the future direction normal to  $\Sigma$ . If  $\varphi$  and  $\phi$  are sufficient regular solutions to (5) such that*

$$\varphi|_{\Sigma} = \phi|_{\Sigma} \quad \text{and} \quad n\varphi|_{\Sigma} = n\phi|_{\Sigma}$$

*then  $\varphi = \phi$  in  $D(\Sigma)$  the domain of dependence of  $\Sigma$ .*

**Remark 2.9.** The *domain of dependence* of  $\Sigma$  is the collection of points  $p \in \mathcal{M}$  such that every inextendible causal curve through  $p$  also intersects  $\Sigma$ . This is exactly the *Cauchy development* of  $\Sigma$ . As the Cauchy development,  $D(\Sigma)$  is globally hyperbolic and  $\Sigma \subset D(\Sigma)$  is a Cauchy hypersurface. This

means that if  $\mathcal{M}$  is a Schwarzschild spacetime (and thus is globally hyperbolic), then  $D(\Sigma) = \mathcal{M}$ .

**THEOREM 2.10** (Well Posedness). *Let  $(\mathcal{M}, g)$  be a globally hyperbolic spacetime and let  $\Sigma$  be a Cauchy hypersurface of  $\mathcal{M}$ . Suppose that  $\varphi \in H_{loc}^s(\Sigma)$  and  $\varphi' \in H_{loc}^{s-1}(\Sigma)$ . Then, there exists a unique  $\psi$  such that  $\square_g \psi = 0$  in  $\mathcal{M}$  and for all spacelike hypersurfaces  $\mathcal{S} \subset \mathcal{M}$ ,*

$$\psi|_{\mathcal{S}} \in H_{loc}^s(\mathcal{S}) \quad n_{\mathcal{S}}\psi|_{\mathcal{S}} \in H_{loc}^{s-1}(\mathcal{S}) \quad \psi|_{\Sigma} = \varphi \quad n\psi|_{\Sigma} = \varphi'.$$

*Remark 2.11.* This theorem states that given some suitable initial data about an event in a Schwarzschild spacetime, there exists a unique solution to the wave equation (condition 1) that is “well behaved” (condition 2) and determines our initial data (condition 3). We will use the well posedness of the scalar wave equation as a general fact while proving its uniform boundedness of the Schwarzschild spacetime.

For the sake of length, the proofs of Theorem 2.8 and Theorem 2.10 are omitted from the paper. Please see [1] for a further explanation.

### 3. The Redshift Effect

An important result in mathematical general relativity was the analytic formalization of the *redshift effect*. In physics, “redshift” is a term usually used to discuss the wavelength of light increasing with respect to an observer while the object emitting the light is moving away. Many people know this redshift phenomenon as the Doppler effect, but it has many different applications throughout geometric optics. In fact in the early 2000s, this property made its way to the mathematical general relativity community, who at the time was in need of new techniques for studying non-Killing vector fields in spacetimes. This section will discuss the heuristics of the redshift effect and then will formalize the redshift dampening effect.

**3.1. Global and Local Redshift.** Consider the following thought experiment. There are two observers,  $A$  and  $B$ , near a black hole, and  $A$  is closer to the event horizon than  $B$  is. We will consider two different scenarios for the observers, and each one will lead to a different formalization of the redshift effect.

**Global Redshift.** Suppose that  $A$  crosses the event horizon and  $B$  does not. Furthermore right before  $A$  crosses, she sends a signal of constant frequency to  $B$ .  $B$  will receive this signal, but if she measures the frequency of it, she will find that it has “shifted to the red.” This phenomenon is called the *global redshift effect*, and it follows from the fact that the proper time of  $A$  before

crossing the event horizon is finite while the proper time of  $B$  is infinite. Its use in general relativity dates back to Oppenheimer and Snyder when they studied the appearance of a collapsing star to faraway observers in [12].

**Local Redshift.** Now, suppose that  $A$  crosses the event horizon and so does  $B$  but at a later time. Again,  $A$  sends a signal of constant frequency to  $B$ . This time, if  $B$  measures the frequency of the signal, she will find that it has “shifted to the red” by a factor depending exponentially on the difference in time that  $A$  and  $B$  crossed the horizon. This exponential factor is called the *surface gravity* of the event horizon, and its value determines the *local redshift effect*. This is the phenomenon Dafermos and Rodnianski formalize in [4] for the Schwarzschild spacetime geometry.

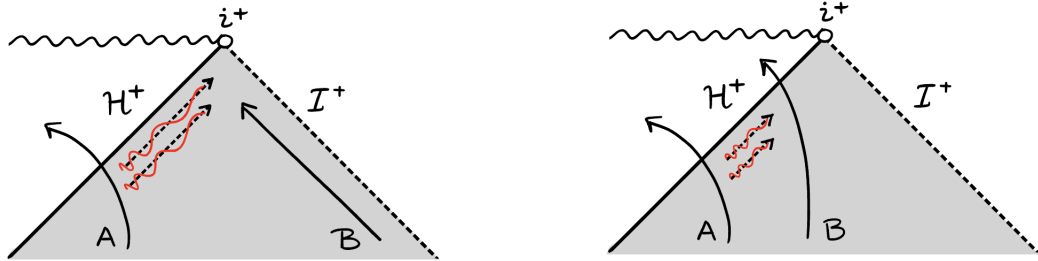


Figure 1. The figure on the left represents *global redshift* while the figure on the right represents *local redshift*.

**3.2. Redshift in Geometric Optics.** In order to formalize the redshift phenomenon, we first need to define the *surface gravity* of a black hole. This term will let us calculate the factor of dampening happening locally in the redshift effect.

**THEOREM 3.1 (Surface Gravity).** *Let  $(\mathcal{M}, g)$  be a black hole spacetime. Then, on the future event horizon of  $\mathcal{M}$ , there exists a constant  $\kappa > 0$  and vector field  $T$  such that*

$$\nabla_T T = \kappa T.$$

**COROLLARY 3.2 (Redshift Effect).** *For a light signal sent at  $t = 0$  from near the future event horizon of  $\mathcal{M}$ , the frequency of the signal is dampened by a factor of  $e^{-\kappa t}$ .*

The formalization of these statements was a major breakthrough in the mathematical study of general relativity. While they are vital to the construction of useful vector fields, the proofs are long and out of scope for the length of this paper. Please read [4] for a detailed discussion.

#### 4. Proof of Theorem 1.1

Uniform boundedness for solutions to the scalar wave equation are achieved through the use of Sobolev inequalities for the solution's energy identities. Recall the main theorem from Section 1.

**THEOREM 1.1.** *Let  $(M, g)$  be a Schwarzschild spacetime and let  $\varphi$  be the smooth solution to the wave equation that satisfies some initial data  $f \in H_{\text{loc}}^m(\Sigma_0)$  given on an initial data set  $(\Sigma_0, \mathfrak{g}, \kappa)$ . Then, there exists a  $C > 0$  such that for all events  $p$  in exterior Schwarzschild region (including the horizons)*

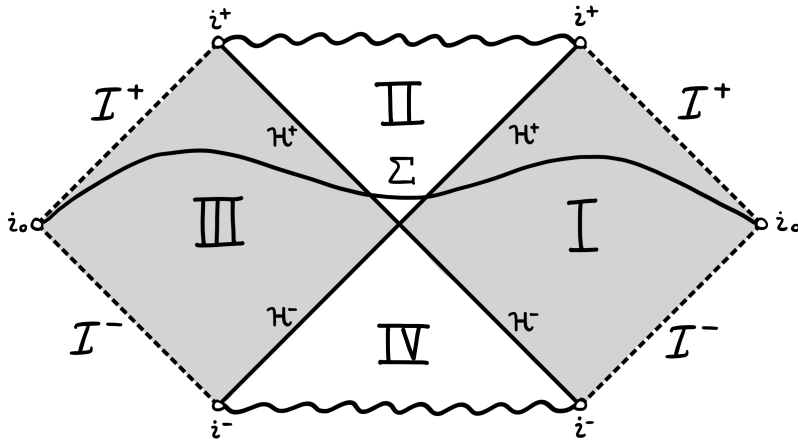
$$|\varphi(p)| \leq C.$$

The proof of this theorem is split into two main steps: pointwise bounds away from the event horizon and pointwise bounds including the event horizon. The first steps will use tools outlined in Sections 2 while the latter will require the non-Killing vector field discussed in Section 3 and a bootstrap argument. Before looking at these steps, however, we will first study the geometric nature of the problem.

4.1. *Schwarzschild Geometry.* Recall from Section 1 that the maximally extended Schwarzschild metric has Kruskal coordinates  $(T, R, \theta, \phi)$  and takes the form

$$g = \frac{32M^3}{r} e^{-r/2M} (-dT^2 + dR^2) + r^2 (d\theta^2 + \sin^2 \theta d\phi^2).$$

Furthermore, since this spacetime is spherically symmetric,  $SO(3)$  acts by isometry. Thus, we can define  $\mathcal{Q} = \mathcal{M}/SO(3)$  and the canonical projection  $\pi : \mathcal{Q} \rightarrow \mathbb{R}^{1+1}$  such that the image of  $\mathcal{Q}$  is compact in  $\mathbb{R}^{1+1}$ . This is the *Penrose diagram* for the maximally extended Schwarzschild metric.



In the diagram, Region I is the *exterior region*, and here  $\{r > 2M, R > T\}$ . This region is isometric to (2). Region II is the *black hole region* where  $\{r < 2M, T < \sqrt{1 + R^2}\}$ . Region III is a *parallel universe*, or the second region isometric to (2) where  $\{r > 2M, R > -T\}$ . Finally, region IV is a *white hole* where  $\{r < 2M, T > -\sqrt{1 + R^2}\}$ .

While the parallel universe and white hole regions sound like science fiction, it is important to note that they arise from the assumption that the spacetime is spherically symmetric. If we wish to generalize our methods to non spherically symmetric spacetimes, then we need to find a way to not include these regions. To do so, consider a different extended Schwarzschild coordinate system: one that covers Regions I and II separately from Regions III and IV. This is called the *Lemaitre coordinate system*, and Regions III and IV are interpreted as being the time reversal of Regions I and II.

Let  $(t, r, \theta, \phi)$  denote the standard Schwarzschild coordinates. Consider the change in coordinates where

$$t^* = t + 2M \log(r - 2M).$$

Then, on  $r > 2M$  the metric in coordinates  $(t^*, r, \theta, \phi)$  takes the form

$$(16) \quad g = - \left(1 - \frac{2M}{r}\right) (dt^*)^2 + \frac{4M}{r} dt^* dr + \left(1 + \frac{2M}{r}\right) dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2).$$

From this equation, we can calculate

$$\begin{aligned} g_{t^*t^*} &= - \left(1 - \frac{2M}{r}\right) & g_{t^*r} &= \frac{2M}{r} & g_{rr} &= 1 + \frac{2M}{r} & g_{AB} &= g_{AB} \\ g^{t^*t^*} &= - \left(1 + \frac{2M}{r}\right) & g^{t^*r} &= \frac{2M}{r} & g^{rr} &= 1 - \frac{2M}{r} & g^{AB} &= g^{AB} \end{aligned}$$

The remaining metric components are zero. Additionally, we can express the covariant wave operator in this metric as

$$(17) \quad \square_g \varphi = - \frac{1}{f^2} \partial_{t^*}^2 \varphi + \frac{\nabla_i f}{f} \nabla^i \varphi + \Delta \varphi \quad \text{where} \quad f = \sqrt{1 - \frac{2M}{r}}.$$

It's important to note that this coordinate system is well behaved when  $r \geq 2M$  and it is isometric to Kruskal coordinates in Regions I and II. We will use these coordinates in the theorem, but we need to understand why we can restrict our analysis to Regions I and II.

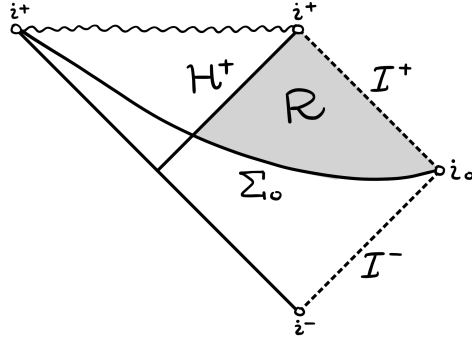
First, the Schwarzschild spacetime  $\mathcal{M}$  is spherically symmetric and  $\Sigma_0$  is a Cauchy hypersurface, so we can assume that  $\Sigma_0$  is also spherically symmetric. Additionally, observe that in the figure above, future null infinity  $\mathcal{I}$  contains

two disjoint regions: the future null infinity of Region I,  $\mathcal{I}_I^+$ , and the future null infinity of Region III,  $\mathcal{I}_{III}^+$ . This is also true for the past null infinity. This means that without loss of generality we can limit ourselves to the region

$$\mathcal{D} = \text{clos}(J^-(\mathcal{I}_I^+) \cap J^+(\mathcal{I}_I^-))$$

since one local Schwarzschild metric (2) covers this region. Since we will only refer to Region I for the remainder of the discussion, we will drop the subscript notation. However, if clarification is needed, it will be noted again. Since evolution of the wave equation is symmetric in time, we may also assume that  $\Sigma_0$  does not intersect the past event horizon  $\mathcal{H}^-$ . Recall that the *past event horizon* is defined to be the past boundary of the future of the past null infinity.

We can further our limitations more. The wave equation is symmetric in time, so we can focus our study of the behavior of  $\varphi$  to  $\mathcal{R} := J^-(\mathcal{I}^+) \cap J^+(\Sigma_0)$ . Recall that the chronological past is bounded by the speed of light. Thus, the domain of dependence for this region is  $\Sigma_0 \cap \mathcal{D}$ , and  $\varphi$  is determined by  $f|_{\Sigma_0 \cap \mathcal{D}}$ . For this reason, it will suffice to prove the theorem assuming that  $\Sigma_0$  is completely contained in exterior region and black hole region of the spacetime. Thus, we can use Lemaitre coordinates for our spacetime. The above reasoning means that we can consider the following figure.



Observe that we can equivalently define  $\mathcal{R} = \{t^* \geq 0, r \geq 2M\}$ . Let  $T = \partial_{t^*}$ . Since  $T$  is Killing,  $K^T[\varphi] = 0$ . Additionally, define

$$n = n^{t^*} \partial_{t^*} + n^r \partial_r \quad \text{where} \quad n^{t^*} = \sqrt{-g^{t^*t^*}} \quad \text{and} \quad n^r = -\frac{g^{t^*r}}{\sqrt{-g^{t^*t^*}}}.$$

This is a unit vector field. To see this, observe that

$$\begin{aligned} g(n, n) &= g_{\mu\nu} n^\mu n^\nu = g_{t^*t^*} (n^{t^*})^2 + 2g_{t^*r} n^{t^*} n^r + g_{rr} (n^r)^2 \\ &= -1 + \frac{4M^2}{r^2} - \frac{8M^2}{r^2} + \frac{4M^2}{r^2} \\ &= -1. \end{aligned}$$

Now, we can calculate

$$\begin{aligned}
g(T, n) &= g(T, n^{t^*} T + n^r \partial_r) \\
&= n^{t^*} g(T, T) + n^r g(T, \partial_r) \\
&= n^{t^*} \left(-1 + \frac{2M}{r}\right) + n^r \left(\frac{2M}{r}\right) \\
&= -n^{t^*} + \frac{2M}{r} (n^{t^*} + n^r) \\
&= -\sqrt{1 + \frac{2M}{r}} + \frac{2M}{r} \left(\sqrt{1 + \frac{2M}{r}} - \frac{2M/r}{\sqrt{1 + 2M/r}}\right) \\
&= -\left(1 + \frac{2M}{r}\right)^{-1/2}
\end{aligned}$$

If we let  $f = 1 + (2M/r)$ , then we can further calculate

$$\begin{aligned}
(18) \quad J_\mu^T[\varphi]n^\mu &= \mathbb{T}(T, n) \\
&= (T\varphi) \left((n^{t^*} T + n^r R)\varphi\right) - \frac{1}{2}g(T, n)|\nabla\varphi|^2 \\
&= \sqrt{f} (\partial_{t^*}\varphi)^2 + \frac{1-f}{\sqrt{f}} (\partial_{t^*}\partial_r)(\varphi) + \frac{1}{2\sqrt{f}} |\nabla\varphi|^2 \\
&= \frac{1}{2} \left(\sqrt{f} (\partial_{t^*}\varphi)^2 + \frac{2-f}{\sqrt{f}} (\partial_r\varphi)^2 + \frac{1}{\sqrt{f}} |\nabla\varphi|^2\right) \\
&= \frac{1}{2\sqrt{1 + \frac{2M}{r}}} \left[\left(1 + \frac{2M}{r}\right) (\partial_{t^*}\varphi)^2 + \left(1 - \frac{2M}{r}\right) (\partial_r\varphi)^2 + |\nabla\varphi|^2\right]
\end{aligned}$$

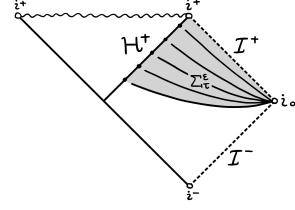
Since  $T$  is a time translation Killing field, we can define  $\phi_\tau$  to be the local flow of  $T$ . We assume that the hypersurface  $\Sigma_0$  is the level set  $\{t^* = 0\}$  of  $\phi_\tau$ . Since Schwarzschild spacetimes are globally hyperbolic, we can define spacelike hypersurfaces  $\Sigma_\tau$  foliating  $\mathcal{R}$  as the pushforward of  $\Sigma_0$ :

$$\mathcal{R} = \bigcup_{\tau \geq 0} \phi_\tau(\Sigma_0).$$

Additionally,  $T$  is timelike in  $\mathcal{R} \setminus \mathcal{H}^+$  and it is lightlike on  $\mathcal{H}^+$ , making it a Killing horizon. Recall from Section 3 the black hole has surface gravity  $\kappa$ : specifically on  $\mathcal{H}^+ = \{r = 2M\}$ ,

$$\kappa = \frac{1}{4M} \quad \text{and} \quad \nabla_T T = \kappa T \quad \text{when} \quad T = \partial_{t^*}.$$

4.2. *Pointwise Bounds away from the Horizon.* We want to use Theorem 2.6, but it only holds when  $J^I$  is non-degenerate. The calculation in (18) shows us that this happens exactly at  $r = 2M$ . So, we will first prove Theorem 1.1 for hypersurfaces that are sufficiently disjoint from the event horizon. To construct these hypersurfaces, first let  $S_\tau = \{t^* = \tau\} \cap \{r = 2M\}$  as the boundary of  $\Sigma_\tau$ .

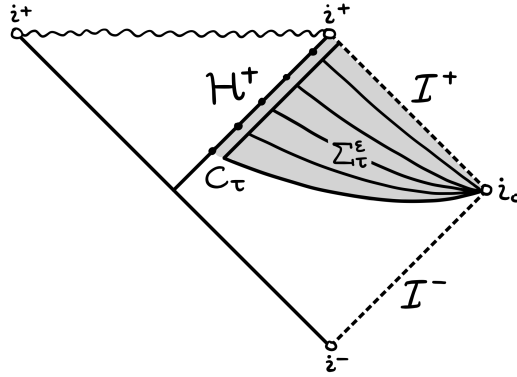


Note that  $S_\tau$  is compact and that each  $\Sigma_\tau$  is asymptotically flat. This means that there exists a compact  $K_\tau = \{r \leq R\} \cap \Sigma_\tau$  for  $R$  large such that  $\Sigma_\tau \setminus K_\tau$  is covered globally by Euclidean coordinates. Let  $\varepsilon > 0$ , and denote

$$B_0^\varepsilon = \Sigma_0 \cap \{r < 2M + \varepsilon\}$$

to be an  $\varepsilon$  neighborhood of  $S_0$ . Define  $\Sigma_0^\varepsilon := \Sigma_0 \setminus B_0^\varepsilon$ . Then,

$$\Sigma_\tau^\varepsilon = \phi_\tau(\Sigma_0^\varepsilon) = \{t^* = \tau\} \cap \{r \geq 2M + \varepsilon\}.$$



PROPOSITION 4.1 (Degenerate Energy Bounds). *For all  $\varepsilon > 0$  there exists  $C_\varepsilon > 0$  that only depends on  $\varepsilon$  such that for all  $\tau$  and for all solutions  $\varphi$  of the homogeneous wave equation*

$$(19) \quad \sup_{\Sigma_\tau^\varepsilon} |\varphi| \leq C_\varepsilon.$$

*Proof.* Consider  $\mathcal{R}_\tau^\varepsilon \subset \mathcal{R}$  such that

$$\mathcal{R}_\tau^\varepsilon = \bigcup_{t^*=0}^{\tau} \Sigma_{t^*}^\varepsilon$$

Then,  $\partial(\mathcal{R}_\tau^\varepsilon) = \Sigma_\tau^\varepsilon \cup C_\tau \cup \Sigma_0^\varepsilon$  where  $C_\tau = \{r = 2M + \varepsilon, 0 \leq t^* \leq \tau\}$ . Furthermore,  $T$  is Killing on  $\mathcal{R}_\tau^\varepsilon$ . This means that  $K^T = 0$  and (13) states

$$(20) \quad \int_{\Sigma_\tau^\varepsilon} J_\mu^T[\varphi] n_{\Sigma_\tau^\varepsilon}^\mu + \int_{C_\tau} J_\mu^T[\varphi] n_{C_\tau}^\mu = \int_{\Sigma_0} J_\mu^T[\varphi] n_{\Sigma_0}^\mu.$$

Since  $T$  and each normal vector are all timelike on  $\mathcal{R}_\tau^\varepsilon$ , it follows from the positivity property that there is a constant  $c \geq 0$  such that

$$J^T[\varphi]n \geq c((\partial_{t^*}\varphi)^2 + (\partial_r\varphi)^2 + |\nabla\varphi|^2).$$

On the other hand, (20)

$$\int_{\Sigma_\tau^\varepsilon} J_\mu^T[\varphi] n_{\Sigma_\tau^\varepsilon}^\mu \leq \int_{\Sigma_0} J_\mu^T[\varphi] n_{\Sigma_0}^\mu.$$

Putting these two inequalities together, it follows that

$$(21) \quad \|\varphi\|_{\dot{H}^1(\Sigma_\tau^\varepsilon)}^2 = \int_{\Sigma_\tau^\varepsilon} ((\partial_{t^*}\varphi)^2 + (\partial_r\varphi)^2 + |\nabla\varphi|^2) \leq c_\varepsilon \int_{\Sigma_0} J_\mu^T[\varphi] n_{\Sigma_0}^\mu \leq C_1(\varepsilon).$$

In the above, the last inequality holds since  $\varphi \in H_{\text{loc}}^m(\Sigma_0)$  for a sufficiently large  $m$ . The value of  $C_1(\varepsilon)$  only depends on  $\varepsilon$  since the integral is uniformly bounded and  $c_\varepsilon$  depends only on  $\varepsilon$ .

We also know that the function  $T\varphi$  also satisfies the homogeneous wave equation, meaning that  $\square_g(T\varphi) = 0$ , since  $T$  is Killing. Note that (21) holds for any solution to the homogeneous wave equation, so it follows that

$$(22) \quad \|T\varphi\|_{\dot{H}^1(\Sigma_\tau^\varepsilon)}^2 = \int_{\Sigma_\tau^\varepsilon} ((\partial_{t^*}^2\varphi)^2 + (\partial_r\partial_{t^*}\varphi)^2 + |\nabla\partial_{t^*}\varphi|^2) \leq C_2(\varepsilon).$$

Recall that in a Schwarzschild spacetime, we can write the covariant wave operator as

$$\square_g\varphi = -\frac{1}{f^2}\partial_{t^*}^2\varphi + \frac{\nabla_i f}{f}\nabla^i\varphi + \Delta\varphi \quad \text{where} \quad f = \sqrt{1 - \frac{2M}{r}}.$$

Since  $\square_g\varphi = 0$ , it follows that

$$\Delta\varphi = \frac{1}{f^2}\partial_{t^*}^2\varphi - \frac{\nabla_i f}{f}\nabla^i\varphi.$$

Furthermore, since  $r \geq 2M + \varepsilon$ ,  $f$  is uniformly bounded. This means that there exists two constants  $D_1 > 0$  and  $D_2 > 0$  such that

$$\|\Delta\varphi\|_{L^2(\Sigma_\tau^\varepsilon)} \leq D_1 \|\partial_{t^*}^2\varphi\|_{L^2(\Sigma_\tau^\varepsilon)} + D_2 \|\nabla\varphi\|_{L^2(\Sigma_\tau^\varepsilon)}.$$

Using (21) and (22) we can further bound

$$\|\Delta\varphi\|_{L^2(\Sigma_\tau^\varepsilon)} \leq D_1 C_2(\varepsilon) + D_2 C_1(\varepsilon) =: C_3(\varepsilon).$$

By the homogeneous Sobolev embedding theorem, there exists a constant  $D_3 > 0$  such that for each event  $p \in \Sigma_\tau^\varepsilon$

$$\begin{aligned} |\varphi(p)| &\leq \|\varphi\|_{L^\infty(\Sigma_\tau^\varepsilon)} \leq D_3 (\|\nabla\varphi\|_{L^2(\Sigma_\tau^\varepsilon)} + \|\nabla^2\varphi\|_{L^2(\Sigma_\tau^\varepsilon)}) \\ &\leq D_4 (\|\nabla\varphi\|_{L^2(\Sigma_\tau^\varepsilon)}^2 + \|\nabla^2\varphi\|_{L^2(\Sigma_\tau^\varepsilon)}^2) \\ &\leq D_4 (C_1(\varepsilon) + \|\partial_{t^*}^2\varphi\|_{L^2(\Sigma_\tau^\varepsilon)}^2 + \|\Delta\varphi\|_{L^2(\Sigma_\tau^\varepsilon)}^2) \\ &\leq D_4 (C_1(\varepsilon) + C_2(\varepsilon) + C_3(\varepsilon)) =: C_\varepsilon. \end{aligned}$$

Note that the second inequality follows from Young's inequality while the remaining inequalities follow from (21) and (22). This proves our theorem.  $\square$

4.3. *Pointwise Bounds including the Horizon.* If we let  $\varepsilon \rightarrow 0$  then  $\Sigma_\tau^\varepsilon \rightarrow \Sigma_\tau$ . However, as we noted before, there is a degeneracy in the energy current at  $r = 2M$ . Previously, we went around this by considering  $\mathcal{R}_\tau^\varepsilon$ , a region sufficiently disjoint from the event horizon. We now want to consider regions that include the event horizon. To do this, we need to construct a well behaved energy current of a non-Killing vector field. This is done through the use of the local redshift effect discussed in Section 3.

PROPOSITION 4.2 (Local Redshift Vector Field). *There exists a vector field  $N$  such that*

- (1)  $(\phi_\tau)_*N = N$  (meaning  $N$  is  $t^*$  independent)
- (2)  $N$  is future directed timelike
- (3)  $N = T$  on  $\Sigma_0 \setminus B$  for some compact  $B$
- (4) there exists an  $\varepsilon > 0$  and  $C > 0$  such that on  $\Sigma_\tau \setminus \Sigma_\tau^\varepsilon$

$$(23) \quad K^N[\varphi] \geq C J^N[\varphi] N$$

*Proof.* To prove this proposition, we will first prove that there exists a timelike vector field  $N_0$  along  $\Sigma_0 \setminus \Sigma_0^{2\varepsilon}$  such that (23) holds. Then, we will extend this vector field so that the proposition holds.

To construct such an  $N_0$ , consider  $T$  along  $S_0$ . Recall that  $T = \partial_{t^*}$  is null here. Now, consider another null vector field

$$Y_0 = 2\partial_{t^*} - 2\partial_r$$

along  $S_0$ . This is in fact null since

$$g(Y_0, Y_0) = 4(g_{t^*t^*} - 2g_{rt^*} + g_{rr}) = 4(0) = 0.$$

Furthermore, observe the following computation.

$$g(T, Y_0) = 2g(\partial_{t^*}, \partial_{t^*}) - 2g(\partial_{t^*}, \partial_r) = 2\left(-1 + \frac{2M}{r}\right) - 2\left(\frac{2M}{r}\right) = -2$$

Now, we want to extend  $Y_0$  off of the horizon to  $Y$  such that for some large  $\sigma > 0$ ,

$$\nabla_Y Y = -\sigma(T + Y) \quad \text{on } S_0.$$

To do this, consider

$$Y = \left(2 + k_1 \left(1 - \frac{2M}{r}\right)\right) \partial_{t^*} + \left(-2 + k_2 \left(1 - \frac{2M}{r}\right)\right) \partial_r$$

where

$$k_1 = 1 + 3M\sigma \quad \text{and} \quad k_2 = 2 - 2M\sigma.$$

We now define  $N_0$  as follows.

$$N_0 := T + Y = \left(3 + k_1 \left(1 - \frac{2M}{r}\right)\right) \partial_{t^*} + \left(-2 + k_2 \left(1 - \frac{2M}{r}\right)\right) \partial_r$$

We will now show that on  $\Sigma_0 \setminus \Sigma_0^{2\epsilon}$  there exists some constant  $C > 0$  such that

$$K^{N_0}[\varphi] \geq C J^{N_0}[\varphi] N_0.$$

First note that since  $T$  is Killing,  $K^T[\varphi] = 0$ . So,  $K^{N_0}[\varphi] = K^Y[\varphi]$ . Second observe that

$$J^{N_0}[\varphi] N_0 = \mathbb{T}(N_0, N_0) = \mathbb{T}(T, T) + 2\mathbb{T}(T, Y) + \mathbb{T}(Y, Y).$$

Second, since  $T$  is Killing,  $K^T[\varphi] = 0$ . So,  $K^{N_0}[\varphi] = K^Y[\varphi]$ . Recall that  $K^Y[\varphi] = {}^{(Y)}\pi_{\mu\nu} \mathbb{T}^{\mu\nu}$ , and in local coordinates note that we can write

$${}^{(Y)}\pi_{\mu\nu} = D_\mu X_\nu + D_\nu X_\mu.$$

This allows us to perform the following calculation.

$$\begin{aligned} K^Y[\varphi] &= {}^{(Y)}\pi_{\mu\nu} \mathbb{T}^{\mu\nu} \\ &= {}^{(Y)}\pi_{t^*t^*} \mathbb{T}^{t^*t^*} + 2{}^{(Y)}\pi_{t^*r} \mathbb{T}^{t^*r} + {}^{(Y)}\pi_{rr} \mathbb{T}^{rr} \\ &= (0) \mathbb{T}^{t^*t^*} + \frac{4Mk_1}{r^2} \mathbb{T}^{t^*r} + \frac{4Mk_2}{r^2} \mathbb{T}^{t^*r} \\ &= \frac{\kappa}{r} (k_1 \mathbb{T}^{t^*r} + k_2 \mathbb{T}^{rr}) \\ &\geq \frac{1}{4} (\mathbb{T}(Y, Y)\kappa + \mathbb{T}(T, Y)\sigma + \mathbb{T}(T, T)\sigma) \\ &\geq C J^{N_0}(\varphi) N_0 \end{aligned}$$

We will now extend  $N_0$ . Let  $\chi$  be the cutoff function such that in  $\Sigma_0^\epsilon \setminus \Sigma_0^{2\epsilon}$ ,  $\chi = 1$ , and elsewhere  $\chi = 0$ . Then, since  $\mathcal{I}^+$  is convex, define a future timelike vector field  $\tilde{N}_0$  such that

$$\tilde{N}_0 = \chi N_0 + (1 - \chi)T.$$

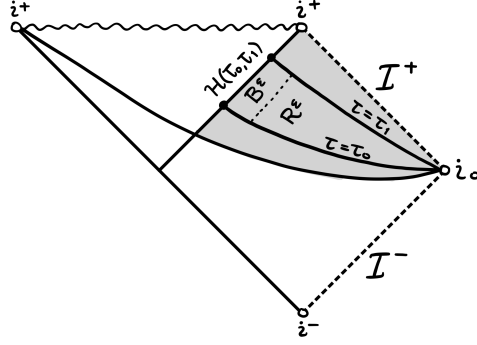
By definition,  $\tilde{N}_0$  satisfies (23) on  $\Sigma_0 \setminus \Sigma_0^\epsilon$ . Define  $N = \phi_\tau(\tilde{N}_0)$  to be the pushforward of  $\tilde{N}_0$  along integral curves of  $T$ . Clearly,  $N$  satisfies the conditions of the proposition.  $\square$

*Proof of Theorem 1.1.* Let  $0 \leq \tau_0 < \tau_1 \leq \tau$ . Denote  $\Sigma_\tau = \Sigma(\tau)$ . Consider the regions

$$\mathcal{R}(\tau_0, \tau_1) = \bigcup_{t^*=\tau_0}^{\tau_1} \Sigma(t^*) \quad \text{and} \quad \mathcal{H}(\tau_0, \tau_1) = \mathcal{H}^+ \cap \{\tau_0 \leq t^* \leq \tau_1\}.$$

Then for the region defined to the right, (13) states

$$(24) \quad \int_{\mathcal{H}(\tau_0, \tau_1)} J^N[\varphi]n + \int_{\Sigma(\tau_1)} J^N[\varphi]n + \int_{\mathcal{R}(\tau_0, \tau_1)} K^N[\varphi] = \int_{\Sigma(\tau_0)} J^N[\varphi]n$$



Recall that we can split  $\mathcal{R}(\tau_0, \tau_1)$  into

$$\mathcal{R}(\tau_0, \tau_1) = \bigcup_{t^*=\tau_0}^{\tau_1} (\Sigma(t^*) \setminus \Sigma^\epsilon(t^*)) \cup \bigcup_{t^*=\tau_0}^{\tau_1} \Sigma^\epsilon(t^*).$$

Denote  $B^\epsilon(\tau_0, \tau_1)$  as the first union and  $\mathcal{R}^\epsilon(\tau_0, \tau_1)$  as the second union. Then we can expand the integral over the interior region as

$$(25) \quad \int_{\mathcal{R}(\tau_0, \tau_1)} K^N(\varphi) = \int_{B^\epsilon(\tau_0, \tau_1)} K^N[\varphi] + \int_{\mathcal{R}^\epsilon(\tau_0, \tau_1)} K^N[\varphi].$$

Combining the integral identities (24) and (25) gives us

$$\int_{\mathcal{H}(\tau_0, \tau_1)} J^N[\varphi]n + \int_{\Sigma(\tau_1)} J^N[\varphi]n + \int_{B^\epsilon(\tau_0, \tau_1)} K^N[\varphi] = \int_{\Sigma(\tau_0)} J^N[\varphi]n - \int_{\mathcal{R}^\epsilon(\tau_0, \tau_1)} K^N[\varphi].$$

Since the energy current is non-negative

$$(26) \quad \int_{\Sigma(\tau_1)} J^N[\varphi]n + \int_{B^\epsilon(\tau_0, \tau_1)} K^N[\varphi] \leq \int_{\Sigma(\tau_0)} J^N[\varphi]n + \int_{\mathcal{R}^\epsilon(\tau_0, \tau_1)} |K^N[\varphi]|.$$

Recall that the coarea formula states that for a real valued Lipschitz function  $u : \mathbb{R}^n \rightarrow \mathbb{R}$  and domain  $\Omega \subset \mathbb{R}^n$

$$\int_{\Omega} |\nabla u| = \int_{\mathbb{R}} H_{n-1}(u^{-1}(t))$$

where  $H_{n-1}$  is the  $(n-1)$ -dimensional Hausdorff measure. Applied to the energy currents of  $N$ , we have

$$(27) \quad \int_{B^\varepsilon(\tau_0, \tau_1)} K^N[\varphi] = \int_{\tau_0}^{\tau_1} \int_{B^\varepsilon(t^*)} J^N[\varphi]n \, dt^* d\tau$$

Then, by Proposition 4.2 there exists a vector field  $N$  such that for some  $C_1 > 0$

$$C_1 \int_{\tau_0}^{\tau_1} \int_{B^\varepsilon(t^*)} J^N[\varphi]n \, dt^* d\tau \leq \int_{B^\varepsilon(\tau_0, \tau_1)} K^N[\varphi].$$

We can further bound this term from below by considering

$$(28) \quad -C_1 \int_{\tau_0}^{\tau_1} \int_{\Sigma(t^*)} J^N[\varphi]n \, dt^* d\tau \leq C_1 \int_{\tau_0}^{\tau_1} \int_{B^\varepsilon(t^*)} J^N[\varphi]n \, dt^* d\tau.$$

On the other hand, specifically on the complement of  $B^\varepsilon(\tau_0, \tau_1)$ , we have

$$\int_{\mathcal{R}^\varepsilon(\tau_0, \tau_1)} |K^N[\varphi]| \leq C_1 \int_{\tau_0}^{\tau_1} \left( \int_{\Sigma^\varepsilon(t^*)} J^N[\varphi]n \right).$$

Since  $N = T$  on  $\Sigma^{\varepsilon'}(0)$  for some  $\varepsilon' > 0$ , we have

$$\int_{\Sigma^\varepsilon(t^*)} J^N(\varphi)n \leq C_2 \int_{\Sigma(t^*)} J^T[\varphi]n \leq \int_{\Sigma(0)} J^T[\varphi]n =: D.$$

Note that the last inequality follows from analysis done in the proof of Proposition 4.1. So, putting all the above together, we derive

$$(29) \quad \int_{\mathcal{R}^\varepsilon(\tau_0, \tau_1)} |K^N[\varphi]| \leq C_1 C_2 (\tau_1 - \tau_0) D =: C_3$$

Combining (26), (28), and (29) gives us

$$(30) \quad \int_{\Sigma(\tau_1)} J^N[\varphi]n - C_1 \int_{\tau_0}^{\tau_1} \int_{\Sigma(t^*)} J^N[\varphi]n \, dt^* d\tau \leq \int_{\Sigma(\tau_0)} J^N[\varphi]n + C_3.$$

Now define the function  $f$  by the following.

$$f(\tau) := \int_{\Sigma(\tau)} J^N(\varphi)n.$$

This means we can rewrite (30) as

$$f(\tau_1) - C_1 \int_{\tau_0}^{\tau_1} f(t^*) \leq f(\tau_0) + C_3.$$

Simple rearranging leads to

$$(31) \quad f(\tau_1) \leq f(\tau_0) + C_3 + C_1 \int_{\tau_0}^{\tau_1} f(t^*).$$

Notice that (31) holds for all  $0 \leq \tau_0 < \tau_1 \leq \tau$ . So, we can take  $\tau_0 = 0$  and  $\tau_1 = \tau$ . Then,

$$(32) \quad f(\tau) \leq f(0) + C_3 + C_1 \int_0^\tau f(t^*)$$

By Grönwall's inequality

$$f(\tau) \leq f(0)e^{\tau C_4}.$$

Using the same argument as in Section 2, it follows that

$$(33) \quad \|\varphi\|_{\dot{H}^1(\Sigma_\tau)} + \|n\varphi\|_{L^2(\Sigma_\tau)} \leq C_5 \left( \|\varphi\|_{\dot{H}^1(\Sigma_0)} + \|n\varphi\|_{L^2(\Sigma_0)} \right).$$

for some constant  $C_5 > 0$  independent of  $\tau$ .

Recall that the Schwarzschild metric is spherically symmetric and asymptotically flat. This means that every hypersurface is also spherically symmetric and asymptotically flat. So by using a localized Sobolev embedding, we can find a constant  $C_6 > 0$  such that for every  $p \in \mathcal{R}$  and every  $\tau \geq 0$

$$(34) \quad \sup_{\Sigma_\tau} |\varphi(p)| \leq C_6 \left( \|\varphi\|_{\dot{H}^2(\Sigma_\tau)} + \|\varphi\|_{\dot{H}^1(\Sigma_\tau)} \right)$$

Combining (33) and (34), we end up finding a constant  $C > 0$  such that

$$\sup_{\Sigma_\tau} |\varphi(p)| \leq C_7 \left( \|\varphi\|_{\dot{H}^1(\Sigma_0)} + \|n\varphi\|_{L^2(\Sigma_0)} \right) \leq C.$$

The last inequality holds since, similar to the analysis for hypersurfaces separate from the event horizon, both of these norms are bounded by a constant depending only on the initial data set.  $\square$

## 5. Conclusion

As mentioned in Section 1, there is a natural generalization of Theorem 1.1. This was the theorem of focus in Dafermos and Rodnianski's original paper [4] where they formalized the local redshift effect. Their paper proves the following theorem.

**THEOREM 5.1** (Dafermos-Rodnianski, 2008). *Let  $\phi$  be a sufficiently regular solution to the wave equation*

$$\square_g \phi = 0$$

*on the (maximally extended) Schwarzschild spacetime  $(\mathcal{M}, g)$ , decaying suitably at spatial infinity on an arbitrary complete asymptotically flat Cauchy surface  $\Sigma$ . Fix hindered and advanced Eddington-Finkelstein coordinates  $u$  and  $v$  on one of the exterior regions. For any achronal hypersurface  $\mathcal{S}$  in the closure of this region, let  $F(\mathcal{S})$  denote the flux of energy through  $\mathcal{S}$ , where energy is here measured with respect to the timelike Killing vector field. Let  $v_+ = \max\{v, 1\}$ ,*

$u_+ = \max\{u, 1\}$ , and  $v_+(\mathcal{S}) = \max\{\inf_{\mathcal{S}} v, 1\}$ ,  $u_+(\mathcal{S}) = \max\{\inf_{\mathcal{S}} u, 1\}$ . We have

$$F(\mathcal{S}) \leq C \left( (v_+(\mathcal{S}))^{-2} + (u_+(\mathcal{S}))^{-2} \right).$$

(We also allow  $\mathcal{S}$  to be a subset of null infinity, interpreted in the obvious limiting sense.) In addition, we have the pointwise decay rates

$$\begin{aligned} |\phi| &\leq C(v_+)^{-1} && \text{in } \overline{J^-(\mathcal{I}^+) \cap J^+(\mathcal{I}^-)} \\ |r\phi| &\leq C_{\hat{R}}(1 + |u|)^{-1/2} && \text{in } \{r \leq \hat{R} < 2M\} \cap J^+(\Sigma). \end{aligned}$$

Although this result is still cited in present literature, 2008 was a long time ago, and there have been many advances to our understanding of stability in different black hole spacetimes. With respect to Schwarzschild spacetimes, linear stability was proven in Dafermos, Holzegel, and Rodnianski in 2016 (see [?]). This was an interim result between our Theorem 1.1 and the general nonlinear case. The general nonlinear case was solved in 2021 by Dafermos, Holzegel, Rodnianski, and Taylor when they proved the nonlinear asymptotic stability of solutions to the Einstein vacuum equation in the exterior region of the black hole (see [?]). They showed that generic vacuum initial data (meaning no assumed symmetry) sufficiently close to Schwarzschild data evolve to a vacuum spacetime that has the following properties:

- (1) It possesses a complete future null infinity  $\mathcal{I}^+$ . The past of the future null infinity  $J^-(\mathcal{I}^+)$  is bounded by a regular future complete event horizon  $\mathcal{H}^+$ .
- (2) It remains close to Schwarzschild in the exterior region to the black hole.
- (3) It asymptotically approaches a member of the Schwarzschild family as (an appropriate notion of) time goes to infinity.

The general nonlinear stability of Kerr spacetimes remains conjectured. However, Giorgi, Klainerman, and Szeftel proved nonlinear stability for small angular momentum in 2022 (see [?] and [?]). These proofs relied on the construction of the *general covariant modulation* (GCM) procedure which constructs a dynamical center of mass frame of the final state. An outline of this procedure can be found in Section 3 of [?]. Most recently, Fang, Giorgi, and Wan have developed a modification and extension of the GCM procedure in [?]. This extension is applicable to solutions of a broad class of matter models, such as the Kerr-Newman spacetimes which describe the geometry of rotating and electrically charged black holes. Current research is focused on proving the nonlinear stability of these spacetimes.

**CONJECTURE 5.2.** (*Nonlinear Stability of Kerr-Newman Spacetimes*) *Electrovacuum initial data sufficiently close to Kerr-Newman initial data have a maximal development with a complete future null infinity and a domain of outer communication which globally approaches a nearby Kerr-Newman solution.*

## References

- [1] ARETAKIS, S. Lecture notes on general relativity, 2013.
- [2] CHRISTODOULOU, D. *The Action Principle and Partial Differential Equations.* (AM-146). Princeton University Press, 2000.
- [3] DAFERMOS, M. The geometry and analysis of black hole spacetimes in general relativity, 2018.
- [4] DAFERMOS, M., AND RODNIANSKI, I. The red-shift effect and radiation decay on black hole spacetimes, 2005.
- [5] DAFERMOS, M., AND RODNIANSKI, I. Lectures on black holes and linear waves, 2008.
- [6] EVANS, L. *Partial Differential Equations.* Graduate studies in mathematics. American Mathematical Society, 2010.
- [7] GALDI, G. P. *An Introduction to the Mathematical Theory of the Navier-Stokes Equations, Second Edition.* Springer, 2011.
- [8] GIORGI, E. The linear stability of reissner–nordström spacetime: The full subextremal range  $|q| < m$ . *Communications in Mathematical Physics* 380, 3 (Nov. 2020), 1313–1360.
- [9] HOLZEGEL, G. Lecture notes vienna, 2014.
- [10] KAY, B. S., AND WALD, R. M. Linear stability of schwarzschild under perturbations which are non-vanishing on the bifurcation 2-sphere. *Classical and Quantum Gravity* 4, 4 (jul 1987), 893.
- [11] KRUSKAL, M. D. Maximal extension of schwarzschild metric. *Physical Review* (1960).
- [12] OPPENHEIMER, J. R., AND SNYDER, H. On continued gravitational contraction. *Phys. Rev.* 56 (Sep 1939), 455–459.

KATHERINE MEKECHUK

Department of Mathematics, Columbia University

*E-mail:* [columbiajournalofundergradmath@gmail.com](mailto:columbiajournalofundergradmath@gmail.com)

<https://journals.library.columbia.edu/index.php/cjum>

# On Günther’s Perturbation Results For Nash’s Isometric Embedding Theorems

By SHIV YAJNIK

## Contents

1. Introduction	1
2. Riemannian Geometry and Isometric Embeddings	3
3. Regularity Issues, Perturbations, and Hölder Norms	11
4. Günther’s Solution	21
5. Consequences and Other Problems	40
6. Reverberations	44
References	50

## 1. Introduction

Very often, the Riemannian manifolds that we are interested in are explicitly defined as submanifolds of Euclidean space. For example, the  $n$ -sphere  $\mathbb{S}^n$  can be written as the real solution space of the equation  $\sum_{k=1}^n x_k^2 = 1$ . This equation expresses  $\mathbb{S}^n$  as a submanifold of  $\mathbb{R}^{n+1}$ , and the round metric on this sphere is induced by the Euclidean metric. So, we can easily imagine, for example, what the geodesics may be on these types of submanifolds: They are the curves that are closest to being straight lines in the Euclidean sense. In the case of the sphere, those are the great circles. Similarly, we can also picture what angles might look like without too much difficulty.

But manifolds are not a-priori defined to have a structure whose coordinates are directly inherited from Euclidean space; they are merely required to be *locally Euclidean*—meaning, to consist of Euclidean “patches” that are sewn together. The version of the sphere that we are most familiar with—the one

---

Received by the editors May 2025.

© 2026 Yajnik, Shiv. This is an open access article distributed under the terms of the Creative Commons BY-NC-ND 4.0 license.

with the round metric—is one that can be *isometrically embedded* into three-dimensional Euclidean space. Informally, isometric embeddings are a way to identify a manifold as a subset of Euclidean space in such a way that preserves distances, angles, and curvature intrinsic to the manifold itself. More formally, isometric embeddings preserve the Riemannian metric on the sphere.

We consider two important questions:

- Can any manifold with any metric be isometrically embedded into a Euclidean space, where the embedding is of some desired differentiability class?
- How high is the dimension of the ambient Euclidean space required to be?

This is a very nontrivial problem: For example, there are many metrics on surfaces and higher dimensional spheres which do not admit codimension 1 isometric embeddings, either at all or in the desired differentiability class, as we will see in Sections 2.2 and 3.1.

John Nash proved that any compact Riemannian manifold of dimension  $n$  can be smoothly isometrically embedded into  $\frac{n(3n+11)}{2}$ -dimensional Euclidean space, and that any Riemannian manifold (not necessarily compact) can be smoothly isometrically embedded into  $\frac{3}{2}n^3 + 7n^2 + \frac{11}{2}n$ -dimensional Euclidean space. Günther, whose methods we will explore in this exposition, improved this bound to

$$\max\left(\frac{n(n+5)}{2}, \frac{n(n+3)}{2} + 5\right)$$

for any Riemannian manifold; this remains the best result for the general case as of 2026.

An isometric embedding is some differentiable map from a manifold  $M$  to Euclidean space  $\mathbb{R}^N$ . As we will see, finding an isometric embedding requires analysis of a system of nonlinear PDE's of first order, which appears in the following form:

$$g_{ij} = \sum_{r=1}^N \partial_i u_r \partial_j u_r \quad 1 \leq i, j \leq n$$

$g_{ij}$  are known, and  $u_1, \dots, u_N$  are unknown.

In Section 2, we provide a quick introduction to some fundamental notions in Riemannian geometry, including isometric embeddings, and some motivating examples. In Section 3, we build our understanding of the isometric embedding problem as one about PDE's and discuss associated regularity issues, including a certain loss of differentiability which makes this problem especially

difficult. The meat of this exposition is contained in Section 4, where we will study the regularity of solutions to the above system of PDE's—in particular, we will examine Matthias Günther's solution to the loss of differentiability problem, which elegantly demonstrates that there exist  $C^s$  solutions  $u$ , given that  $g_{ij}$  is  $C^s$ -differentiable for  $1 \leq s \leq \infty$ . Then, I include a summary of how this perturbation result is applied to prove Günther's version of the isometric embedding theorem.<sup>1</sup> We will conclude with a summary of consequences of the Nash embedding problem, results of isometric embeddings in more specific cases, and a short introduction to global embedding problems in other areas of differential geometry in Sections 5 and 6.

*Prerequisites.* Sections 2, 3, and 4 assume linear algebra, some real analysis (up to the Arzela–Ascoli theorem) and basic knowledge of manifolds, for example Chapters 1-3, Chapter 5, and Appendices A and B of [57]. Sections 5 and 6 may require some more familiarity with geometry to fully grasp, but I have included citations to literature that more thoroughly discuss many of the areas mentioned.

*Acknowledgments.* I would like to thank Professor Mu-Tao Wang of Columbia University's Math Department for his endless support and guidance in overseeing my senior thesis during my very difficult final year of college. I also thank Professor Wang for introducing me to this fascinating subject. I am very grateful to have had the opportunity to do a project with him. I appreciate the Columbia Undergraduate Journal of Mathematics for their support and the referees for reviewing my thesis. As a result of their suggestions, this exposition is now a significantly expanded version of my senior thesis, which I wrote in 2025. This version, from 2026, serves as a better introduction to the topics surrounding the isometric embedding problem and includes more details about how Günther applied his perturbation result to prove his version of John Nash's isometric embedding theorem.

## 2. Riemannian Geometry and Isometric Embeddings

2.1. *An Overview of Fundamentals from Riemannian Geometry.* As in [57], I assume that our manifolds are Hausdorff, second countable, and smooth.

We introduce the following notation:

- The space of smooth functions is denoted as  $C^\infty(M)$ .

---

<sup>1</sup>Two other useful sources on Nash's and Günther's work on the isometric embedding problem are [2] by Ben Andrews and [38] by Siyuan Lu.

- The tangent space at a point  $p \in M$  is denoted  $T_pM$ , and the tangent bundle is denoted  $TM$ .
- The cotangent space at a point  $p \in M$  is denoted  $T_p^*M$ , and the cotangent bundle is denoted  $T^*M$ .
- For a vector bundle  $E$  and for  $\lambda > 0$ ,  $C^{k,\lambda}(M, E)$ , denotes the space of  $C^{k,\lambda}$  sections of  $E$ , and  $C^\infty(M, E)$  denotes the smooth sections of  $E$ .<sup>2</sup>

In this section, we provide a brief overview of relevant concepts from Riemannian geometry.

First, we review tensor products:

DEFINITION 2.1. *For any commutative ring  $R$ , the tensor product  $A \otimes B$  of (left)  $R$ -modules  $A$  and  $B$  contains elements  $a \otimes b$  with  $a \in A$  and  $b \in B$  that obey the following relations<sup>3</sup>:*

$$\begin{aligned} ra \otimes b &= r(a \otimes b) = (a \otimes rb) \\ (a_1 + a_2) \otimes b &= a_1 \otimes b + a_2 \otimes b \\ a \otimes (b_1 + b_2) &= a \otimes b_1 + a \otimes b_2 \end{aligned}$$

The spaces  $C^\infty(M, TM)$  and  $C^\infty(M, T^*M)$  of vector fields and differential 1-forms, respectively, are both  $C^\infty(M)$ -modules, so we can take tensor products of them. Conveniently, elements of  $C^\infty(M, T^*M) \otimes C^\infty(M, T^*M)$ , of the form  $\omega \otimes \tau$  with  $\omega, \tau \in C^\infty(M, T^*M)$ , can be expressed at every point  $p \in M$  by the  $\mathbb{R}$ -bilinear map

$$\begin{aligned} (\omega \otimes \tau)_p &: T_pM \otimes T_pM \rightarrow \mathbb{R} \\ (\omega \otimes \tau)_p(\mathbf{u} \otimes \mathbf{v}) &= \omega(\mathbf{u})\tau(\mathbf{v}) \end{aligned}$$

The above corresponds to the definition of 2-tensors given in [57]. Letting  $p$  vary across  $M$ , we observe  $(\omega \otimes \tau)(U \otimes V) = \omega(U)\tau(V) \in C^\infty(M)$  (with  $U, V \in C^\infty(M, TM)$ ). We can repeatedly take tensor products as above to get  $k$ -tensors as described in [57].

DEFINITION 2.2. *Define the vector bundle*

$$T_s^r M := TM^{\otimes r} \otimes T^*M^{\otimes s}$$

---

<sup>2</sup>For further discussion on the meaning of  $C^{k,\lambda}$  (Hölder differentiability), see the beginning of Section 3.3.

<sup>3</sup>For more details, in [3], see the beginning of Chapter 1 for definition of commutative rings and Chapter 2 for the definition of modules and tensor products of modules. For the sake of intuition, one can think of a module as being similar to a vector space, but with “scalar multiplication” by elements of a commutative ring such as  $C^\infty(M)$  instead of elements of a field such as  $\mathbb{R}$ .

whose space of sections is<sup>4</sup>

$$C^\infty(M, TM)^{\otimes r} \otimes C^\infty(M, T^*M)^{\otimes s}$$

consisting of  $(r, s)$ -tensors.

DEFINITION 2.3. A Riemannian metric  $g : M \rightarrow T^*M^{\otimes 2}$  is a symmetric  $(0, 2)$ -tensor such that  $g(p)$  defines an inner product on the tangent space  $T_pM$ . A Riemannian manifold  $(M, g)$  is a manifold  $M$  equipped with a Riemannian metric  $g$ .

The bundle of symmetric  $(0, 2)$ -tensors is denoted  $\text{Sym}^2(T^*X)$ . The associated space of tensors is  $C^\infty(M)$ -spanned by elements of the form

$$dx_i dx_j = \frac{1}{2}(dx_i \otimes dx_j + dx_j \otimes dx_i)$$

We will write  $dx^2$  as shorthand for  $dx dx$ .

Any smooth manifold admits a Riemannian metric.  $g_{ij} = g(\partial_i, \partial_j)$ <sup>5</sup> are smooth functions such that at every point, the matrix with entries  $g_{ij}(p)$  is a positive-definite symmetric matrix. We write  $g^{ij}$  to denote the entries of the inverse matrix.

We measure rate of change in vector fields with the following device, defined by the Riemannian metric.

DEFINITION 2.4. Given a Riemannian manifold  $(M, g)$ , its Levi-Civita connection is the unique map  $\nabla : C^\infty(M, TM) \times C^\infty(M, TM) \rightarrow C^\infty(M, TM)$  which satisfies:

- (i)  $C^\infty(M)$ -linearity in the first argument:  $\nabla_{fX+gY}Z = f\nabla_XZ + g\nabla_YZ$  for  $X, Y, Z \in C^\infty(M, TM)$  and  $f, g \in C^\infty(M, TM)$
- (ii)  $\mathbb{R}$ -linearity in the second argument:  $\nabla_Z(\alpha X + \beta Y) = \alpha\nabla_ZX + \beta\nabla_ZY$  for  $X, Y, Z \in C^\infty(M, TM)$  and  $\alpha, \beta \in \mathbb{R}$
- (iii) Leibniz rule:  $\nabla_X fY = X(f)Y + f\nabla_XY$
- (iv) Torsion freeness:  $\nabla_XY - \nabla_YX = [X, Y]$
- (v) Metric compatibility:  $Z(g(X, Y)) = g(\nabla_ZX, Y) + g(X, \nabla_ZY)$

The Levi-Civita connection is uniquely defined by the Koszul formula:

$$(1) \quad 2g(\nabla_XY, Z) = X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y)) - g([X, Z], Y) - g([Y, Z], X) + g([X, Y], Z)$$

---

<sup>4</sup>  $A^{\otimes r}$  is shorthand for  $r$  tensor products of  $A$ .

<sup>5</sup>  $\partial_i = \frac{\partial}{\partial x_i}$

We write  $\nabla_i$  to represent  $\nabla_{\partial_i}$ . The covariant derivatives  $\nabla_i$  do not commute in general; in fact, the commutator of two covariant derivatives is a curvature term.

We have that<sup>6</sup>  $\nabla_i \partial_j = \Gamma_{ij}^l \partial_l$  for some  $\Gamma_{ij}^l \in C^\infty(M)$ .  $\Gamma_{ij}^l$  are called the Christoffel symbols, and they can be computed from the metric by the following formula:

$$(2) \quad \Gamma_{ij}^l = \frac{1}{2} \sum_{k=1}^n g^{lk} (\partial_j g_{ik} + \partial_i g_{kj} - \partial_k g_{ij})$$

Below are definitions for Riemannian and Ricci curvature, respectively, which will appear in the paper:

DEFINITION 2.5. *Riemannian curvature  $R$  is defined by*

$$R(X, Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X, Y]}Z$$

The above definition expresses Riemannian curvature as a  $(1, 3)$  tensor. We can also define Riemannian curvature as a  $(0, 4)$  tensor given by

$$R(W, X, Y, Z) = g(R(W, X)Y, Z)$$

DEFINITION 2.6. *Ricci curvature is the bilinear form  $Q$  given by<sup>7</sup>*

$$\text{Ric}(p)(\mathbf{x}, \mathbf{y}) = \text{tr}(\mathbf{v} \mapsto R(\mathbf{x}, \mathbf{v})\mathbf{y}) = \sum_{k=1}^n R(p)(\mathbf{e}_k, \mathbf{x}, \mathbf{y}, \mathbf{e}_k)$$

where  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  forms an orthonormal basis for  $T_p M$ .

In terms of local coordinates, we have

$$\begin{aligned} R(\partial_i, \partial_j)\partial_k &= R_{ijk}{}^m \partial_m = [\nabla_i, \nabla_j]\partial_k & R(\partial_i, \partial_j, \partial_k, \partial_l) &= R_{ijkl} = R_{ijk}{}^m g_{ml} \\ \text{Ric}(\partial_i, \partial_j) &= R_{ij} = R_{kij}{}^k = R_{kijl} g^{kl} \end{aligned}$$

Notice how Riemannian curvature has four indices, and Ricci curvature comes from multiplying by metric tensor elements and summing over pairs of indices so that Ricci curvature has only two indices. This is an example of *contraction*—that is, the Ricci curvature tensor is a contraction of the Riemannian curvature tensor.

For  $X \in C^\infty(M, TM)$ , let the  $\mathbb{R}$ -linear map  $\nabla_X : C^\infty(M, TM) \rightarrow C^\infty(M, TM)$  be given by the Levi-Civita connection.  $\nabla_X$  extends to a unique  $\mathbb{R}$ -linear map  $C^\infty(M, T_s^r M) \rightarrow C^\infty(M, T_s^r M)$  on  $(r, s)$ -tensors satisfying

<sup>6</sup>We invoke the Einstein summation convention here.

<sup>7</sup>We use the Einstein summation convention here.

- $\nabla_X(c(S)) = c(\nabla_X(S))$  for any tensor  $S$  and any contraction  $c$ .
- Leibniz rule:  $\nabla_X(S \otimes T) = (\nabla_X S) \otimes T + S \otimes (\nabla_X T)$  for any tensors  $S, T$ .

For  $(0, s)$  tensors, letting  $Y_1, \dots, Y_s \in C^\infty(M, TM)$ , we have

$$\nabla_X T(Y_1, \dots, Y_s) = X(T(Y_1, \dots, Y_s)) - \sum_{i=1}^s T(Y_1, \dots, Y_{i-1}, \nabla_X Y_i, Y_{i+1}, \dots, Y_s)$$

Finally, we define isometric embeddings, which are the subject of this paper:

**DEFINITION 2.7.** *Let  $(M_1, g_1)$  and  $(M_2, g_2)$  be two manifolds with  $\dim M_2 \geq \dim M_1$ . A smooth function  $u : M_1 \rightarrow M_2$  is an embedding if the Jacobian of  $u$  is of full rank (making  $u$  an immersion) and if  $u$  is homeomorphic onto its image. This embedding is isometric if  $u^* g_2 = g_1$ .*

We are interested in finding isometric embeddings of Riemannian manifolds  $(M, g)$  into  $(\mathbb{R}^N, g_E)$  where  $g_E$  is the standard Euclidean metric

$$g_E = \sum_{k=1}^N dx_k^2$$

If  $u$  is an isometric embedding of  $M$  into  $\mathbb{R}^N$ , then  $u(M)$  is a submanifold of  $\mathbb{R}^N$  with the same geodesics and curvature as that of  $M$ , so in essence, one is placing  $M$  inside of Euclidean space without any deformation or damage to the original Riemannian structure of  $M$ .

**2.2. Examples I: Metrics on Spheres.** To get a feel for what manifolds with a given Riemannian structure could look like, we take a short detour to list a few examples of Riemannian metrics and the resulting curvature. These examples are of different metrics on the same types of manifolds: spheres. We illustrate a few of the ways that differences between these structures manifest; in particular, we will see that while one metric may isometrically embed into an Euclidean space of a given dimension  $N$ , there may be plenty of other metrics that do not.

Below, we define two more notions of curvature that may be more readily visualizable than the original Riemannian curvature tensor.

**DEFINITION 2.8.** *Let  $(M, g)$  be a Riemannian manifold,  $R$  the curvature tensor from the Levi-Civita connection, and  $\sigma$  a 2 dimensional subspace of*

$T_p M$ . We define the sectional curvature of  $\sigma$  to be

$$K(\sigma, p) := \frac{R(p)(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1)}{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle \langle \mathbf{e}_2, \mathbf{e}_2 \rangle - \langle \mathbf{e}_1, \mathbf{e}_2 \rangle^2}$$

Where  $\{\mathbf{e}_1, \mathbf{e}_2\}$  is a basis of  $\sigma$  and  $\langle \cdot, \cdot \rangle = g(p)(\cdot, \cdot)$ .

One can show that the sectional curvature determines the Riemannian curvature tensor.

DEFINITION 2.9. Where  $\text{Ric}$  is the Ricci curvature tensor, we define scalar curvature to be the smooth function given by

$$S(p) := \frac{1}{n(n-1)} \sum_{l=1}^n \text{Ric}(p)(\mathbf{e}_l, \mathbf{e}_l)$$

such that  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  is an orthonormal basis.<sup>8</sup>

Scalar curvature contains less information but is often simpler to compute and provides perhaps the most friendly measure for how bent a space is around a point.

*Example 2.10 (The Round Metric).* The standard round metric  $g_{\text{can}}$  on the sphere  $\mathbb{S}^n$  can be expressed as the pullback of the metric  $g_E$  on  $\mathbb{R}^{n+1}$  under the inclusion map  $\iota : \mathbb{S}^n \hookrightarrow \mathbb{R}^{n+1}$  when  $\mathbb{S}^n$  is written as the level set  $\sum_{k=1}^n x_k^2 = 1$ . For example, on  $\mathbb{S}^2$ , in local coordinates on the northern hemisphere (i.e.  $z > 0$ ), the metric can be written as

$$g_{\text{can}} = \left(1 + \frac{x^2}{1 - x^2 - y^2}\right) dx^2 + \left(1 + \frac{y^2}{1 - x^2 - y^2}\right) dy^2 - \left(\frac{2xy}{1 - x^2 - y^2}\right) dx dy$$

There is a much nicer expression for  $g_{\text{can}}$  in spherical coordinates  $(r, \theta_1, \theta_2)$ . If we choose the neighborhood of the sphere where  $\theta_1 \in (0, \pi)$  and  $\theta_2 \in (0, 2\pi)$ , we can express the metric on the corresponding ambient neighborhood  $\mathbb{R} \times (0, \pi) \times (0, 2\pi) \subseteq \mathbb{R}^3$  in spherical coordinates by taking the pullback of  $dx^2 + dy^2 + dz^2$  under the coordinate transformation:

$$(r, \theta_1, \theta_2) \mapsto (r \cos \theta_1, r \sin \theta_1 \cos \theta_2, r \sin \theta_1 \sin \theta_2)$$

We get

$$g_E = dr^2 + r^2 d\theta_1^2 + r^2 \sin^2 \theta_1 d\theta_2^2$$

---

<sup>8</sup>In some other conventions, the definition of scalar curvature leaves out the factor of  $\frac{1}{n(n-1)}$ .

Since the sphere is given by just  $r = 1$ , pulling back  $g_E$  via the inclusion map  $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3 \setminus \{0\}$  reveals that the metric on the sphere can be written as

$$g_{\text{can}} = d\theta_1^2 + \sin^2 \theta_1 d\theta_2^2$$

Using (2) again, we can compute the Christoffel symbols:  $\Gamma_{12}^2 = \Gamma_{21}^2 = \cot \theta_1$ ,  $\Gamma_{22}^1 = -\sin \theta_1 \cos \theta_1$ , and the rest are 0. So we have

$$\begin{aligned}\nabla_1 \partial_1 &= 0 \\ \nabla_2 \partial_1 &= \nabla_1 \partial_2 = \cot \theta_1 \partial_2 \\ \nabla_2 \partial_2 &= -\sin \theta_1 \cos \theta_1 \partial_1\end{aligned}$$

With the orthonormal basis  $\mathbf{e}_1 = \partial_1$  and  $\mathbf{e}_2 = \frac{\partial_2}{\sin \theta_1}$ , we compute

$$R(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1) = \frac{g_{\text{can}}(\nabla_1 \nabla_2 \partial_2 - \nabla_2 \nabla_1 \partial_2, \partial_1)}{\sin^2 \theta_1} = -\frac{g_{\text{can}}(\nabla_1(\sin \theta_1 \cos \theta_1 \partial_1) - \nabla_2(\cot \theta_1 \partial_2), \partial_1)}{\sin^2 \theta_1} = 1$$

We get that the sectional curvature  $K$  and the scalar curvature  $S$  are:

$$\begin{aligned}K &= \frac{R(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1)}{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle \langle \mathbf{e}_2, \mathbf{e}_2 \rangle - \langle \mathbf{e}_1, \mathbf{e}_2 \rangle^2} = 1 \\ S &= \frac{1}{2} (R(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1) + R(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1)) = 1\end{aligned}$$

Hence, the 2-sphere with round metric has constant sectional curvature 1, as does every  $n$ -sphere with the round metric. Round spheres also have scalar curvature 1, which is to be expected: scalar curvature here will agree with sectional curvature if the sectional curvature is constant.

Since the round metric on  $\mathbb{S}^n$  is the pullback of the standard metric on  $\mathbb{R}^{n+1}$  along inclusion, we have automatically that  $(\mathbb{S}^n, g_{\text{can}})$  smoothly isometrically embeds into  $(\mathbb{R}^{n+1}, g_E)$ .

*Example 2.11* (Berger Spheres). The *quaternions* are a unital associative  $\mathbb{R}$ -algebra of elements of the form  $z = v + w\mathbf{i} + x\mathbf{j} + y\mathbf{k}$  with  $w, x, y, z \in \mathbb{R}$  and relations  $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$ . These can be thought of as an extension of the complex numbers. We define  $\bar{z} = v - w\mathbf{i} - x\mathbf{j} - y\mathbf{k}$ .

$\mathbb{S}^3$  can be expressed as the set of unit quaternions, given by  $|z| = z\bar{z} = 1$ . The unit quaternions form a noncommutative *Lie group*. In this case, this means that the unit quaternions have a manifold structure combined with an associative arithmetic operation—quaternion multiplication—with identity element 1 and a multiplicative inverse for every element. On  $\mathbb{S}^3$ , multiplication  $\cdot : \mathbb{S}^3 \times \mathbb{S}^3 \rightarrow \mathbb{S}^3$  is smooth, and so is the associated inverse map  $()^{-1} : \mathbb{S}^3 \rightarrow \mathbb{S}^3$ . Quaternion multiplication, unlike multiplication of complex numbers, is *non-commutative*.

On a Lie group, every tangent space is canonically isomorphic to the tangent space at the identity by pushforward along left multiplication by group elements; the tangent space at the identity is referred to as the *Lie algebra*.

The Lie algebra  $T_1\mathbb{S}^3$  of  $\mathbb{S}^3$  consists of imaginary quaternions, of the form  $w\mathbf{i} + x\mathbf{j} + y\mathbf{k}$ . This Lie algebra is generated by the basis  $\mathbf{i}$ ,  $\mathbf{j}$ , and  $\mathbf{k}$  satisfying

$$(3) \quad [\mathbf{i}, \mathbf{j}] = 2\mathbf{k}, \quad [\mathbf{j}, \mathbf{k}] = 2\mathbf{i}, \quad [\mathbf{k}, \mathbf{i}] = 2\mathbf{j}$$

Because of the canonical isomorphism of tangent spaces at every point, we can define a metric on  $\mathbb{S}^3$  by specifying an inner product on the Lie algebra and defining the inner product at other points simply to be the pushforward of the inner product along left multiplication of group elements.

We define a metric in this way by declaring an orthonormal basis to be  $\varepsilon\mathbf{i}$ ,  $\mathbf{j}$ , and  $\mathbf{k}$  for a constant  $\varepsilon \in \mathbb{R}^+$ . This defines a family of metrics  $g^\varepsilon$  on  $\mathbb{S}^3$ ;  $(\mathbb{S}^3, g^\varepsilon)$  are known as the Berger Spheres, after Marcel Berger.<sup>9</sup>  $g^1$  is the round metric. Using the Koszul formula (1), we now compute the covariant derivatives and the curvature tensor. Let  $\nabla^\varepsilon$  denote the Levi-Civita connection of  $g^\varepsilon$ .

$$\begin{aligned} g^\varepsilon(\nabla_i^\varepsilon \mathbf{i}, \mathbf{i}) &= g^\varepsilon(\nabla_i^\varepsilon \mathbf{i}, \mathbf{j}) = g^\varepsilon(\nabla_i^\varepsilon \mathbf{i}, \mathbf{k}) = 0 \implies \nabla_i^\varepsilon \mathbf{i} = 0 \\ g^\varepsilon(\nabla_j^\varepsilon \mathbf{j}, \mathbf{i}) &= g^\varepsilon(\nabla_j^\varepsilon \mathbf{j}, \mathbf{j}) = 0, \quad g^\varepsilon(\nabla_j^\varepsilon \mathbf{j}, \mathbf{k}) = 2 - \frac{1}{\varepsilon^2} \implies \nabla_j^\varepsilon \mathbf{j} = \left(2 - \frac{1}{\varepsilon^2}\right) \mathbf{k} \implies \nabla_j^\varepsilon \mathbf{i} = -\frac{1}{\varepsilon^2} \mathbf{k} \\ g^\varepsilon(\nabla_i^\varepsilon \mathbf{k}, \mathbf{i}) &= 0, \quad g^\varepsilon(\nabla_i^\varepsilon \mathbf{k}, \mathbf{j}) = -2 + \frac{1}{\varepsilon^2}, \quad g^\varepsilon(\nabla_i^\varepsilon \mathbf{k}, \mathbf{k}) = 0 \implies \nabla_i^\varepsilon \mathbf{k} = \left(-2 + \frac{1}{\varepsilon^2}\right) \mathbf{j} \implies \nabla_k^\varepsilon \mathbf{i} = \frac{1}{\varepsilon^2} \mathbf{j} \\ g^\varepsilon(\nabla_j^\varepsilon \mathbf{j}, \mathbf{i}) &= g^\varepsilon(\nabla_j^\varepsilon \mathbf{j}, \mathbf{j}) = g^\varepsilon(\nabla_j^\varepsilon \mathbf{j}, \mathbf{k}) = 0 \implies \nabla_j^\varepsilon \mathbf{j} = 0 \\ g^\varepsilon(\nabla_j^\varepsilon \mathbf{k}, \mathbf{i}) &= \frac{1}{\varepsilon^2}, \quad g^\varepsilon(\nabla_j^\varepsilon \mathbf{k}, \mathbf{j}) = 0, \quad g^\varepsilon(\nabla_j^\varepsilon \mathbf{k}, \mathbf{k}) = 0 \implies \nabla_j^\varepsilon \mathbf{k} = \mathbf{i} \implies \nabla_k^\varepsilon \mathbf{j} = -\mathbf{i} \\ g^\varepsilon(\nabla_k^\varepsilon \mathbf{k}, \mathbf{i}) &= g^\varepsilon(\nabla_k^\varepsilon \mathbf{k}, \mathbf{j}) = g^\varepsilon(\nabla_k^\varepsilon \mathbf{k}, \mathbf{k}) = 0 \implies \nabla_k^\varepsilon \mathbf{k} = 0 \end{aligned}$$

With the orthonormal basis  $\mathbf{e}_1 = \varepsilon\mathbf{i}$ ,  $\mathbf{e}_2 = \mathbf{j}$ , and  $\mathbf{e}_3 = \mathbf{k}$ , we find

$$\begin{aligned} R(\varepsilon\mathbf{i}, \mathbf{j}, \mathbf{j}, \varepsilon\mathbf{i}) &= \varepsilon^2 g^\varepsilon(\nabla_i^\varepsilon \nabla_j^\varepsilon \mathbf{j} - \nabla_j^\varepsilon \nabla_i^\varepsilon \mathbf{j} - \nabla_{[\mathbf{i}, \mathbf{j}]}^\varepsilon \mathbf{j}, \mathbf{i}) = \frac{1}{\varepsilon^2} \\ R(\varepsilon\mathbf{i}, \mathbf{k}, \mathbf{k}, \varepsilon\mathbf{i}) &= \varepsilon^2 g^\varepsilon(\nabla_i^\varepsilon \nabla_k^\varepsilon \mathbf{k} - \nabla_k^\varepsilon \nabla_i^\varepsilon \mathbf{k} - \nabla_{[\mathbf{i}, \mathbf{k}]}^\varepsilon \mathbf{k}, \mathbf{i}) = \frac{1}{\varepsilon^2} \\ R(\mathbf{j}, \mathbf{k}, \mathbf{k}, \mathbf{j}) &= g^\varepsilon(\nabla_j^\varepsilon \nabla_k^\varepsilon \mathbf{k} - \nabla_k^\varepsilon \nabla_j^\varepsilon \mathbf{k} - \nabla_{[\mathbf{j}, \mathbf{k}]}^\varepsilon \mathbf{k}, \mathbf{j}) = 4 - \frac{3}{\varepsilon^2} \end{aligned}$$

---

<sup>9</sup>Berger spheres are included as Example 1.3.5. and discussed in Section 4.4.3 of [47]. Some of the calculations here can be found there as well.

We find that the sectional curvatures  $K(\sigma)$  at  $\sigma = \text{Span}_{\mathbb{R}}(\mathbf{i}, \mathbf{j})$ ,  $\text{Span}_{\mathbb{R}}(\mathbf{i}, \mathbf{k})$ ,  $\text{Span}_{\mathbb{R}}(\mathbf{j}, \mathbf{k})$  and the scalar curvature  $S$  is given by:

$$\begin{aligned} K(\text{Span}_{\mathbb{R}}(\mathbf{i}, \mathbf{j})) &= \frac{R(\varepsilon\mathbf{i}, \mathbf{j}, \mathbf{j}, \varepsilon\mathbf{i})}{\langle \varepsilon\mathbf{i}, \varepsilon\mathbf{i} \rangle \langle \mathbf{j}, \mathbf{j} \rangle - \langle \varepsilon\mathbf{i}, \mathbf{j} \rangle^2} = \frac{1}{\varepsilon^2} \\ K(\text{Span}_{\mathbb{R}}(\mathbf{i}, \mathbf{k})) &= \frac{R(\varepsilon\mathbf{i}, \mathbf{k}, \mathbf{k}, \varepsilon\mathbf{i})}{\langle \varepsilon\mathbf{i}, \varepsilon\mathbf{i} \rangle \langle \mathbf{k}, \mathbf{k} \rangle - \langle \varepsilon\mathbf{i}, \mathbf{k} \rangle^2} = \frac{1}{\varepsilon^2} \\ K(\text{Span}_{\mathbb{R}}(\mathbf{j}, \mathbf{k})) &= \frac{R(\mathbf{j}, \mathbf{k}, \mathbf{k}, \mathbf{j})}{\langle \mathbf{j}, \mathbf{j} \rangle \langle \mathbf{k}, \mathbf{k} \rangle - \langle \mathbf{j}, \mathbf{k} \rangle^2} = 4 - \frac{3}{\varepsilon^2} \end{aligned}$$

$$S = \frac{1}{6}(2R(\varepsilon\mathbf{i}, \mathbf{j}, \mathbf{j}, \varepsilon\mathbf{i}) + 2R(\varepsilon\mathbf{i}, \mathbf{k}, \mathbf{k}, \varepsilon\mathbf{i}) + 2R(\mathbf{j}, \mathbf{k}, \mathbf{k}, \mathbf{j})) = \frac{1}{6} \left( 8 - \frac{2}{\varepsilon^2} \right)$$

The above curvature data shows that the metrics  $g^\varepsilon$  are very different from the round metric  $g^1$ :

- The scalar curvature can be zero or negative depending on the value of  $\varepsilon$ .
- The sectional curvature  $K(p, \sigma)$  is non-constant with respect to  $\sigma$  if  $\varepsilon \neq 1$ .

Also unlike the round metric, infinitely many of the metrics  $g^\varepsilon$  for  $\varepsilon > 0$  do not admit *conformal* embeddings,<sup>10</sup> let alone isometric embeddings, by Corollary 2.2 of Chiakuei Peng and Zizhou Tang in [46]; they prove this by an argument involving the Chern–Simons 3-form.

### 3. Regularity Issues, Perturbations, and Hölder Norms

Recall that given a metric  $g$  on  $M$ , we are to find an embedding  $u : M \rightarrow \mathbb{R}^N$  such that  $u^*g_E = g$ . Equivalently, we may look at this problem in the following way: Any immersion  $u = (u_1, \dots, u_N) : M \rightarrow \mathbb{R}^N$  induces a tensor  $g_u = u^*g_E$  given by

$$(g_u)_{ij} = \partial_i u \cdot \partial_j u = \sum_{r=1}^N \partial_i u_r \partial_j u_r$$

Proving the isometric embedding theorem amounts to finding  $u$  such that

$$(4) \quad g_u = g$$

We will see in this section that our problem is one in analysis as well as geometry.

The following is the precise version of Nash's Isometric Embedding Theorems for compact Riemannian manifolds:

---

<sup>10</sup>Isometric embeddings have to be conformal embeddings, but conformal embeddings are weaker than isometric embeddings, as they preserve only angles, not necessarily distance.

THEOREM 3.1 ([42] Theorem 2). *Any compact Riemannian  $n$ -manifold with a  $C^s$  metric has a  $C^s$  isometric embedding in  $\mathbb{R}^{n(3n+11)/2}$  for  $3 \leq s \leq \infty$ .*

THEOREM 3.2 ([42] Theorem 3). *Any Riemannian manifold with a  $C^s$  metric has a  $C^s$  embedding in  $\mathbb{R}^{\frac{3}{2}n^3+7n^2+\frac{11}{2}n}$  for  $3 \leq s \leq \infty$ .*

Nash's isometric embedding theorems are not only geometric results: They are also *regularity*<sup>11</sup> results in the context of PDE's, and (4) is a nonlinear system of PDE's. In this paper, we will explore Günther's version of this result, which is similarly demanding in terms of regularity.

3.1. *Examples II:  $C^2$  vs  $C^1$  Embeddings on Spheres and Surfaces.* If the dimension  $N$  of the ambient Euclidean space is too low, a  $C^s$  metric may admit an isometric embedding of a lower differentiability class into  $\mathbb{R}^N$ , but not necessarily a  $C^s$  embedding. In particular, we will see that  $C^2$  is a substantially more rigid differentiability class than  $C^1$ .

We start with some definitions and a very interesting result about  $C^1$ -isometric embeddings by Nash and Nicolaas Kuiper:

DEFINITION 3.3. *A strictly short map  $f$  of Riemannian manifolds  $(M_1, g_1)$ ,  $(M_2, g_2)$  satisfies  $\|df(\mathbf{v})\|_{g_2(f(p))} < \|\mathbf{v}\|_{g_1(p)}$  for all  $p \in M_1$ ,  $\mathbf{v} \in T_p M_1$ .<sup>12</sup>*

The most general definition of a strictly short map is on metric spaces  $(X_1, d_1)$ , and  $(X_2, d_2)$ , where strict shortness of a map  $f : X_1 \rightarrow X_2$  is defined by the property  $d_1(x, y) > d_2(f(x), f(y))$  for all  $x, y \in X_1$ . On a Riemannian manifold, there is a notion of distance:

DEFINITION 3.4. *Let  $(M, g)$  be a Riemannian manifold. Let  $\gamma : [0, 1] \rightarrow M$  be a path between two points  $p_1$  and  $p_2$  (i.e.  $\gamma(0) = p_1$ ,  $\gamma(1) = p_2$ ). The length  $\ell(\gamma)$  of  $\gamma$  is given by*

$$\ell(\gamma) = \int_0^1 \|\gamma'(t)\|_{g(\gamma(t))} dt$$

The distance  $\text{dist}_g(p_1, p_2)$  between  $p_1$  and  $p_2$  with respect to the metric  $g$  is defined to be

$$\text{dist}_g(p_1, p_2) = \inf\{\ell(\gamma) : \gamma \text{ is piecewise smooth, } \gamma(0) = p_1, \gamma(1) = p_2\}$$

<sup>11</sup>*Regularity* is an umbrella term we use to refer to “niceness” of solutions to PDE's; examples of “niceness” include  $C^s$  differentiability and  $L_p$  integrability.

<sup>12</sup> $\|\cdot\|_{g(p)}$  and  $\|\cdot\|_{g_i(p)}$  are the metric-induced norms on  $M_i$  given by  $\|\mathbf{v}\|_{g_i(p)} = \sqrt{g_i(p)(\mathbf{v}, \mathbf{v})}$ .

Hence, shortness of maps on Riemannian manifolds holds only if the map is short on the underlying metric spaces with the distance metrics induced by the Riemannian metric tensors.

**THEOREM 3.5** (Nash–Kuiper, [19] Theorem 21.2.1). *For two Riemannian manifolds  $(M_1, g_1)$  and  $(M_2, g_2)$ <sup>13</sup> with  $\dim M_i = N_i$ , if  $N_1 < N_2$ , then any strictly short immersion (not necessarily isometric) can be  $C^0$ -approximated by a uniformly convergent sequence of isometric  $C^1$ -embeddings.*

This means that the  $C^1$  class of isometric embeddings is *very* flexible and that a large class of manifolds admit  $C^1$ -isometric embeddings of codimension just 1, for example. But many of these manifolds will not admit isometric embeddings of higher differentiability class.

*Example 3.6.* (Berger Sphere, revisited) In the previous section, we introduced the Berger metrics on the 3-sphere and mentioned that a corollary of Peng–Tang’s result in [46] is that infinitely many of these metrics admit no conformal  $C^2$ -embeddings into  $\mathbb{R}^4$ . But, we can easily find a short immersion of the Berger sphere into  $\mathbb{R}^4$ : Given a Berger metric  $g^\varepsilon$  on the sphere with  $\varepsilon \neq 1$ , we can use inclusion of  $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$  as the unit quaternions, then scale this inclusion so that the radius of the immersed sphere in  $\mathbb{R}^4$  is less than  $\varepsilon$ . In other words, we define our immersion  $f$  by

$$z \xrightarrow{f} rz \quad r < \varepsilon$$

where  $z$  is a unit quaternion. By comparing the Berger metrics  $g^\varepsilon$  with the metric on the immersed sphere, which is  $rg_{\text{can}}$ , it is simple to check that  $f$  is a short immersion. Therefore, we can apply Nash–Kuiper to conclude that there is a  $C^1$ -isometric embedding of the 3-sphere with the Berger metric  $g^\varepsilon$  into  $\mathbb{R}^4$  for any constant  $\varepsilon$ .

We will now show that any closed orientable surface admits metrics with isometric  $C^1$ -embeddings into  $\mathbb{R}^3$  but no isometric  $C^2$ -embeddings into  $\mathbb{R}^3$ .

---

<sup>13</sup>Closedness need not be assumed here.

It is known by the uniformization theorem<sup>14</sup> that every closed orientable surface admits a metric with Gaussian curvature 1, 0, or  $-1$ ; in particular, the sphere admits a metric with positive curvature 1, the torus admits a flat metric, and the  $g$ -holed tori for  $g \geq 2$  admit hyperbolic metrics (i.e. curvature  $-1$ ).<sup>15</sup>

*Example 3.7 ( $g$ -Holed Tori).* To see that an isometric  $C^1$ -embedding exists for the flat metric on the torus and the hyperbolic metric on the  $g$ -holed tori for  $g \geq 2$ , we can consider *normal coordinates*; the following is an example of what we mean.<sup>16</sup>

Take a small disk from the flat torus, a small disk from the hyperbolic surface, and a small disk from the corresponding embedded surface. Hyperbolic surfaces are locally modeled on the *Poincaré disk*<sup>17</sup>  $D_H$  with hyperbolic metric

$$g_H = \frac{4dx^2 + 4dy^2}{(1 - (x^2 + y^2))^2}$$

The flat torus is modelled on the flat disk<sup>18</sup>  $D_E$  with Euclidean metric

$$g_E = dx^2 + dy^2$$

On the embedded surfaces, one can take the coordinates on a small disk  $D_S$  to be orthogonal projection  $\pi_\perp$  onto the tangent plane so that locally the metric on  $D_S$  is of the form

$$g_S = d(x \circ \pi_\perp)^2 + d(y \circ \pi_\perp)^2 = \rho_1(x, y)dx^2 + \rho_2(x, y)dy^2$$

---

<sup>14</sup>See Chapter 11 [51] for more on the uniformization theorem, which states that every simply connected Riemann surface is biholomorphic to  $\mathbb{S}^2$ , the complex plane  $\mathbb{C}$ , or the upper half plane  $\mathbb{H}$ , which admit metrics with constant sectional (or Gaussian) curvature 1, 0, and  $-1$ , respectively. Tori can be realized as  $\mathbb{C}/\Lambda$  where  $\Lambda$  is an integer lattice in  $\mathbb{C}$ , and  $g$ -holed tori for  $g \geq 2$  can be realized as the quotient of  $H$  by a finite subgroup of  $PSL(2, \mathbb{R})$ , and these inherit the Riemannian structures (and therefore curvature as well) from their universal covers  $\mathbb{C}$  and  $\mathbb{H}$ .

<sup>15</sup>*Gaussian curvature* is the product of the minimum and maximum possible curvature of curves passing through a point rendered as cross sections of the surface with a plane containing a vector normal to the surface. In two dimensions, it is equal to the sectional curvature.

<sup>16</sup>For a more rigorous explanation of geodesics and normal coordinates, see Sections 14 and 15 in [56].

<sup>17</sup>The Poincaré disk is (biholomorphically) isometric to the upper half plane with metric  $\frac{dx^2 + dy^2}{y^2}$ .

<sup>18</sup>Usually, we depict a torus as a square with parallel edges identified, but for the sake of this discussion, we take the local picture to be a disk; one can think of this disk as a subset of the square.

for some smooth functions  $\rho_i$ . One can find that curves of minimal distance—or geodesics: On  $D_H$ , these are circular arcs perpendicular to the boundary of the disk and diameters, and on  $D_E$ , these are lines. The image of  $D_S$  through orthogonal projection is  $D_E$ , and one can find by staring at this picture that for three points  $p_1, p_2$ , and  $p_3$  on the disk from the embedded surface,

$$\text{dist}_{g_S}(p_1, p_2) > \text{dist}_{g_S}(p_1, p_3) \iff \text{dist}_{g_E}(\pi_\perp(p_1), \pi_\perp(p_2)) > \text{dist}_{g_E}(\pi_\perp(p_1), \pi_\perp(p_3))$$

The quantity

$$|\text{dist}_{g_S}(p_1, p_2) - \text{dist}_{g_E}(\pi_\perp(p_1), \pi_\perp(p_2))|$$

has an upper bound for all  $p_1, p_2$  by boundedness of the disk.

After staring some more—in particular, in the case of the hyperbolic surfaces, comparing  $(D_H, g_H)$  with  $(D_E, g_E)$ —one can conclude that obtaining short immersions of the flat torus and hyperbolic  $g$ -holed tori into  $\mathbb{R}^3$  is possible if one is able globally apply a shrinking of  $D_S$  to the entire embedded surface; we can do this simply by globally shrinking the embedded surfaces in  $\mathbb{R}^3$ , so Nash–Kuiper applies.

However, these metrics on the  $g$ -holed tori admit no  $C^2$ -embeddings into  $\mathbb{R}^3$  because any smooth hypersurface of  $\mathbb{R}^3$  with metric induced by the Euclidean metric must have at least one point with positive Gaussian curvature—in other words, it must be locally strictly convex around at least one point.<sup>19</sup> Therefore, we know automatically that the  $g$ -holed tori admit metrics that do not  $C^2$ -embed isometrically into  $\mathbb{R}^3$  with the Euclidean metric for all  $g \geq 1$ .

Next is an example of a metric on  $\mathbb{S}^2$  admitting no isometric  $C^2$ -embedding into  $\mathbb{R}^3$ .

*Example 3.8* (Greene’s Sphere, [24]). Consider the stereographic projection of  $\mathbb{S}^2 \setminus \{p_S\}$  onto  $\mathbb{R}^2$ , that takes southern hemisphere  $H_2 = \{z < 0\}$  onto the open unit disk on the  $xy$ -plane and the equator  $Z = \{z = 0\}$  onto the unit circle on the  $xy$ -plane.  $\mathbb{S}^2 \setminus \{p_S\}$  inherits a flat metric  $g_1$  from  $\mathbb{R}^2$ .

Then, for  $\varepsilon > 0$ , define the open annulus  $A_\varepsilon \subseteq \mathbb{S}^2 \setminus \{p_S\}$  by

$$A_\varepsilon = \{p \in \mathbb{S}^2 \setminus \{p_S\} : \inf_{q \in E} \text{dist}_{g_1}(q, p) < \varepsilon\}$$

Let  $f_\varepsilon : S \setminus \{p_S\}$  be a smooth bump function adjusted so that, where  $H_1 = \{z > 0\}$  is the northern hemisphere:

- $f_\varepsilon(p) = 1$  for  $p \in H_1 \cup A_{\varepsilon/2}$
- $f_\varepsilon(p) < \varepsilon$  for  $p \in \mathbb{S}^2 \setminus (H_1 \cup A_\varepsilon)$

---

<sup>19</sup>See Exercise 16 in Section 3.3 of [13].

- $0 < f_\varepsilon(p) \leq 1$  for  $p \in \mathbb{S}^2$

One can extend  $g_1|_{H_1 \cup A_{1/2}}$  to all of  $\mathbb{S}^2$  by partition of unity; call this new metric  $g_2$ . Then let  $g = f_\varepsilon g_2$  for  $\varepsilon$  small enough so that

$$\sup_{p_1, p_2} \text{dist}_g(p_1, p_2) \leq \frac{1}{2}$$

$g$  here is the  $C^\infty$  metric constructed by Robert Greene in [24] that admits no  $C^2$ -isometric embedding into  $\mathbb{R}^3$  with the Euclidean metric.

One can apply the idea of normal coordinates in a manner similar to that of Example 3.7 to show that one can find a short immersion of the sphere by shrinking the embedded sphere in  $\mathbb{R}^3$ , so by Nash–Kuiper, the sphere with the above metric  $g$  admits an isometric  $C^1$ -embedding.

His idea was to deform and flatten the lower hemisphere in such a way that the structures are still smooth, but all of the distances with respect to the sphere’s metric between points on the equator shrink—namely, all at most  $\frac{1}{2}$ .

Since the shortest path between two equatorial points traverses the flat part of the structure on the sphere, it must therefore be true that the distance between these points’ images in  $\mathbb{R}^3$  must also be at most  $\frac{1}{2}$ .

However, Greene proved that an obstruction to  $C^2$  isometric embeddings is that there must exist points on the equator whose images are at least  $\sqrt{3}$  apart in  $\mathbb{R}^3$ .

One fascinating project inspired by the Nash–Kuiper theorem is the [Hevea Project](#), which is dedicated to visualizing surfaces with metrics that admit only isometric  $C^1$ -embeddings into  $\mathbb{R}^3$  based on explicit constructions of the closed forms of such metrics. As of 2026, the project includes renderings of the flat torus, the hyperbolic plane— $\mathbb{R} \times \mathbb{R}^+$  endowed with the metric  $\frac{dx^2 + dy^2}{y^2}$ —and the reduced sphere—an object obtained by using Nash–Kuiper to perform a deformation in the  $C^1$ -family approximating the round metric so that we end up with a sphere of arbitrarily small radius but with the geodesic length preserved from the original unit sphere.

**3.2. Linearizing the Problem by Choice of Variations.** Returning to the general isometric embedding problem, we recall that, given a  $C^s$  embedding of a Riemannian  $n$ -manifold  $(M, g)$  into  $(\mathbb{R}^N, g_E)$ , the aim is to find a  $C^s$

embedding  $u$  such that

$$g_{ij} = (g_u)_{ij} = \sum_{r=1}^N \partial_i u_r \partial_j u_r \quad 1 \leq i, j \leq n$$

The above system of PDE's is very nonlinear; one cannot learn much by trying to solve directly for  $u$ . One useful method to get the correct regularity result for the above system is to consider the following *perturbation problem*: If  $v = (v_1, \dots, v_N) : M \rightarrow \mathbb{R}^N$  is a variation in the embedding inducing a change—or *perturbation*— $f$  in the metric, we get the following system of first-order PDE:

$$g_{u+v} = g_u + f \iff \sum_{r=1}^N \partial_i (u_r + v_r) \partial_j (u_r + v_r) = \sum_{r=1}^N \partial_i u_r \partial_j u_r + f_{ij} \quad 1 \leq i, j \leq n$$

$$(5) \iff \sum_{r=1}^N \partial_i u_r \partial_j v_r + \sum_{r=1}^N \partial_i v_r \partial_j u_r + \sum_{r=1}^N \partial_i v_r \partial_j v_r = h_{ij} \quad 1 \leq i, j \leq n$$

More precisely, if the perturbation  $f$  is  $C^s$ , we need to prove that there exists a  $C^s$  variation  $v$  inducing  $f$ .<sup>20</sup>

The system of PDE's (5) is *still* nonlinear. But our redefinition of our problem in terms of perturbations allows us to employ some tricks.

We define the embedding  $u$  to be *free* if the vectors

$$\begin{bmatrix} \partial_i u_1 \\ \vdots \\ \partial_i u_N \end{bmatrix} \quad i \in \{1, \dots, n\} \quad \begin{bmatrix} \partial_{ij} u_1 \\ \vdots \\ \partial_{ij} u_N \end{bmatrix} \quad 1 \leq i \leq j \leq n$$

are linearly independent.

Nash simplified this problem by considering *normal* variations: That is, our  $C^s$  variation  $v$  satisfies

$$\sum_{r=1}^N v_r \cdot \partial_j u_r = 0 \quad 1 \leq j \leq n$$

---

<sup>20</sup>Technically, as we will see in the statement of Günther's perturbation result (Theorem 4.1, the proof is for *sufficiently small*  $C^s$ -differentiable  $f$ ).

Then, by product rule,

$$\implies \sum_{r=1}^N \partial_i v_r \partial_j u_r + \sum_{r=1}^N v_r \partial_{ij} u_r = 0$$

Then, when  $v$  is a normal variation, our PDE becomes

$$-2 \sum_{r=1}^N v_r \partial_{ij} u_r + \sum_{r=1}^N \partial_i v_r \partial_j v_r = f_{ij} \quad 1 \leq i, j \leq n$$

Noting that the second term on the left is quadratic in  $v$  (hence for  $v$  small, it contributes negligibly), if our perturbation  $f = tf_0$ , the infinitesimal variation  $w = \partial_t v|_{t=0}$  will satisfy

$$(6) \quad \sum_{r=1}^N w_r \partial_{ij} u_r = -\frac{1}{2}(f_0)_{ij}, \quad \sum_{r=1}^N w_r \partial_j u_r = 0 \quad 1 \leq i, j \leq n$$

This is seen by taking the derivative with respect to  $t$  in the PDE and the normal condition on  $v$ .  $u$  is a known embedding, and the variation  $w_r$  is unknown—Nash has *linearized* this problem by constructing a linear system of equations for  $w_r$ .

There remains a key issue, however. If one looks carefully at the PDE, one would notice that if  $v$  is of class  $C^s$ , then  $f = g_{u+v} - g_u$  is of class  $C^{s-1}$  because of the first order partial derivatives; correspondingly, if  $h$  is of class  $C^{s-1}$ , then the normal infinitesimal variation  $w$  must also be of class  $C^{s-1}$ . So in this process, we have lost a degree of differentiability, which poses an obstacle to being able to estimate our Riemannian metric in the class  $C^s$ . We refer to this as the “loss of differentiability” problem.

Nash’s original solution to this problem emerged in the following way: Viewing  $v \mapsto g_{u+v} - g_u$  as a smooth map  $C^s(M, NM) \rightarrow C^{s-1}(M, \text{Sym}^2(T^*M))$  of Banach spaces,<sup>21</sup> one can think of our objective as attempting to prove a type of inverse function theorem. But our observations show that we cannot use the conventional inverse function theorem from real analysis to return a  $C^s$  variation from a  $C^s$  perturbation  $f$  because the inverse of the derivative of this map, given by the infinitesimal variation  $w$ , will not be bounded in  $C^s$ .

Nash proved the required inverse function theorem in [42]. Nash’s solution to the loss of differentiability problem turns out to be a momentous feat in mathematics; it ultimately allowed Nash to solve the perturbation problem

---

<sup>21</sup> $NM$  here is the *normal bundle* of  $M$  as an embedded submanifold into  $\mathbb{R}^N$ , defined to be the orthogonal complement of  $TM$  in  $T\mathbb{R}^N$ .

and prove Theorems 3.1 and 3.2. Nash's proof also had implications for other problems in geometry and analysis, which will be briefly discussed in Section 5.

The proof of this inverse function theorem is monstrous. Günther discovered a way to avoid the loss of differentiability, however. His solution will be the subject of Section 4.

3.3. *Hölder Norms.* Before we begin discussing a solution to the general isometric embedding problem, however, we need more definitions: In order to prove obtain a perturbation result with the right regularity, one must provide estimates in the proper  $C^s$ -differentiable class. To make estimations precise, it would be especially convenient if the spaces of functions we are working in are complete so that we can use tools such as Arzela–Ascoli to do approximations.

DEFINITION 3.9 (Hölder-differentiability). *Functions of class  $C^{s,\lambda}(D)$  on some domain  $D \subset \mathbb{R}$  are those  $f$  whose  $k$ th derivative must satisfies the following Hölder condition:*

$$\exists C \geq 0 : \forall x, y, |f(x) - f(y)| \leq C \|x - y\|^\lambda$$

This definition generalizes to functions on  $\mathbb{R}^n$ .

Unlike the space of  $C^s = C^{s,0}$  functions, the space of  $C^{s,\lambda}$  functions has a complete norm-induced topology for  $0 < \lambda < 1$ . Let  $0 < \lambda < 1$  be fixed. I use the following definition for  $C^{s,\lambda}$  norms given by Günther in [28] (generally referred to as Hölder norms):

DEFINITION 3.10. *For a real function  $u : B \rightarrow \mathbb{R}$  where  $B$  is an open  $m$ -ball around the origin, and for  $0 < \lambda < 1$ , we define the  $C^{0,\lambda}$  (or  $C^\lambda$ ) norm  $\|\cdot\|_0$  and the  $C^{s,\lambda}$  norm  $\|\cdot\|_s$  by*

$$\|u\|_0^B := \sup_{x \in B} |u(x)| + \sup_{x, y \in B, x \neq y} |x - y|^{-\lambda} |u(x) - u(y)|$$

$$\|u\|_s^B := \|u\|_0^B + \sum_{i_1, \dots, i_s=1}^m \|\partial_{i_1 \dots i_s} u\|_0^B$$

One simple but important fact that we will use to absorb lower-order norms into higher order norms is that

$$(7) \quad \|u\|_s^B \leq K \|u\|_{s+1}^B$$

Intuitively, this makes sense because  $C^{s+1,\lambda}$  is a better regularity than  $C^{s,\lambda}$ . To show this using the definition, we observe that

$$\begin{aligned} |\partial_{i_1 \dots i_s} u(x) - \partial_{i_1 \dots i_s} u(y)| &\leq \sum_j \left| \int_B \partial_{j i_1 \dots i_s} u(x) - \partial_{j i_1 \dots i_s} u(y) dx_j \right| \\ &\leq \text{Vol}(B) \sum_j \sup_{x,y \in B, x \neq y} |\partial_{j i_1 \dots i_s} u(x) - \partial_{j i_1 \dots i_s} u(y)| \end{aligned}$$

From this, we can conclude that

$$\begin{aligned} &\sum_{i_1, \dots, i_s=1}^m \sup_{x,y \in B, x \neq y} \frac{|\partial_{i_1 \dots i_s} u(x) - \partial_{i_1 \dots i_s} u(y)|}{|x - y|^\lambda} \\ &\leq K \sum_{i_1, \dots, i_{s+1}=1}^m \sup_{x,y \in B, x \neq y} \frac{|\partial_{i_1 \dots i_{s+1}} u(x) - \partial_{i_1 \dots i_{s+1}} u(y)|}{|x - y|^\lambda} \end{aligned}$$

Which proves  $\|u\|_s^B \leq K \|u\|_{s+1}^B$ .

**DEFINITION 3.11.** For a function  $v = (v_1, \dots, v_q) : B \rightarrow \mathbb{R}^q$ , we define its  $C^{s,\lambda}$  norm to be

$$\|v\|_s^B := \sum_{j=1}^q \|v_j\|_s^B$$

An important property that follows directly from the definitions is that for  $0 \leq r \leq s$ ,

$$\|u\|_s^B \leq \|u\|_0^B + \sum_{i_1, \dots, i_r=1} \|\partial_{i_1 \dots i_r} u\|_{s-r}^B \leq \|u\|_r^B + \|u\|_s^B$$

Now, we extend the concept of Hölder norms to manifolds using *partitions of unity*:

**DEFINITION 3.12.** Given a manifold  $M$  with an open cover  $\{U_\alpha\}$ , a *partition of unity subordinate to  $\{U_\alpha\}$*  is a collection of maps  $\{\psi_\alpha : U_\alpha \rightarrow \mathbb{R}\}$  such that  $\psi_\alpha \geq 0$ ,  $\{\text{supp } \psi_\alpha\}$  is locally finite, and  $\sum_\alpha \psi_\alpha = 1$ .

**DEFINITION 3.13.** Given a manifold  $M$  with a  $\{\psi_\alpha\}$  subordinate to a  $C^\infty$  atlas  $\{U_\alpha\}$ :

$$\|v\|_s = \sum_{\alpha,j} \|\psi_\alpha v_j\|_s^B$$

Property (7) holds true on manifolds as well:

$$(8) \quad \|u\|_s \leq K \|u\|_{s+1}^B$$

Our norms have two other useful properties: For two functions  $u, v$  on  $M$ , and for  $0 \leq r < s$ ,

$$(9) \quad \|uv\|_s \leq C_s \|u\|_s \|v\|_s$$

$$(10) \quad \|uv\|_s \leq K(\|u\|_s \|v\|_r + \|u\|_r \|v\|_s) + C_s \|u\|_{s-1} \|v\|_{s-1}$$

Now that we have defined the proper norms and listed a few useful properties, we are ready to tackle the geometric PDE problem of isometric embeddings.

#### 4. Günther's Solution

We have now arrived at the center of our discussion, Günther's solution to Nash's isometric Embedding Problem. Günther's perturbation result comes in two versions, Theorem 4.1 and Theorem 4.12. We will discuss in detail the proof of Theorem 4.1, then we will briefly sketch the proof of variant Theorem 4.12 and indicate how it can be applied to the isometric embedding theorem, which requires Theorem 4.15.

4.1. *Günther's First Perturbation Result and the Associated Fixed-Point Problem.* From now until the end of Section 4.4, we assume that  $M$  is a compact Riemannian manifold. We define  $E(u)(h, f)$  for a free mapping  $u \in C^\infty(M, \mathbb{R}^N)$ ,  $h \in C^{s, \lambda}(M, T^*M)$ , and  $f \in C^{s, \lambda}(M, \text{Sym}^2(T^*M))$  to be the unique solution  $v$  to the system

$$(11) \quad \nabla_i u \cdot v = h_i \quad \nabla_i \nabla_j u \cdot v = f_{ij}$$

so that for all  $x \in M$ ,  $\|v(x)\|$  is minimal. One can define  $v = E(u)(h, f)$  point-wise by pseudo-inverse.

Given our linear system, we can write  $v = (v_1, \dots, v_N)$  as

$$v_\alpha = A^{\alpha, i}(u)h_i + B^{\alpha, ij}(u)f_{ij} \quad \alpha = 1, \dots, N$$

where  $A^{\alpha, i}(u)$  is a smooth contravariant vector field and  $B^{\alpha, ij}(u)$  is a smooth symmetric  $(0, 2)$  tensor field on  $M$ . Then, define

$$D(u) := \max \left( \sum_{\alpha=1}^N \|A^\alpha(u)\|_2, \sum_{\alpha=1}^N \|B^\alpha(u)\|_2 \right)$$

Günther's first perturbation result states:

**THEOREM 4.1** ([28] Theorem / [27] Theorem 1). *Let  $M$  be a compact Riemannian manifold. For a free mapping  $u \in C^\infty(M, \mathbb{R}^N)$ ,  $f \in C^{s, \lambda}(M, \text{Sym}^2(T^*M))$  with  $s \geq 2$  (resp.  $f \in C^\infty(M, \text{Sym}^2(T^*M))$ ), there exists some  $\theta > 0$  independent of  $u, s, f$  such that if*

$$D(u)\|E(u)(0, f)\|_2 \leq \theta$$

then there exists a  $v \in C^{s,\lambda}(M, \mathbb{R}^N)$  (resp.  $v \in C^\infty(M, \mathbb{R}^N)$ ) with  $\|v\|_2 \leq \|E(u)(0, f)\|_2$  such that

$$\partial_i u \cdot \partial_j v + \partial_j u \cdot \partial_i v + \partial_i v \cdot \partial_j v = f_{ij}$$

The purpose of the next few sections is to build up to the proof of this theorem, which will be provided in Section 4.4.

To prove this theorem, Günther’s strategy was to express  $v$  as a solution to a fixed-point problem, as motivated by the following definition and theorem<sup>22</sup>:

DEFINITION 4.2. *Let  $(X, d)$  be a metric space. Then  $\mathcal{F} : X \rightarrow X$  is a contraction mapping if there exists a Lipschitz constant  $\theta < 1$  such that  $d(\mathcal{F}(x), \mathcal{F}(y)) \leq \theta d(x, y)$  for all  $x, y \in X$ .*

THEOREM 4.3 (Fixed Point Theorem, Theorem 5.1 in [23]). *If  $(X, d)$  is a nonempty complete metric space and  $E : X \rightarrow X$  is a contraction mapping, then  $E$  admits a unique fixed point  $y$  such that  $E(y) = y$ , which may be realized as the limit of the recursively defined sequence  $x_{n+1} = E(x_n)$ .*

Günther’s “contraction mapping” is the pseudoinverse  $E(u)$ : By Lemma 4.10, which will be proven in Section 4.3, we are able to express the desired variation  $v$  as a solution to a fixed-point problem  $v = E(u)(h(v), f(v))$ . Proving Theorem 4.1 essentially amounts to finding a Lipschitz-type constant  $\theta$ .

We will assume the Hölder condition—that is, wherever  $\lambda$  appears,  $0 < \lambda < 1$ . We need the completeness of Hölder spaces to prove the convergence of a recursive sequence of the form  $v_{k+1} = E(u)(h(v_k), f(v_k))$  by applying the Arzela–Ascoli theorem.

Critical in this set up is that the LHS of (11) are *linear* in  $v$ ; recall that in the previous section, the system (6) was linear not in  $v$  but in the infinitesimal variation  $w$ . Because  $u$  is a free mapping, all the derivatives  $\nabla_i u$  and  $\nabla_i \nabla_j u$  are linearly independent, so solutions exist if the ambient dimension  $N$  is large enough, given that the RHS is in the correct subspace. Günther realized that if one can express our desired variation  $v$  as a solution to such a fixed-point problem, one avoids *any* loss in differentiability.

Once we know that we can express the variation  $v$  in this way, we need to be able to estimate  $v$  in the correct differentiability class. In other words, we

---

<sup>22</sup>In [23], the statement of Theorem 4.3 is given for Banach spaces, but observing the proof reveals that the structure of a normed vector space is not required; the proof works for any nonempty complete metric space.

need to be able to bound the pseudo-inverse. We observe that

$$\begin{aligned} \|v_\alpha\|_2 &\leq \|A^{\alpha,i}(u)h_i\|_2 + \|B^{\alpha,ij}(u)f_{ij}\|_2 \leq C_2\|A^{\alpha,i}(u)\|_2\|h_i\|_2 + C_2\|B^{\alpha,ij}(u)\|_2\|f_{ij}\|_2 \\ (12) \quad &\implies \|E(u)(h, f)\|_2 \leq KD(u)(\|h\|_2 + \|f\|_2) \end{aligned}$$

This inequality almost generalizes for  $s > 2$ :

LEMMA 4.4 ([28] Lemma 5). *For our current assumptions on  $u$ ,  $f$ , and  $h$  and  $s \geq 3$ ,*

$$\|E(u)(h, f)\|_s \leq KD(u)(\|h\|_s + \|f\|_s) + C(u, s)(\|h\|_{s-1} + \|f\|_{s-1})$$

$K$  is independent of  $u$  and  $s$ .

*Proof.* Note that by (10) and (8),

$$\begin{aligned} \|A^\alpha(u)h\|_s &\leq K(\|A^\alpha(u)\|_s\|h\|_2 + \|A^\alpha(u)\|_2\|h\|_s) + C_s\|A^\alpha(u)\|_{s-1}\|h\|_{s-1} \\ &\leq KD(u)\|h\|_s + C(u, s)\|h\|_{s-1} \end{aligned}$$

Analogously,  $\|B^\alpha(u)f\|_s \leq KD(u)\|f\|_s + C(u, s)\|f\|_{s-1}$ . So we can conclude that

$$\|E(u)(h, f)\|_s \leq \|A^\alpha(u)h\|_s + \|B^\alpha(u)f\|_s \leq KD(u)(\|h\|_s + \|f\|_s) + C(u, s)(\|h\|_{s-1} + \|f\|_{s-1})$$

□

4.2. *Strictly Elliptic Operators and The Screened Poisson Equation.* Elliptic operators, which are particular types of 2nd order differential operators, are nice because much is known about their solutions to their associated PDE's. The convenient properties of the Laplacian as an elliptic operator are key to Günther's proof of Theorem 4.1.

$$(13) \quad \Delta x - \nu x = f \quad \nu \geq 0$$

Consider equation (13) on the unit ball in Euclidean space. This is known as the *screened Poisson equation*. Given boundary conditions, the above PDE has a unique solution, so the operator  $\Delta - \nu\mathbf{1}$  has a well-defined bounded inverse  $(\Delta - \nu\mathbf{1})^{-1} : C^{s,\lambda} \rightarrow C^{s+2,\lambda}$ .

This is true in general for a class of *strictly elliptic operators*  $\mathcal{L}$ , defined below:

DEFINITION 4.5. *A second-order differential operator*

$$\mathcal{L}u = a_{ij}\partial_{ij}u + b_i\partial_iu + cu$$

*is strictly elliptic if  $a_{ij}$  satisfy  $a_{ij}(x)\xi_i\xi_j \geq \kappa\|\xi\|^2$  for all  $\xi \in \mathbb{R}^n$  and some  $\kappa > 0$ .*

We have the following result for strictly elliptic operators:

THEOREM 4.6. *If  $\mathcal{L}$  is a strictly elliptic 2nd order differential operator*

$$\mathcal{L}u := a_{ij}\partial_{ij}u + b_i\partial_iu + cu$$

*satisfying  $\|a_{ij}\|_0, \|b_i\|_0, \|c\|_0 \leq \Lambda$  for some  $\Lambda > 0$ , then for a bounded open set  $E \subset \mathbb{R}^n$  with smooth boundary, then there exists  $\eta \geq 0$  such that for  $\nu \geq \eta$ , the associated Dirichlet problem*

$$\begin{cases} \mathcal{L}u - \nu u = f & \bar{E} \\ u = \varphi & \partial E \end{cases}$$

*has a unique solution  $u \in C^{s+2,\lambda}$  for  $f \in C^{s,\lambda}$ .*

To understand why this is true, see for instance Theorem 3 of Section 6.2.2 in [21], which shows uniqueness, and Schauder's global estimate, which shows that the solution is in the right differentiability class (Theorem 6.6 in [23]). Necessary for Schauder's global estimate is the interior estimate, which we will state and use shortly.

On our compact Riemannian manifold  $M$ , we let  $g_0$  be an auxiliary metric on  $M$  and define the covariant Laplacian on tensors by<sup>23</sup>  $\Delta = \nabla^k \nabla_k = g_0^{kl} \nabla_l \nabla_k$ . As in the case of  $\mathbb{R}^n$ ,  $\Delta - \mathbf{1}$  on  $M$  also has a well-defined bounded inverse  $G := (\Delta - \mathbf{1})^{-1}$ .<sup>24</sup> Recall that by definition of bounded operators on Banach spaces, we have for any  $C^{s,\lambda}$  tensor field  $y$ ,

$$(14) \quad \|Gy\|_{s+2} \leq C_s \|y\|_s$$

We can get more specific estimates on  $G$  by applying an important and useful result about elliptic operators:

THEOREM 4.7 (Schauder's Interior Estimate, Theorem 6.2 in [23]). *If  $\mathcal{L}$  is a strictly elliptic 2nd order differential operator on Euclidean functions of  $n$  variables*

$$\mathcal{L}u := a_{ij}\partial_{ij}u + b_i\partial_iu + cu$$

*satisfying  $\|a_{ij}\|_0, \|b_i\|_0, \|c\|_0 \leq \Lambda$  for some  $\Lambda > 0$ , then for a bounded open set  $E \subset \mathbb{R}^n$ , if  $u \in C^2(E)$  solves*

$$\mathcal{L}u = f$$

*for some  $f \in C^{0,\lambda}$ , then  $u \in C^{2,\lambda}$  and*

$$\|u\|_2^E \leq K(n, \lambda, \kappa, \Lambda) \left( \|u\|_0^E + \|f\|_0^E \right)$$

<sup>23</sup>We use the Einstein summation convention here and in the proofs of this section.

<sup>24</sup>In the original paper, Günther used the Laplace operator defined by Lichnerowicz in [35]; we will be using the covariant Laplacian, as used in [2].

We can apply Schauder's interior estimate to manifolds by identifying coordinate neighborhoods on  $M$  with open sets of  $\mathbb{R}^n$ . We get the following result:

LEMMA 4.8 ([28] Lemma 3). *For  $h \in C^{s,\lambda}(M, T^*M)$  and  $f \in C^{s,\lambda}(\text{Sym}^2(T^*M))$  with  $s \geq 1$ ,*

$$\|Gh\|_{s+2} \leq K\|h\|_s + C_s\|h\|_{s-1}$$

$$\|Gf\|_{s+2} \leq K\|f\|_s + C_s\|f\|_{s-1}$$

*Proof.* I prove this for  $f \in \text{Sym}^2(T^*M)$  using local coordinates; the proof for vector fields is analogous.

As before, let  $\{(U_\alpha, \varphi_\alpha)\}$  be an atlas, and  $\{\psi_\alpha\}$  be a smooth partition of unity subordinate to  $\{U_\alpha\}$ . The covariant Laplace operator is written in local coordinates as  $g_0^{ab}\partial_a\partial_b$ . Let  $z = Gf$  and  $\partial^{(s)} = \partial_{i_1} \dots \partial_{i_s}$  for an arbitrary collection of indices  $i_1, \dots, i_s$ .

$$\begin{aligned} & g_0^{ab}\partial_a\partial_b\partial^{(s)}(\psi_\alpha z_{ij}) \\ = & \partial_{i_s} \left( g_0^{ab}\partial^{(s-1)}\partial_a\partial_b(\psi_\alpha z_{ij}) \right) - \left( \partial_{i_s} g_0^{ab} \right) \partial^{(s-1)}\partial_a\partial_b(\psi_\alpha z_{ij}) \\ = & \partial_{i_s} \left( \partial_{i_{s-1}} \left( g_0^{ab}\partial^{(s-2)}\partial_a\partial_b(\psi_\alpha z_{ij}) \right) \right) - \left( \partial_{i_s}\partial_{i_{s-1}} g_0^{ab} \right) \left( \partial^{(s-2)}\partial_a\partial_b(\psi_\alpha z_{ij}) \right) \\ & - \left( \partial_{i_{s-1}} g_0^{ab} \right) \left( \partial_{i_s}\partial^{(s-2)}\partial_a\partial_b(\psi_\alpha z_{ij}) \right) - \left( \partial_{i_s} g_0^{ab} \right) \partial^{(s-1)}\partial_a\partial_b(\psi_\alpha z_{ij}) \\ & \vdots \end{aligned}$$

In the expression after the first equals sign, the derivative is of order  $s+1$  on  $z_{ij}$ , and the last three terms in the expression after the second equals sign have derivatives of order  $\leq s+1$  on  $z_{ij}$ . This pattern continues as we move  $g_0^{ab}$  inside the derivatives using the product rule. This means

$$g_0^{ab}\partial_a\partial_b\partial^{(s)}(\psi_\alpha z_{ij}) = \partial^{(s)}(\psi_\alpha f_{ij}) + T_{ij}^{\alpha,(s)} z$$

where  $T_{ij}^\alpha$  is a differential operator of  $\leq s+1$ .

Recall that we define the  $C^{s,\lambda}$  norms on a (compact) manifold by  $\|f\|_s := \sum_{\alpha,i,j} \|\psi_\alpha f_{ij}\|_s^B$ . The norms defined under different partitions of unity are equivalent. Choosing a partition of unity  $\tilde{\psi}_\alpha$  so that  $\tilde{\psi}_\alpha(B') = 1$  for an open ball  $B' \subseteq B$ , we get  $\|f\|_s^{B'} \leq \|\tilde{\psi}_\alpha f_{ij}\|_s^B \leq \|f\|_s$ . Let  $B_1, B_2$  be open balls so that  $\text{supp } \psi_\alpha \subset \bar{B}_1 \subseteq B_2 \subseteq \bar{B}_2 \subseteq B$ . Similarly, we get

$$(15) \quad \sum_{(s)} \left\| T_{ij}^{A,(s)} z \right\|_0^{B_2} \leq \sum_{(s)} \left\| T_{ij}^{A,(s)} z \right\|_0 \leq C_s \|z\|_{s+1}$$

The Laplacian  $g_0^{ab}\partial_a\partial_b$  is a linear elliptic differential operator, which means that we can apply Schauder's interior estimate (Theorem 4.7):

$$(16) \quad \|\partial^{(s)}(\psi_\alpha z_{ij})\|_2^{B_1} \leq K \left( \left\| \partial^{(s)}(\psi_\alpha f_{ij}) + T_{ij}^{\alpha,(s)} z \right\|_0^{B_2} + \left\| \partial^{(s)}(\psi_\alpha z_{ij}) \right\|_0^{B_2} \right)$$

Note that because  $g_0$  is an auxiliary metric defined independently of  $u$ , the constant  $K$  also does not depend on  $u$ .

Applying definitions of norms, we get

$$(17) \quad \|\psi_\alpha z_{ij}\|_0^{B_1} \leq \|z\|_{s+1}$$

$$(18) \quad \sum_{(s)} \|\partial^{(s)}(\psi_\alpha z_{ij})\|_0^{B_2} \leq \|\psi_\alpha z_{ij}\|_s^B \leq C_s \|z\|_{s+1}$$

By (15), (16), (17), and (18):

$$\begin{aligned} & \|\psi_\alpha z_{ij}\|_{s+2}^B \\ & \leq \|\psi_\alpha z_{ij}\|_0^{B_1} + \sum_{(s)} \|\partial^{(s)}(\psi_\alpha z_{ij})\|_2^{B_1} \\ & \leq \|\psi_\alpha z_{ij}\|_0^{B_1} + K \sum_{(s)} \left( \left\| \partial^{(s)}(\psi_\alpha f_{ij}) + T_{ij}^{\alpha,(s)} z \right\|_0^{B_2} + \left\| \partial^{(s)}(\psi_\alpha z_{ij}) \right\|_0^{B_2} \right) \\ & \leq K \left( \|\psi_\alpha f_{ij}\|_0^{B_2} + \sum_{(s)} \left\| \partial^{(s)}(\psi_\alpha f_{ij}) \right\|_0^{B_2} \right) + \|\psi_\alpha z_{ij}\|_0^{B_1} + K \sum_{(s)} \left( \left\| T_{ij}^{\alpha,(s)} z \right\|_0^{B_2} + \left\| \partial^{(s)}(\psi_\alpha z_{ij}) \right\|_0^{B_1} \right) \\ & \leq K \|\psi_\alpha f_{ij}\|_s^B + C_s \|z\|_{s+1} \end{aligned}$$

$$\begin{aligned} \implies \|Gf\|_{s+2} &= \sum_{\alpha,i,j} \|\psi_\alpha z_{ij}\|_{s+2}^B \leq K \sum_{\alpha,i,j} \|\psi_\alpha f_{ij}\|_s + C_s \|Gf\|_{s+1} \\ &= K \|f\|_s + C_s \|Gf\|_{s+1} \leq K \|f\|_s + C'_s \|f\|_{s-1} \end{aligned}$$

□

4.3. *Günther's Lemmas.* Now, we provide the proofs of the lemmas necessary to set up the fixed-point problem for Theorem 4.1. The expression of the desired variation  $v$  as a solution to a fixed-point problem follows from Lemmas 4.9 and 4.10; Lemma 4.11 provides the remaining estimates necessary to solve the fixed-point problem.

First, we prove a useful identity. Now, define the vector field  $N(v)$  and the symmetric  $(0, 2)$  tensor field  $L(v)$  by:

$$(19) \quad N_i(v) := -\Delta v \cdot \nabla_i v$$

(20)

$$L_{ij}(v) := 2\nabla^k \nabla_i v \cdot \nabla_k \nabla_j v - 2\Delta v \cdot \nabla_i \nabla_j v - \nabla_i v \cdot \nabla_j v + R_i^p \nabla_p v \cdot \nabla_j v + R_j^p \nabla_p v \cdot \nabla_i v$$

$$1 \leq i, j \leq n$$

The above two definitions are rather unintuitive, but as we will see in our first two lemmas, they are helpful in writing the desired variation as a solution to a fixed-point problem.

LEMMA 4.9 ([28] Lemma 1). *For  $u \in C^\infty(M, \mathbb{R}^N)$ ,  $v \in C^2(M, \mathbb{R}^N)$ , and  $1 \leq i, j \leq n$ ,*

$$\begin{aligned} & (\Delta - \mathbf{1})(\nabla_i u \cdot \nabla_j v + \nabla_j u \cdot \nabla_i v + \nabla_i v \cdot \nabla_j v) \\ &= \nabla_i ((\Delta - \mathbf{1})(\nabla_j u \cdot v) - N_j(v)) + \nabla_j ((\Delta - \mathbf{1})(\nabla_i u \cdot v) - N_i(v)) \\ & \quad - 2(\Delta - \mathbf{1})(\nabla_i \nabla_j u \cdot v) + L_{ij}(v) + R_j^p \nabla_p (\nabla_i u \cdot v) + R_i^p \nabla_p (\nabla_j u \cdot v) \end{aligned}$$

*Remark.* Because the Laplacian does not commute with covariant differentiation, we get some curvature terms:<sup>25</sup>

$$\begin{aligned} & [\Delta, \nabla_j] \\ &= g_0^{kl} \nabla_k \nabla_l \nabla_j - \nabla_j g_0^{kl} \nabla_k \nabla_l \\ &= g_0^{kl} (\nabla_k \nabla_l \nabla_j - \nabla_j \nabla_k \nabla_l) \\ &= g_0^{kl} (\nabla_k [\nabla_l, \nabla_j] + \nabla_k \nabla_j \nabla_l - \nabla_j \nabla_k \nabla_l) \\ &= g_0^{kl} (\nabla_k [\nabla_l, \nabla_j] + [\nabla_k, \nabla_j] \nabla_l) \end{aligned}$$

For a sufficiently differentiable function  $a$ , we get

$$\begin{aligned} & \implies [\Delta, \nabla_j]a \\ &= g_0^{kl} (\nabla_k [\nabla_l, \nabla_j]a + [\nabla_k, \nabla_j] \nabla_l a) \\ &= g_0^{kl} (R_{kjl}^p \nabla_p a) \\ &= g_0^{kl} (R_{kjlq} g_0^{pq} \nabla_p a) \\ &= R_{jq} g_0^{pq} \nabla_p a \\ &= R_j^p \nabla_p a \end{aligned}$$

---

<sup>25</sup>That the metric commutes with the covariant derivative is equivalent to the metric compatibility condition on the Levi-Civita connection:

$$\nabla_Z(g_0(X, Y)) = Z(g_0(X, Y)) - g_0(\nabla_Z X, Y) - g_0(X, \nabla_Z Y) = 0$$

So in particular, when expanding  $\nabla_j(g_0^{kl} \nabla_k \nabla_l) = (\nabla_j g_0^{kl}) \nabla_k \nabla_l + g_0^{kl} \nabla_j \nabla_k \nabla_l$ , the term  $\nabla_j g_0^{kl} \nabla_k \nabla_l$  vanishes because  $\nabla_j g_0^{kl} = 0$ .

Regarding the proof, some relevant calculations are shown in Section 6.3 of [2]. Nevertheless, we go step-by-step in the proof below.

*Proof.* Using the product rule, we have:

$$\nabla_j(\nabla_i u \cdot v) + \nabla_i(\nabla_j u \cdot v) - 2\nabla_i \nabla_j u \cdot v = \nabla_i u \cdot \nabla_j v + \nabla_j u \cdot \nabla_i v$$

Applying this observation as well as the commutation formula from the remark, we get

$$\begin{aligned} & (\Delta - \mathbf{1})(\nabla_j(\nabla_i u \cdot v) + \nabla_i(\nabla_j u \cdot v)) \\ = & \nabla_j((\Delta - \mathbf{1})(\nabla_i u \cdot v)) + R_j^p \nabla_p(\nabla_i u \cdot v) + \nabla_i((\Delta - \mathbf{1})(\nabla_j u \cdot v)) + R_i^p \nabla_p(\nabla_j u \cdot v) \\ & (\Delta - \mathbf{1})(\nabla_i v \cdot \nabla_j v) \\ = & \nabla^k(\nabla_k \nabla_i v \cdot \nabla_j v + \nabla_i v \cdot \nabla_k \nabla_j v) - \nabla_i v \cdot \nabla_j v \\ = & \Delta \nabla_i v \cdot \nabla_j v + 2\nabla_k \nabla_i v \cdot \nabla^k \nabla_j v + \nabla_i v \cdot \Delta \nabla_j v - \nabla_i v \cdot \nabla_j v \\ = & \nabla_i(\Delta v \cdot \nabla_j v) + \nabla_j(\Delta v \cdot \nabla_i v) - 2\Delta v \cdot \nabla_i \nabla_j v + R_i^p \nabla_p v \cdot \nabla_j v + R_j^p \nabla_p v \cdot \nabla_i v \\ & + 2\nabla_k \nabla_i v \cdot \nabla^k \nabla_j v - \nabla_i v \cdot \nabla_j v \end{aligned}$$

Recalling the definition of  $N(v)$  and  $L(v)$  and combining the above calculations we get

$$\begin{aligned} & (\Delta - \mathbf{1})(\nabla_i u \cdot \nabla_j v + \nabla_j u \cdot \nabla_i v + \nabla_i v \cdot \nabla_j v) \\ = & (\Delta - \mathbf{1})(\nabla_j(\nabla_i u \cdot v) + \nabla_i(\nabla_j u \cdot v) - 2\nabla_i \nabla_j u \cdot v + \nabla_i v \cdot \nabla_j v) \\ = & \nabla_i((\Delta - \mathbf{1})(\nabla_j u \cdot v) - N_j(v)) + \nabla_j((\Delta - \mathbf{1})(\nabla_i u \cdot v) - N_i(v)) \\ & - 2(\Delta - \mathbf{1})(\nabla_i \nabla_j u \cdot v) + L_{ij}(v) + R_j^p \nabla_p(\nabla_i u \cdot v) + R_i^p \nabla_p(\nabla_j u \cdot v) \end{aligned}$$

□

Now, define the symmetric  $(0, 2)$ -tensor field  $M$ :

$$(21) \quad M_{ij}(v) = L_{ij}(v) + R_j^p \nabla_p (GN(v))_i + R_i^p \nabla_p (GN(v))_j$$

LEMMA 4.10 ([28] Lemma 2). For  $u \in C^\infty(M, \mathbb{R}^N)$ ,  $v \in C^{2,\lambda}(M, \mathbb{R}^N)$ ,  $f \in C^{2,\lambda}(\text{Sym}^2(T^*M))$ , and  $1 \leq i, j \leq n$ , if

$$\nabla_i u \cdot v = (GN(v))_i \quad \nabla_i \nabla_j u \cdot v = -\frac{1}{2}f_{ij} + \frac{1}{2}(GM(v))_{ij}$$

Then we have

$$\nabla_i u \cdot \nabla_j v + \nabla_j u \cdot \nabla_i v + \nabla_i v \cdot \nabla_j v = f_{ij}$$

*Proof.* If we make these substitutions in the identity from Lemma 4.9,

$$\nabla_i((\Delta - \mathbf{1})(\nabla_j u \cdot v) - N_j(v)) + \nabla_j((\Delta - \mathbf{1})(\nabla_i u \cdot v) - N_i(v))$$

vanishes. Then we have

$$\begin{aligned} & (\Delta - \mathbf{1})(\nabla_i u \cdot \nabla_j v + \nabla_j u \cdot \nabla_i v + \nabla_i v \cdot \nabla_j v) \\ &= -2(\Delta - \mathbf{1})(\nabla_i \nabla_j u \cdot v) + M_{ij}(v) \\ &= (\Delta - \mathbf{1})(f_{ij} - (GM(v))_{ij}) + M_{ij}(v) \\ &= (\Delta - \mathbf{1})(f_{ij}) \\ &\implies \nabla_i u \cdot \nabla_j v + \nabla_j u \cdot \nabla_i v + \nabla_i v \cdot \nabla_j v = f_{ij} \end{aligned}$$

□

By this lemma, we can prove Theorem 4.1 by solving the following fixed-point problem:

$$v = E(u) \left( GN(v), -\frac{1}{2}f + \frac{1}{2}GM(v) \right)$$

Now, we provide estimates for  $N$  and  $M$ :

LEMMA 4.11 ([28] Lemma 4). For  $v \in C^{s+2,\lambda}(M, \mathbb{R}^N)$  with  $s \geq 0$ ,

$$\|N(v)\|_s, \|M(v)\|_s \leq K\|v\|_2\|v\|_{s+2} + C_s\|v\|_{s+1}^2$$

For  $v_1, v_2 \in C^{2,\lambda}(M, \mathbb{R}^N)$ ,

$$\|N(v_1) - N(v_2)\|_0, \|M(v_1) - M(v_2)\|_0 \leq K(\|v_1\|_2 + \|v_2\|_2)\|v_1 - v_2\|_2$$

*Proof.* Let  $v \in C^{s+2,\lambda}(M, \mathbb{R}^N)$ . Applying the definitions and properties of  $C^{s,\lambda}$  norms gives us

$$(22) \quad \|\nabla_i \nabla_j v\|_0 \leq K\|v\|_2$$

$$(23) \quad \|\nabla_i \nabla_j v\|_s \leq K\|v\|_{s+2} + C_s\|v\|_{s+1}$$

We observe that by (23), (22), (10), and (8)

$$\begin{aligned} & \|N(v)\|_s \\ &= \|\Delta v \cdot \nabla_i v\|_s \\ &\leq K(\|\Delta v\|_s\|\nabla_i v\|_0 + \|\Delta v\|_0\|\nabla_i v\|_s) + C_s\|\Delta v\|_{s-1}\|\nabla_i v\|_{s-1} \\ &\leq K((K\|v\|_{s+2} + C_s\|v\|_{s+1})K\|v\|_1 + K\|v\|_2\|v\|_{s+1}) + C_s(K\|v\|_{s+1} + C_{s-1}\|v\|_s)K\|v\|_s \\ &\leq K'\|v\|_2\|v\|_{s+2} + C_s\|v\|_{s+1}^2 \end{aligned}$$

Next, we observe that by (22), (23), and (9),

$$(24) \quad \|\nabla_i v \cdot \nabla_j v\|_s \leq C_s\|v\|_{s+1}^2$$

$$\begin{aligned}
& \|\nabla_p GN(v)\|_s \leq K \|GN(v)\|_{s+1} \leq K' \|N(v)\|_{s-1} \leq K'' \|v\|_{s+1} \|v\|_2 + C_s \|v\|_s^2 \\
(25) \quad & \implies \|\nabla_p GN(v)\|_s \leq K \|v\|_{s+1} \|v\|_2 + C_s \|v\|_s^2
\end{aligned}$$

$$\begin{aligned}
& \|\nabla_i \nabla_j v \cdot \nabla_k \nabla_l v\|_s \\
& \leq K (\|\nabla_i \nabla_j v\|_s \|\nabla_k \nabla_l v\|_0 + \|\nabla_i \nabla_j v\|_0 \|\nabla_k \nabla_l v\|_s) + C_s \|\nabla_i \nabla_j v\|_{s-1} \|\nabla_i \nabla_j v\|_{s-1} \\
& \leq K' \|v\|_{s+2} \|v\|_2 + C_s \|v\|_{s+1}^2
\end{aligned}$$

$$(26) \quad \implies \|\nabla_i \nabla_j v \cdot \nabla_k \nabla_l v\|_s \leq K \|v\|_{s+2} \|v\|_2 + C_s \|v\|_{s+1}^2$$

These above are all of the types of terms that appear in the definition of  $M(v)$ . Some are multiplied by a Ricci curvature term, but because we are in the setting of compact manifolds, these terms are bounded, and the Ricci curvature is defined by the auxiliary metric  $g_0$ , independent of  $u$ . Hence we have by (24), (25), (26), and the triangle inequality that

$$\|M(v)\|_s \leq K \|v\|_2 \|v\|_{s+2} + C_s \|v\|_{s+1}^2$$

Now, we see that by triangle inequality, (9), and (22),

$$\begin{aligned}
& \|N(v_1) - N(v_2)\|_0 \\
& = \|\Delta v_1 \cdot \nabla_i v_1 - \Delta v_2 \cdot \nabla_i v_2\|_0 \\
& = \|\Delta(v_1 - v_2) \cdot \nabla_i v_1 + \Delta v_2 \cdot \nabla_i(v_1 - v_2)\|_0 \\
& \leq \|\Delta(v_1 - v_2) \cdot \nabla_i v_1\|_0 + \|\Delta v_2 \cdot \nabla_i(v_1 - v_2)\|_0 \\
& \leq K \|\Delta(v_1 - v_2)\|_0 \|\nabla_i v_1\|_0 + K \|\Delta v_2\|_0 \|\nabla_i(v_1 - v_2)\|_0 \\
& \leq K \|v_1 - v_2\|_2 (\|v_1\|_2 + \|v_2\|_2) \\
(27) \quad & \implies \|N(v_1) - N(v_2)\|_0 \leq K \|v_1 - v_2\|_2 (\|v_1\|_2 + \|v_2\|_2)
\end{aligned}$$

For very similar reasons

$$\begin{aligned}
& \|\nabla_i v_1 \cdot \nabla_j v_1 - \nabla_i v_2 \cdot \nabla_j v_2\|_0 \\
& = \|\nabla_i(v_1 - v_2) \cdot \nabla_j v_1 + \nabla_i v_2 \cdot \nabla_j(v_1 - v_2)\|_0 \\
& \leq K \|v_1 - v_2\|_2 (\|v_1\|_2 + \|v_2\|_2)
\end{aligned}$$

$$\begin{aligned}
& \|\nabla_i \nabla_j v_1 \cdot \nabla_k \nabla_l v_1 - \nabla_i \nabla_j v_2 \cdot \nabla_k \nabla_l v_2\|_0 \\
& = \|\nabla_i \nabla_j(v_1 - v_2) \cdot \nabla_k \nabla_l v_1 + \nabla_i \nabla_j v_2 \cdot \nabla_k \nabla_l(v_1 - v_2)\|_0 \\
& \leq K \|v_1 - v_2\|_2 (\|v_1\|_2 + \|v_2\|_2)
\end{aligned}$$

By (27) and (14), we see that

$$\begin{aligned} \|\nabla_p GN(v_1) - \nabla_p GN(v_2)\|_0 &\leq K \|GN(v_1) - GN(v_2)\|_1 \\ &\leq C_0 \|N(v_1) - N(v_2)\|_0 \\ &\leq K \|v_1 - v_2\|_2 (\|v_1\|_2 + \|v_2\|_2) \end{aligned}$$

Again, these are all of the types of terms that appear in the definition of  $M(v)$ , so we get

$$\|M(v_1) - M(v_2)\|_0 \leq K \|v_1 - v_2\|_2 (\|v_1\|_2 + \|v_2\|_2)$$

□

4.4. *Proving the First Perturbation Result.* We have found all necessary estimations and are now ready to prove the first perturbation result.

*Proof of Theorem 4.1.* Recall that by Lemma 4.10, this theorem may be proved by solving the following fixed point problem:

$$v = E(u) \left( GN(v), -\frac{1}{2}f + \frac{1}{2}GM(v) \right)$$

We use a method of iteration. Define the following sequence:

$$v_0 = 0 \quad v_{k+1} = E(u) \left( GN(v_k), -\frac{1}{2}f + \frac{1}{2}GM(v_k) \right)$$

It suffices to prove that there exists a  $\theta$  so that the sequence  $v_k$  converges in  $C^{s,\lambda}$  for  $s \geq 2$ .

First, assume  $s = 2$ .  $E(u)$  functions as a pseudo-inverse, so  $E(u)$  is  $\mathbb{R}$ -linear in  $C^{s,\lambda}(M, T^*M) \oplus C^{s,\lambda}(M, \text{Sym}^2(T^*M))$ . As a result, we have the following triangle inequality:

$$(28) \quad \|E(u)(h^1 + h^2, f^1 + f^2)\|_2 \leq \|E(u)(h^1, f^1)\|_2 + \|E(u)(h^2, f^2)\|_2$$

Using (12), (28), Lemma 4.11, (14), and (8), one can obtain:

$$\begin{aligned}
& \|v_{k+1}\|_2 \\
&= \left\| E(u) \left( GN(v_k), -\frac{1}{2}f + \frac{1}{2}GM(v_k) \right) \right\|_2 \\
&\leq \left\| E(u) \left( GN(v_k), \frac{1}{2}GM(v_k) \right) \right\|_2 + \frac{1}{2} \|E(u)(0, f)\|_2 \\
&\leq KD(u) \left( \|GN(v_k)\|_2 + \frac{1}{2} \|GM(v_k)\|_2 \right) + \frac{1}{2} \|E(u)(0, f)\|_2 \\
&\leq KD(u) (C_0 \|N(v_k)\|_0 + C_0 \|M(v_k)\|_0) + \frac{1}{2} \|E(u)(0, f)\|_2 \\
&\leq KD(u) (2K \|v_k\|_2^2 + 2C'_0 \|v_k\|_1^2) + \frac{1}{2} \|E(u)(0, f)\|_2 \\
&\leq K_1 D(u) \|v_k\|_2^2 + \frac{1}{2} \|E(u)(0, f)\|_2 \\
(29) \quad & \implies \|v_{k+1}\|_2 \leq K_1 D(u) \|v_k\|_2^2 + \frac{1}{2} \|E(u)(0, f)\|_2
\end{aligned}$$

$\|E(u)(0, f)\|_{s'}$  is determined by  $f$ , which is independent of the variation, so  $\|E(u)(0, f)\|_{s'}$  is finite. Now, we prove by induction that if

$$2K_1 D(u) \|E(u)(0, f)\|_2 \leq 1 \implies \|v_k\|_2 \leq \|E(u)(0, f)\|_2$$

The base case  $v_0 = 0$  is obvious; continuing inductively, we have by (29):

$$\begin{aligned}
& \|v_{k+1}\|_2 \\
&\leq K_1 D(u) \|v_k\|_2^2 + \frac{1}{2} \|E(u)(0, f)\|_2 \\
&\leq K_1 D(u) \|E(u)(0, f)\|_2^2 + \frac{1}{2} \|E(u)(0, f)\|_2 \\
&\leq K_1 D(u) \cdot \frac{1}{2K_1 D(u)} \|E(u)(0, f)\|_2 + \frac{1}{2} \|E(u)(0, f)\|_2 \\
&= \|E(u)(0, f)\|_2
\end{aligned}$$

So for all  $k$ , we have

$$(30) \quad \|v_k\|_2 \leq \|E(u)(0, f)\|_2$$

Now, one can examine

$$\|v_{k+1} - v_k\|_2 = \left\| E(u) \left( GN(v_k), -\frac{1}{2}f + \frac{1}{2}GM(v_k) \right) - E(u) \left( GN(v_{k-1}), -\frac{1}{2}f + \frac{1}{2}GM(v_{k-1}) \right) \right\|_2$$

Let  $x(v_k) = GN(v_k)$  and  $y(v_k) = -\frac{1}{2}f + \frac{1}{2}GM(v_k)$ . We use the trick of adding and subtracting  $E(u)(x(v_{k-1}), y(v_k))$ :

$$\begin{aligned} & \|v_{k+1} - v_k\|_2 \\ & \leq \|E(u)(x(v_k), y(v_k)) - E(u)(x(v_{k-1}), y(v_k))\|_2 + \|E(u)(x(v_{k-1}), y(v_k)) - E(u)(x(v_{k-1}), y(v_{k-1}))\|_2 \end{aligned}$$

Then, I claim that

$$(31) \quad \|E(u)(h^1, f) - E(u)(h^2, f)\|_2 \leq C_2 D(u) \|h^1 - h^2\|_2$$

$$(32) \quad \|E(u)(h, f^1) - E(u)(h, f^2)\|_2 \leq C_2 D(u) \|f^1 - f^2\|_2$$

To see (31) and (32), recall that any solution  $v = E(u)(h, f)$  can be written as

$$v_\alpha = A^{\alpha,i}(u)h_i + B^{\alpha,ij}(u)f_{ij}$$

Since  $A^{\alpha,i}$  and  $B^{\alpha,ij}$  depend only on  $u$ , by (9),

$$\begin{aligned} & \|E(u)(h^1, f) - E(u)(h^2, f)\|_2 \\ & = \|(A^\alpha(u)h^1 + B^\alpha(u)f) - (A^\alpha(u)h^2 + B^\alpha(u)f)\|_2 \\ & \leq C_2 \|A^\alpha(u)\|_2 \|h^1 - h^2\|_2 \\ & \leq C_2 D(u) \|h^1 - h^2\|_2 \\ & \|E(u)(h, f^1) - E(u)(h, f^2)\|_2 \\ & = \|(A^\alpha(u)h + B^\alpha(u)f^1) - (A^\alpha(u)h + B^\alpha(u)f^2)\|_2 \\ & \leq C_2 \|B^\alpha(u)\|_2 \|f^1 - f^2\|_2 \\ & \leq C_2 D(u) \|f^1 - f^2\|_2 \end{aligned}$$

Now, it follows from (31), (32), (14), and Lemma 4.11 that

$$\begin{aligned} & \|v_{k+1} - v_k\|_2 \\ & \leq \|E(u)(x(v_k), y(v_k)) - E(u)(x(v_{k-1}), y(v_k))\|_2 + \|E(u)(x(v_{k-1}), y(v_k)) - E(u)(x(v_{k-1}), y(v_{k-1}))\|_2 \\ & \leq C_2 D(u) \|x(v_k) - x(v_{k-1})\|_2 + C_2 D(u) \|y(v_k) - y(v_{k-1})\|_2 \\ & = C_2 D(u) \|G(N(v_k) - N(v_{k-1}))\|_2 + \frac{1}{2} C_2 D(u) \|G(M(v_k) - M(v_{k-1}))\|_2 \\ & \leq C_2 C_0^1 D(u) \|N(v_k) - N(v_{k-1})\|_0 + C_2 C_0^2 D(u) \|G(M(v_k) - M(v_{k-1}))\|_0 \\ & \leq K_2 D(u) (\|v_k\|_2 + \|v_{k-1}\|_2) (\|v_k - v_{k-1}\|_2) \end{aligned}$$

Then, by (30), we can replace  $\|v_k\|_2 + \|v_{k-1}\|_2$  with  $2\|E(u)(0, f)\|_2$  to get:

$$\|v_{k+1} - v_k\|_2 \leq 2K_2 D(u) \|E(u)(0, f)\|_2 \|v_k - v_{k-1}\|_2$$

Then if  $4K_2 D(u) \|E(u)(0, f)\|_2 \leq 1$ , then the sequence  $\{v_k\}$  is Cauchy, and by completeness of the Hölder space  $C^{2,\lambda}(M, \mathbb{R}^N)$ ,  $\{v_k\}$  converges in  $C^{2,\lambda}$  to some  $v = E(u)(GN(v_k), -\frac{1}{2}f + \frac{1}{2}GM(v))$ .

To show that  $v \in C^{s,\lambda}(M, \mathbb{R}^N)$ , by Arzela–Ascoli, it suffices to show that the sequence  $\{v_k\}_{k \geq 0}$  is bounded in  $C^{s,\lambda}(M, \mathbb{R}^N)$ , since we already have that  $v_k \xrightarrow{C^{2,\lambda}} v$ .

This may be demonstrated inductively. The base case  $C^{2,\lambda}$  is already done. Now, suppose the sequence is bounded in  $C^{k,\lambda}$  for  $k \leq s' - 1$ . Using (28), Lemma 4.4, Lemma 4.8, and (8), we get

$$\begin{aligned}
\|v_{k+1}\|_{s'} &= \left\| E(u) \left( GN(v_k), -\frac{1}{2}f + \frac{1}{2}GM(v_k) \right) \right\|_{s'} \\
&\leq \left\| E(u) \left( GN(v_k), \frac{1}{2}GM(v_k) \right) \right\|_{s'} + \frac{1}{2} \|E(u)(0, f)\|_{s'} \\
&\leq K'_3 D(u) \left( \|GN(v_k)\|_{s'} + \frac{1}{2} \|GM(v_k)\|_{s'} \right) + C'(u, s) \left( \|GN(v_k)\|_{s'-1} + \frac{1}{2} \|GM(v_k)\|_{s'-1} \right) + \frac{1}{2} \|E(0, f)\|_{s'} \\
&\leq K''_3 D(u) \left( K \|N(v_k)\|_{s'-2} + C_s \|N(v_k)\|_{s'-3} + \frac{1}{2} K \|M(v_k)\|_{s'-2} + \frac{1}{2} C_s \|M(v_k)\|_{s'-3} \right) \\
&\quad + C'(u, s) \left( K \|N(v_k)\|_{s'-3} + C_s \|N(v_k)\|_{s'-4} + \frac{1}{2} K \|M(v_k)\|_{s'-3} + \frac{1}{2} C_s \|M(v_k)\|_{s'-4} \right) + \frac{1}{2} \|E(0, f)\|_{s'} \\
&\leq K'_3 D(u) (K' \|v_k\|_{s'} \|v_k\|_2 + C_2 \|v_k\|_{s'-1}^2) + C'(u, s) (K' \|v_k\|_{s'-1} \|v\|_2 + C_2 \|v_k\|_{s'-2}^2) + \frac{1}{2} \|E(0, f)\|_{s'} \\
&\leq K_3 D(u) (\|E(u)(0, f)\|_2 \|v_k\|_{s'}) + C(u, s) (\|v_k\|_{s'-1}^2) + \frac{1}{2} \|E(u)(0, f)\|_{s'}
\end{aligned}$$

Now, if  $2K_3 D(u) \|E(u)(0, f)\|_2 \leq 1$ , then we get by repeated substitution as below:

$$\begin{aligned}
&\|v_{k+1}\|_{s'} \\
&\leq \frac{1}{2} \|v_k\|_{s'} + C(u, s) \|v_k\|_{s'-1}^2 + \frac{1}{2} \|E(u)(0, f)\|_{s'} \\
&\leq \frac{1}{2} \left( \frac{1}{2} \|v_{k-1}\|_{s'} + C(u, s) \|v_{k-1}\|_{s'-1}^2 + \frac{1}{2} \|E(u)(0, f)\|_{s'} \right) + C(u, s) \|v_k\|_{s'-1}^2 + \frac{1}{2} \|E(u)(0, f)\|_{s'} \\
&\quad \vdots \\
&\leq 2C(u, s) \sup_{l \geq 1} \|v_l\|_{s'-1}^2 + \|E(u)(0, f)\|_{s'}
\end{aligned}$$

By inductive hypothesis,  $\sup_{l \geq 1} \|v_l\|_{s'-1}^2$  exists, and  $\|E(u)(0, f)\|_{s'}$  is a constant. Therefore, the sequence  $\{v_k\}$  is also bounded in  $C^{s',\lambda}(M, \mathbb{R})$ , which means by induction,  $\{v_k\}$  is bounded in  $C^{s,\lambda}(M, \mathbb{R})$  for any  $s \geq 3$ , given that  $2K_3 D(u) \|E(u)(0, f)\|_2 \leq 1$ .

We have proven:

- (i) If  $2K_1D(u)\|E(u)(0, f)\|_2 \leq 1$  and  $4K_2D(u)\|E(u)(0, f)\|_2 \leq 1$ , the sequence  $\{v_k\}$  converges in  $C^{2,\lambda}(M, \mathbb{R}^N)$ .
- (ii) If  $2K_3D(u)\|E(u)(0, f)\|_2 \leq 1$ , the  $\{v_k\}$  is bounded  $C^{s,\lambda}$  for every  $s \geq 3$ .

Therefore, we set

$$\theta = \frac{1}{\max(2K_1, 4K_2, 2K_3)}$$

□

4.5. *How to Apply the Perturbation Result to the Original Isometric Embedding Problem.* During this time, we have been assuming that  $N$  is of “high enough dimension,” but it is now natural to ask how high  $N$  needs to be.

Using his trick involving “ $Y$  and  $Z$  embeddings,” Nash proved that the existence of a free embedding is guaranteed if  $N \geq \frac{n(n+5)}{2}$  [42, pp. 52-59]. Alternatively, this is proven via a spherical fibre bundle argument by Mikhail Gromov and Vladimir Rokhlin in 2.5.3 of [26].<sup>26</sup> There are  $\frac{n(n+3)}{2}$  first and second derivatives, but it turns out that assuming  $N \geq \frac{n(n+3)}{2}$  is not enough to guarantee a solution to Günther’s linearized problem. As indicated in the statement of Theorem 4.15, 5 extra dimensions are required.

Günther discovered that there is a variant of Theorem 4.1 applicable to the isometric embedding problem:

**THEOREM 4.12** ([27] Theorem 2). *Let  $B \subseteq \mathbb{R}^n$  be the open unit ball and  $B_1, B_2 \subseteq \mathbb{R}^n$  open sets with  $\bar{B}_1 \subseteq B_2$ ,  $\bar{B}_2 \subseteq B$ . Let  $u \in C^\infty(\bar{B}, \mathbb{R}^N)$  be a free mapping and  $f = (f_{ij}) \in C^{s,\lambda}(B, \mathbb{R}^{n(n+1)/2})$  with  $2 \leq s \leq \infty$  and  $\text{supp } f \subseteq B_1$ .*

*There exists some  $\theta > 0$  independent of  $u, s, f$  such that if*

$$D(u)\|E(u)(0, f)\|_2 \leq \theta$$

*then there exists a  $v \in C^{s,\lambda}(B, \mathbb{R}^N)$  with  $\text{supp } v \subseteq B_2$  and*

$$\partial_i u \cdot \partial_j v + \partial_j u \cdot \partial_i v + \partial_i v \cdot \partial_j v = f_{ij}$$

To prove Theorem 4.12, we need to set up a fixed point problem analogous to that in the proof of Theorem 4.1, which means that we need new versions of Günther’s lemmas from Section 4.3, all of which require our bounded operator  $G : C^{s,\lambda} \rightarrow C^{s+2,\lambda}$ , defined to be  $(\Delta - \mathbb{1})^{-1}$ . Fortunately, because we are in the setting of the unit ball in Euclidean space, to find such an operator is

---

<sup>26</sup>This argument of Gromov-Rokhlin is explored more in detail of section 2.1 of Ben Andrews’s notes [2].

easier; we can use simply the inverse of the Laplace operator  $\Delta$ , which as we saw in Section 4.2, is well-defined and bounded.

Because of our current setting in the unit ball of  $\mathbb{R}^n$ , derivatives commute, so automatically by Schauder's interior estimate (Theorem 4.7), for every tensor  $y$ , we have an analogue of Lemma 4.8:

$$\|\Delta^{-1}y\|_{s+2} \leq K\|y\|_s$$

Günther employs the following trick in [27]: Let  $\rho$  be a smooth bump function such that  $\text{supp } \rho = B_2$  and  $\rho(B_1) = \{1\}$ . The condition that  $v$  takes the form  $v = \rho^2 w$  for some bounded function  $w$  forces  $v$  to have compact support; recall that the manifolds for Theorem 4.1 are compact. Therefore, for the new theorem we are trying to uncover,  $v$  should have compact support as it did for Günther's lemmas and Theorem 4.1.

Analogously to (19), (20), (21), we define the vector field  $N$  and (0, 2) symmetric tensor field  $M$ :

$$N_j(w) := -\Delta^{-1}(\partial_j w \cdot \rho \Delta w) - (\partial_j \rho)(w \cdot w)$$

$$\begin{aligned} M_{ij}(w) &:= \frac{1}{2} \rho \Delta^{-1}((\Delta \rho)(\partial_i w \cdot \partial_j w) + 2\rho(\partial_{ki} w \cdot \partial_{kj} w)) + \rho \Delta^{-1}((\partial_k \rho)((\partial_{ki} w \cdot \partial_j w) + (\partial_i w \cdot \partial_{kj} w))) \\ &\quad - \frac{1}{2} \rho \Delta^{-1}((\partial_i \rho)(\Delta w \cdot \partial_j w) + (\partial_j \rho)(\partial_i w \cdot \Delta w) + 2\rho(\Delta w \cdot \partial_{ij} w)) \end{aligned}$$

The lemma below is a new version of Lemma 4.9:

LEMMA 4.13. *Let  $u \in C^\infty(\bar{B}, \mathbb{R}^N)$  and  $v \in C^2(B, \mathbb{R}^N)$  of the form  $v = \rho^2 w$  for some bounded function  $w$  and smooth bump function  $\rho$  with  $\text{supp } \rho = B_2$  and  $\rho(B_1) = \{1\}$ . For  $1 \leq i, j \leq n$ , on  $B_2$ , we have*

$$\begin{aligned} &\partial_i u \cdot \partial_j v + \partial_j u \cdot \partial_i v + \partial_i v \cdot \partial_j v \\ &= \rho^3 \partial_j \left( \partial_i u \cdot \frac{w}{\rho} - N_i(w) \right) + \rho^3 \partial_i \left( \partial_j u \cdot \frac{w}{\rho} - N_j(w) \right) - 2\rho^2 \partial_{ij} u \cdot w + 2\rho^2 M_{ij}(w) \end{aligned}$$

Similarly, relevant calculations are shown in Section 6.5 of [2]; we provide them here as well.

*Proof.* Once again, using the product rule, we have

$$\partial_j(\partial_i u \cdot v) + \partial_i(\partial_j u \cdot v) - 2\partial_{ij} u \cdot v = \partial_i u \cdot \partial_j v + \partial_j u \cdot \partial_i v$$

Then, we observe that

$$\begin{aligned}
& \partial_j(\partial_i u \cdot v) \\
&= \partial_{ji} u \cdot v + \partial_i u \cdot \partial_j v \\
&= \rho^3 \left( \partial_{ji} u \cdot \frac{w}{\rho} \right) + (3\rho \partial_j \rho) (\partial_i u \cdot w) + \rho^3 \left( \partial_i u \cdot \partial_j \left( \frac{w}{\rho} \right) \right) \\
&= \rho^3 \partial_j \left( \partial_i u \cdot \frac{w}{\rho} \right) + (3\rho \partial_j \rho) (\partial_i u \cdot w)
\end{aligned}$$

$$\begin{aligned}
& \partial_i v \cdot \partial_j v \\
&= (2\rho(\partial_i \rho)w + \rho^2 \partial_i w) \cdot (2\rho(\partial_j \rho)w + \rho^2 \partial_j w) \\
&= 4\rho^2(\partial_i \rho)(\partial_j \rho)(w \cdot w) + 2\rho^3(\partial_i \rho)(w \cdot \partial_j w) + 2\rho^3(\partial_j \rho)(\partial_i w \cdot w) + \rho^4(\partial_i w \cdot \partial_j w)
\end{aligned}$$

We notice that

$$\begin{aligned}
& \Delta(\rho \partial_i w \cdot \partial_j w) \\
&= \partial_k(\partial_k \rho(\partial_i w \cdot \partial_j w)) + \partial_k(\rho(\partial_k(\partial_i w \cdot \partial_j w))) \\
&= (\Delta \rho)(\partial_i w \cdot \partial_j w) + 2(\partial_k \rho)(\partial_k(\partial_i w \cdot \partial_j w)) + \rho(\Delta(\partial_i w \cdot \partial_j w)) \\
&= (\Delta \rho)(\partial_i w \cdot \partial_j w) + 2(\partial_k \rho)(\partial_{ki} w \cdot \partial_j w) + 2(\partial_k \rho)(\partial_i w \cdot \partial_{kj} w) + \rho(\partial_k(\partial_{ki} w \cdot \partial_j w + \partial_i w \cdot \partial_{kj} w)) \\
&= (\Delta \rho)(\partial_i w \cdot \partial_j w) + 2(\partial_k \rho)(\partial_{ki} w \cdot \partial_j w) + 2(\partial_k \rho)(\partial_i w \cdot \partial_{kj} w) + \rho(\partial_i \Delta w \cdot \partial_j w) \\
&\quad + 2\rho(\partial_{ki} w \cdot \partial_{kj} w) + \rho(\partial_i w \cdot \partial_j \Delta w) \\
&= (\Delta \rho)(\partial_i w \cdot \partial_j w) + 2(\partial_k \rho)(\partial_{ki} w \cdot \partial_j w) + 2(\partial_k \rho)(\partial_i w \cdot \partial_{kj} w) + \partial_i(\rho \Delta w \cdot \partial_j w) \\
&\quad + 2\rho(\partial_{ki} w \cdot \partial_{kj} w) + \partial_j(\partial_i w \cdot \rho \Delta w) - (\partial_i \rho)(\Delta w \cdot \partial_j w) - (\partial_j \rho)(\partial_i w \cdot \Delta w) - 2\rho(\Delta w \cdot \partial_{ij} w)
\end{aligned}$$

So if we apply  $\Delta^{-1}$  to the above expression and then multiply both sides by  $\rho^3$ , we have an alternative expression for  $\rho^4(\partial_i w \cdot \partial_j w)$ . We also observe the following product rule trick:

$$\frac{1}{2} \partial_i((\partial_j \rho)(w \cdot w)) - \frac{1}{2}(\partial_{ij} \rho)(w \cdot w) = (\partial_j \rho)(w \cdot \partial_i w)$$

So we get

$$\begin{aligned}
& \partial_i u \cdot \partial_j v + \partial_j u \cdot \partial_i v + \partial_i v \cdot \partial_j v \\
&= \rho^3 \partial_j \left( \partial_i u \cdot \frac{w}{\rho} \right) + (3\rho \partial_j \rho) (\partial_i u \cdot w) + \rho^3 \partial_i \left( \partial_j u \cdot \frac{w}{\rho} \right) + (3\rho \partial_i \rho) (\partial_j u \cdot w) - 2\rho^2 \partial_{ij} u \cdot w \\
&+ 4\rho^2 (\partial_i \rho) (\partial_j \rho) (w \cdot w) + 2\rho^3 (\partial_i \rho) (w \cdot \partial_j w) + 2\rho^3 (\partial_j \rho) (\partial_i w \cdot w) + \rho^3 \partial_i (\Delta^{-1} (\rho \Delta w \cdot \partial_j w)) \\
&+ \rho^3 \partial_j (\Delta^{-1} (\partial_i w \cdot \rho \Delta w)) + \rho^3 \Delta^{-1} ((\Delta \rho) (\partial_i w \cdot \partial_j w) + 2\rho (\partial_{ki} w \cdot \partial_{kj} w)) \\
&+ 2\rho^3 \Delta^{-1} ((\partial_k \rho) ((\partial_{ki} w \cdot \partial_j w) + (\partial_i w \cdot \partial_{kj} w))) - \rho^3 \Delta^{-1} ((\partial_i \rho) (\Delta w \cdot \partial_j w) + (\partial_j \rho) (\partial_i w \cdot \Delta w) + 2\rho (\Delta w \cdot \partial_{ij} w)) \\
&= \rho^3 \partial_j \left( \partial_i u \cdot \frac{w}{\rho} + \Delta^{-1} (\partial_i w \cdot \rho \Delta w) + (\partial_i \rho) (w \cdot w) \right) + \rho^3 \partial_i \left( \partial_j u \cdot \frac{w}{\rho} + \Delta^{-1} (\rho \Delta w \cdot \partial_j w) + (\partial_j \rho) (w \cdot w) \right) \\
&- 2\rho^2 \partial_{ij} u \cdot w + (3\rho \partial_j \rho) (\partial_i u \cdot w) + (3\rho \partial_i \rho) (\partial_j u \cdot w) + 4\rho^2 (\partial_i \rho) (\partial_j \rho) (w \cdot w) - 2\rho^3 (\partial_{ij} \rho) (w \cdot w) \\
&+ \rho^3 \Delta^{-1} ((\Delta \rho) (\partial_i w \cdot \partial_j w) + 2\rho (\partial_{ki} w \cdot \partial_{kj} w)) + 2\rho^3 \Delta^{-1} ((\partial_k \rho) ((\partial_{ki} w \cdot \partial_j w) + (\partial_i w \cdot \partial_{kj} w))) \\
&- \rho^3 \Delta^{-1} ((\partial_i \rho) (\Delta w \cdot \partial_j w) + (\partial_j \rho) (\partial_i w \cdot \Delta w) + 2\rho (\Delta w \cdot \partial_{ij} w)) \\
&= \rho^3 \partial_j \left( \partial_i u \cdot \frac{w}{\rho} - N_i(w) \right) + \rho^3 \partial_i \left( \partial_j u \cdot \frac{w}{\rho} - N_j(w) \right) - 2\rho^2 \partial_{ij} u \cdot w + 2\rho^2 M_{ij}(w)
\end{aligned}$$

□

Then, we can write the new version of Lemma 4.10:

LEMMA 4.14. For  $u, v, p, w$  defined as in Lemma 4.13, for  $f \in C^{2,\lambda}(B, \mathbb{R}^{n(n+2)/2})$ , and for  $1 \leq i, j \leq n$  with  $\text{supp } f \subseteq B_1$ , if

$$\partial_i u \cdot w = \rho N_i(w) \quad \partial_{ij} u \cdot w = M_{ij}(w) - \frac{1}{2} f_{ij}$$

Then

$$\partial_i u \cdot \partial_j v + \partial_j u \cdot \partial_i v + \partial_i v \cdot \partial_j v = f_{ij}$$

*Proof.* This follows from substituting as we did in the proof of Lemma 4.14 and recognizing that  $\rho(\text{supp } f) = \{1\}$ . □

The rest of the versions of Günther's lemmas and the resulting proof of Theorem 4.12 is essentially repetition from the previous section, so instead of proving everything (again), I include a brief summary of how to get the result.

For  $v \in C^{s+2,\lambda}(B, \mathbb{R}^n)$ , we can apply the boundedness of  $\Delta^{-1}$  by Schauder and properties (8), (9), (10), (24), and (25) of Hölder norms to get estimates on  $N$  and  $M$  similar to those of Lemma 4.11, since  $N$  and  $M$  are linear combinations of dot products of derivatives of  $w$  of order no more than 2 multiplied by derivatives of  $\rho$ . Multiplication by  $\rho$ 's derivatives changes nothing as every derivative of  $\rho$  bounded.

We will use the same pseudo-inverse  $E(u)$  from before<sup>27</sup> and define  $D(u)$  in the same way; (12) and Lemma 4.4 also continues to apply in the new case.

Finally, Theorem 4.12 can be proven by solving the fixed point problem via a similar iterative process:

$$w = E(u) \left( \rho N(w), -\frac{1}{2}f + M(w) \right)$$

Recall that what we are looking for is  $v = \rho^2 w$ .

As mentioned before, Theorem 4.12 can be applied to prove the following theorem:

**THEOREM 4.15** ([27] Theorem 3). *Let  $(M, g)$  be a smooth Riemannian manifold, not necessarily compact. Let  $g$  be a  $C^\infty$  metric,  $N \geq \frac{n(n+3)}{2} + 5$ ,  $u_0 \in C^\infty(M, \mathbb{R}^N)$  such that the norms  $\|\cdot\|_g$  and  $\|\cdot\|_{g_{u_0}}$  induced by  $g$  and  $g_{u_0}$  respectively satisfy  $\|\cdot\|_{g_{u_0}} < \|\cdot\|_g$ , and  $\delta \in C^0(M, \mathbb{R}^+)$ . Then, there exists an embedding  $u \in C^\infty(M, \mathbb{R}^N)$  such that  $g_u = g$  and  $\|u(p) - u_0(p)\| \leq \delta(p)$  for all  $p \in M$ .*

Recall that by Nash in [42] or Gromov–Rokhlin in 2.5.3 of [26], a free embedding is guaranteed to exist if  $N \geq \frac{n(n+5)}{2}$ . Such a free embedding can be modified so that the condition on the induced metric for Theorem 4.15 is satisfied. Therefore, an immediate result from 2.5.3 in [26] and Theorem 4.15 is that any smooth Riemannian manifold can be smoothly isometrically embedded into Euclidean space of dimension

$$\max \left( \frac{n(n+5)}{2}, \frac{n(n+3)}{n} + 5 \right)$$

with the Euclidean metric.

A full proof of Theorem 4.15 in English is provided in Section 4 of [38] by Siyuan Lu.<sup>28</sup> It uses the following technique developed by Nash, which Andrews in [2] refers to as the *Nash twist*.

In [42, pp. 54–56], Nash showed that one can construct a sequence of  $C^\infty$  functions  $f_1, \dots, f_{n(n+3)/2}$  on  $M$  such that the symmetric  $(0, 2)$ -tensors defined by

$$(33) \quad \partial_i f_k \cdot \partial_j f_k$$

<sup>27</sup>The definition of  $E(u)$  did not require  $M$  to be compact.

<sup>28</sup>The original proofs were given in German by Günther in [29]. In Section 4 of [38], Theorem 3.5 is Theorem 4.15.

are a generating set for the space of symmetric  $(0, 2)$ -tensors at every point on  $M$ . By the condition  $\|\cdot\|_{g_{u_0}} < \|\cdot\|_g$ ,  $g - g_{u_0}$  can be written as a  $C^\infty(M)$ -linear combination of tensors of the form (33) with positive  $C^\infty$  coefficients  $a_k$ .

If we define  $y_\varepsilon : M \rightarrow \mathbb{R}^{2m}$  (the “twist”) by

$$\begin{aligned} y_\varepsilon^k &:= \frac{a_k}{\varepsilon} \sin(\varepsilon f_k) \\ y_\varepsilon^{m+k} &:= \frac{a_k}{\varepsilon} \cos(\varepsilon f_k) \end{aligned}$$

The family of induced metrics

$$(34) \quad (g^\varepsilon)_{ij} := \sum_{k=1}^m \left( \partial_i y_\varepsilon^k \cdot \partial_j y_\varepsilon^k + \partial_i y_\varepsilon^{m+k} \cdot \partial_j y_\varepsilon^{m+k} \right) = \frac{1}{\varepsilon^2} \sum_{k=1}^m \partial_i a_k \cdot \partial_j a_k$$

can approximate  $g - g_{u_0}$ .<sup>29</sup>

Furthermore, one can choose local coordinates so that in such a neighborhood  $U_1$ , each term in the decomposition of  $g - g_{u_0}$  has the form  $a^4(dx_1 \otimes dx_1)$  (in other words,  $f_k = x_1$ ) for some  $a \in C^\infty(U_1)$  with compact support (so  $\text{supp } h \subseteq U_1$ ). So the goal becomes to show that  $u_0$  can be perturbed to a free embedding  $\tilde{u}_0$  such that

$$\partial_i \tilde{u}_0 \cdot \partial_j \tilde{u}_0 = \partial_i u_0 \cdot \partial_j u_0 + a^4 \delta_{i1} \delta_{j1}$$

The coordinate neighborhood  $U^{(1)}$  can be identified with the unit ball  $B \in \mathbb{R}^n$  so that Theorem 4.12 can be applied to obtain the right perturbation. Theorem 3.13 in [38] is an alteration of Theorem 4.12 specifically aimed at finding this perturbation.

## 5. Consequences and Other Problems

In this section, I briefly introduce one major consequence of the John Nash’s work in isometric embeddings: the Nash–Moser theorem. Then, I include some examples of other problems to which isometric embeddings apply.

5.1. *The Nash–Moser Theorem.* As mentioned previously, Nash’s proof of his isometric embedding theorem, which involved constructing a non-standard inverse function theorem, turned out to be a very powerful result in analysis. There exists a generalization of Nash’s proof into a new inverse / implicit function theorem on *Fréchet spaces*, which are complete Hausdorff metrizable locally convex topological vector spaces (For a further introduction to this

---

<sup>29</sup>The derivation of 34 is not difficult; it is included in Section 3.1 of Andrews’s notes [2].

concept, see [30]). The first to produce an abstraction of Nash's results to Fréchet was Jacob Schwartz in [50] (published in 1960), but it was Jürgen Moser who first showed in [40] and [41] (published in 1966) that Nash's methods could be applied to other problems—in Moser's case, the nonlinear theory of positive symmetric systems in celestial mechanics. The result became known as the Nash-Moser theorem; below is Richard Hamilton's formulation in [30] (published in 1982):

**THEOREM 5.1 (Nash–Moser).** *Let  $\mathcal{F}$  and  $\mathcal{G}$  be tame Fréchet spaces and  $P : \mathcal{F} \rightarrow \mathcal{G}$  a smooth tame map. Suppose  $U \subseteq \mathcal{F}$  is an open set and, for all  $f \in U$ ,  $k \in \mathcal{G}$ , the equations for the derivative  $DP(f)h = k$  each have a unique solution  $h = VP(f)k$  so that the family of inverses  $VP : U \times \mathcal{G} \rightarrow \mathcal{F}$  is a smooth tame map. Then  $P$  is locally invertible, and each local inverse  $P^{-1}$  is a smooth tame map.*

The tameness condition on the spaces and mappings has to do with gradings on Fréchet spaces. [30] includes the full explanation of how calculus is done on Fréchet spaces as well as the proof of the Nash–Moser theorem (in particular, see Part II). Later on, Hamilton used the Nash–Moser result to prove the existence of Ricci flow in the 1980s, which eventually led to Perelman's proof of the Thurston geometrization conjecture (and therefore the Poincaré conjecture) in the early 2000s.

In the cases of both Riemannian isometric embeddings and Ricci flow however, mathematicians later discovered more elegant methods which avoided the monstrosities involved in proving the Nash–Moser result. As we saw in Section 4, in 1987, Günther did so and was able to improve the dimension of the ambient Euclidean space to  $\max\left(\frac{n(n+3)}{2} + 5, \frac{n(n+5)}{2}\right)$ .

**5.2. Isometric Embeddings of Surfaces into Euclidean Space.** Some interesting results have been developed specifically in the case of surfaces, which we explored a bit in Section 3.1. Mikhail Gromov was able to demonstrate that any compact surface can be isometrically embedded into  $\mathbb{R}^5$  (see 3.2.4(D) in [25]).<sup>30</sup> Furthermore, any compact *orientable* surface can be isometrically embedded in  $\mathbb{R}^4$  (see 3.2.4(C), in particular the exercise, in [25]). In Example 3.8, we discussed Greene's metric on  $\mathbb{S}^2$  for which there is no  $C^2$  isometric embedding into  $\mathbb{R}^3$ . Louis Nirenberg solved Weyl's problem in [43], however, showing that a 2-sphere with any  $C^4$  metric inducing everywhere positive Gauss curvature can be isometrically embedded into  $\mathbb{R}^3$ . Alexei Pogorelov independently

---

<sup>30</sup>In comparison, Günther's proof allows surfaces to be isometrically embedded into  $\mathbb{R}^{10}$ .

proved something more general in [48]: A 2-sphere can be isometrically embedded in a 3-manifold whose curvature at every point is less than the lower bound for the Gaussian curvature induced by the metric on the 2-sphere at every point. For example, this means that any 2-sphere with a metric inducing Gaussian curvature greater than  $-1$  can be isometrically embedded into hyperbolic 3-space.

5.3. *Isometric Embeddings in General Relativity.* In general relativity, we are concerned with studying solutions  $g$  to the Einstein equation

$$(35) \quad \text{Ric}g - \frac{1}{2}Sg = 8\pi T$$

where the unknown  $g$ —the gravitational field—is a Lorentzian metric,<sup>31</sup>  $S$  is the scalar curvature of  $g$ , and  $T$  is the stress-energy tensor, which encodes some type of “momentum density” information responsible for generating the gravitational field.<sup>32</sup> Of great importance in general relativity is the equivalence principle: Accelerating inertial reference frames—such as an elevator accelerating upward at 9.81...m/s—are indistinguishable from reference frames in gravitational fields—for example, standing still on Earth. The upshot is that gravity is not a force but an *effect* caused by the curvature of spacetime; the gravitational field is a *global* structure defined on the space-time manifold, and a mass moving around within a gravitational reference frame is not actually responsible for the curvature of space time. This phenomenon leads to ambiguity as to how to *locally* define mass density—that is, the problem of *quasi-local mass*. One may be able to find such information by doing an energy integral of  $T$  over some “space-like” 3-submanifold  $\Omega$  describing a physical system. By conservation laws, expressed as  $\nabla^\mu T_{\mu\nu} = 0$ , this information depends only on the boundary  $\partial\Omega$ , a surface whose metric is induced by that of  $\Omega$ . Therefore, we can look at problems of quasi-local mass from the point of view of isometric embeddings of surfaces.<sup>33</sup> For this problem, it is useful to consider surfaces that admit isometric embeddings into Minkowski space ( $\mathbb{R}^{1,3} = \mathbb{R}^4$  with the metric  $-dt^2 + dx^2 + dy^2 + dz^2$ ). Weyl’s embedding theorem is useful here: See the work of Chiu-Chu Liu, Mu-Tao Wang, and Shing-Tung Yau in [36] and

---

<sup>31</sup>A Lorentzian metric  $g$  is similar to a Riemannian metric, but with signature  $(-, +, +, +)$ ; in other words, at every point  $p$ , the matrix  $g_{ij}(p)$  has one negative eigenvalue and three positive eigenvalues.

<sup>32</sup>One can think of this as a generalization of mass density in Newtonian mechanics and magnetic flux; the components  $T^{\mu\nu}$  measures *momentum flux* (think momentum “flow”) in the  $\mu$  direction through a 3-submanifold describing a physical system whose normal is in the  $\nu$  direction. The idea of a stress-energy tensor extends to other types of physical systems beyond the context of gravitation.

<sup>33</sup>For a more thorough explanation of the problem of quasi-local mass, see [59].

[60] to observe how this gets applied.

There is a specific case of the Einstein equation called the *Einstein vacuum equation*:

$$(36) \quad \text{Ric}g = 0$$

Solutions to (36) are called *gravitational instantons*. Maciej Dunajski and Paul Tod explored conformal and isometric embeddings of anti-self dual instantons<sup>34</sup> into Euclidean space in [15].

5.4. *Isometric Embeddings and Nonlinear Wave Systems.* The Einstein vacuum equation (36) is a type of *nonlinear wave system*. The use of isometric embedding results extend to other nonlinear wave systems as well. For example, Terence Tao has used Günther’s solution to obtain results for solutions  $u$  to defocusing nonlinear wave systems of the form

$$(37) \quad -\partial_{tt}u + \sum_{i=1}^d \partial_{ii}u = \text{grad}_{\mathbb{R}^m}F(u)$$

where  $\mathbb{R}^{d+1}$  has the Minkowski metric and  $\mathbb{R}^m$  has the Euclidean metric. In particular, Tao in [53] used Günther’s solution to the isometric embedding problem to find that there exists a smooth potential  $F$  where Equation (37) admits solutions with finite-time singularities if the Euclidean dimension  $m$  is at least 76. He did this by deriving an isometric-embedding-type condition for  $u$  from the constraints of the system, where the stress energy tensor is the target “metric”, and accordingly adapting the Nash embedding theorem.<sup>35</sup> In the zero-dimensional case, which is the well equation

$$(38) \quad -\partial_{tt}u = \text{grad}_{\mathbb{R}^m}F(u)$$

Tao in [54] also used Günther’s result to find solutions  $u$  that are global flows on some compact manifold. The idea is to find solutions to

$$\partial_t u = X(u)$$

the ODE for the global flow of a nonvanishing vector field  $X$  on a Riemannian manifold that push forward along an isometric embedding to solutions for (38).

---

<sup>34</sup>These are instantons  $g$  for which  $\star g = -g$ , where  $\star$  is the Hodge operator, defined in Section 5.1 of [58].

<sup>35</sup>See Section 3 of [53] for more details.

5.5. *Isometric Embeddings and Calibrated Geometry.* Isometric embedding theorems also have an interesting relationship with calibrated geometry, the study of submanifolds of minimal volume, where the volume is restricted by a particular closed form (or *calibration*). One can think of this as an extension of minimal surface problems to higher dimensions. Calibrated submanifolds of interest include special Lagrangian submanifolds of Calabi–Yau (Ricci-flat Kähler) manifolds, associative and co-associative submanifolds of  $G_2$ -manifolds, and Cayley submanifolds of Spin(7)-manifolds (for a further introduction to these concepts, see [37]). Robert Bryant in [9] and Colleen Robles and Sema Selur in [49] proved the following:

- (i) The interior of any closed, oriented, smooth 3-manifold can be isometrically embedded as an associative submanifold of a  $G_2$ -manifold. Furthermore, any closed, oriented, *analytic* 3-manifold can be isometrically embedded as a special Lagrangian submanifold of a Calabi–Yau 3-manifold.
- (ii) The interior of any closed, oriented, smooth 4-manifold with trivial bundle of self-dual 2-forms can be isometrically embedded as a Cayley submanifold of a Spin(7)-manifold. Furthermore, any closed, oriented, *analytic* 4 manifold with trivial bundle of self-dual 2-forms can be embedded as a co-associative submanifold of a  $G_2$ -manifold.

Although  $\mathbb{R}^7$  admits a  $G_2$ -structure, and  $\mathbb{R}^8$  admits Calabi-Yau and Spin(7)-structures, the above does not imply that the generic 3- or 4-manifold with those specified conditions above can be isometrically embedded into  $\mathbb{R}^7$  or  $\mathbb{R}^8$ .

5.6. *A Brief Note about Surface Matching.* Isometric embedding problems also have a place in surface matching problems in computer vision. In particular, the language of isometries and isometric embeddings applies to methods in comparing surfaces in the context of multidimensional scaling. See the work of Ron Kimmel and Asi Elad in [16] and Alexander Bronstein, Michael Bronstein, and Kimmel in [8] for more details.

## 6. Reverberations

We have spent this exposition discussing the *global* isometric embedding problem. Global embedding problems are present in just about every subfield of differential geometry and has also featured interactions with topology and algebraic geometry. In this exposition, we discussed global embeddings of Riemannian isometric embeddings that respect the Riemannian metric. In general, the goal of global embedding problems is to study embeddings that respect some special structure; in this section, we include (without proof) several results for the interested reader from symplectic geometry, contact geometry,

complex geometry, and CR-geometry. A few of these results can be observed from the point of view of isometric embeddings.

6.1. *Holomorphic Embeddings.* Complex manifolds are those that locally resemble  $\mathbb{C}^N$  and have holomorphic transition functions between coordinate charts. This forces the manifold to be orientable. A *holomorphic embedding* requires that the restriction of the coordinates of the ambient space to those of the manifold be holomorphic. Furthermore, a hermitian metric is defined in the same way as a Riemannian metric tensor, except that the inner product at each point is *hermitian* and is of the form  $g = g_{j\bar{k}} dz_j d\bar{z}_k$ . To each hermitian metric, we can associate the differential form  $\omega = g_{j\bar{k}} dz_j \wedge d\bar{z}_k$ . If  $d\omega = 0$ , then the manifold is called *Kähler*,  $g$  is referred to as the *Kähler metric*, and  $\omega$  is referred to as the *Kähler form*.

Kunihiko Kodaira's embedding theorem states that one can embed a compact Kähler manifold into some  $\mathbb{C}\mathbb{P}^n$  only if there exists a positive line bundle (see [4] and [5] for a proof by Donaldson). These can be realized as holomorphic isometric embeddings of the Kähler metric on the complex manifold into the Fubini–Study metric on  $\mathbb{C}\mathbb{P}^n$ .<sup>36</sup>

Complex geometry and algebraic geometry have a great amount of intersection, and one often moves between the language of the two when studying complex varieties—that is, zero sets of homogeneous polynomials in complex variables. From the theory of algebraic curves, it is known that every compact Riemann surface (i.e. complex 1-manifold) can be algebraically—and therefore holomorphically, since an algebraic embedding is defined by polynomials—embedded in  $\mathbb{C}\mathbb{P}^3$  because of 1-1 correspondence between smooth projective algebraic curves and compact Riemann surfaces (see for example Corollary IV.3.6 and Theorem B.3.1 in [31]). By the theory of Weil divisors—in particular, the equivalence of Weil divisors and line bundles on algebraic curves—all Riemann surfaces meet the conditions for Kodaira's embedding theorem. By Chow's theorem, from which it follows that every complex submanifold of  $\mathbb{C}\mathbb{P}^N$  is an algebraic variety ([11] Theorem V), the Kähler manifolds that can be realized as *algebraic subvarieties* of  $\mathbb{P}^n$ —that is, the zero set of some collection of homogeneous multivariate polynomials in  $n + 1$  variables—are precisely the

---

<sup>36</sup> $\mathbb{S}^{2n+1}$  is not a complex manifold, but one can express the round metric in terms of complex coordinates by taking the pullback along inclusion of  $dz_k d\bar{z}_k$  on  $\mathbb{C}^{n+1}$ . This metric induces another metric on the quotient  $\mathbb{S}^{2n+1}/U(1) \cong \mathbb{C}\mathbb{P}^n$ . The induced metric on  $\mathbb{C}\mathbb{P}^n$  turns out to be Kähler and is known as the *Fubini–Study* metric.

ones that admit positive line bundles.

One cannot holomorphically embed a compact manifold into affine complex space  $\mathbb{C}^n$  because such an embedding would have to simultaneously achieve maximum modulus and be open. However, Heinrich Behnke and Karl Stein proved that noncompact Riemann surfaces can be embedded holomorphically as submanifolds of affine complex space, or Stein manifolds (see Corollary 26.8 and Theorem 28.6 in [22]).

Isometric embedding problems are also studied in the context of Teichmüller theory. The *Teichmüller Space* of a Riemann surface  $S$  is the moduli space of complex structures.<sup>37</sup> Simultaneously, Teichmüller spaces can be viewed as the moduli space of hyperbolic metrics. This is related to the fact that most Riemann surfaces admit hyperbolic structures (i.e. metrics with constant sectional curvature  $-1$ ) by the uniformization theorem. Teichmüller spaces themselves have the structure of a complex manifold, and studying their geometry is important in understanding the deformation theory on a Riemann surface. Some recent work has been done concerning isometric embeddings of Teichmüller spaces—see the work of Frederik Benirschke and Carlos A. Serván in [6], in which they were able to classify all isometric embeddings of Teichmüller spaces as branched coverings for complex dimension at least 2.

6.2. *Symplectic and Contact Embeddings.* Symplectic manifolds are even-dimensional manifolds with a non-degenerate closed 2-form; such manifolds are studied from the point of view of *area* as opposed to distance and curvature in Riemannian geometry. Taking powers of the symplectic form produces a nonvanishing volume form, so symplecticity also forces orientability. Contact manifolds are manifolds of dimension  $2k + 1$  that admit a smooth maximally non-integrable  $2k$ -distribution, or equivalently there exists a one form  $\alpha$  such that  $\alpha \wedge (d\alpha)^n$  is nowhere vanishing.<sup>38</sup> Contact geometry is often viewed as the odd-dimensional analogue to symplectic geometry, though the translation is not always clean, as we will see in parts of this section.<sup>39</sup> [25] has chapters

---

<sup>37</sup>The existence of (holomorphic) complex structures is equivalent to, among other things, the existence of an automorphism  $J$  on the tangent bundle that satisfies  $J^2 = -\text{Id}$  and on which the *Nijenhuis tensor* of  $J$ , given below, is identically 0.

$$N_J(X, Y) = [X, Y] + J([JX, Y] + [X, JY]) - [JX, JY]$$

<sup>38</sup> $(d\alpha)^n$  means  $d\alpha$  wedged  $n$  times.

<sup>39</sup>For instance, not all contact manifolds are orientable; see [12] for a discussion of nonorientable contact manifolds.

on global embeddings in symplectic and contact geometry. Under symplectic (resp. contact) embeddings, symplectic (resp. contact) form of the ambient space pulls back to the symplectic (resp. contact) form of the manifold.

In symplectic geometry, Gromov’s non-squeezing theorem showed that a ball cannot be squeezed via symplectic embedding into a cylinder unless the radius of the sphere is at most the radius of the cylinder, which showed the strength of symplecticity as a condition (see Theorem 12.1.1 in [39]).

In [18] and [17], Yakov Eliashberg, Sang Seon Kim, and Leonid Polterovich studied the analogue of the nonsqueeze problem in contact geometry. The *prequantization space*  $\mathbb{R}^{2n} \times \mathbb{S}^1$  has a contact structure, so one considers the problem of squeezing  $B \times \mathbb{S}^1$  into  $C \times \mathbb{S}^1$ , where  $B$  is a ball and  $C$  is a cylinder. The results do not translate in general. There is a non-squeeze result if the difference between the radii of the cylinder and the ball is large enough, but results for when the difference is small are still to be determined at the time of writing. Where the difference between contact-squeezing and symplectic-squeezing is more apparent is in the case of two balls: If the radii are both smaller than 1, the balls *can* be contact-squeezed into each other; this is not true in the symplectic case.

Mentioned previously was Kodaira’s embedding theorem of Kähler manifolds into complex projective space. Kähler manifolds are simultaneously complex and symplectic—in particular, the Kähler form is symplectic. For general symplectic manifolds (not necessarily complex), by Gromov and David Tischler, *every* symplectic manifold admits a symplectic isometric embedding into  $\mathbb{C}\mathbb{P}^N$ , where the metric on the symplectic manifold is defined based on the symplectic form,<sup>40</sup> and the metric on  $\mathbb{C}\mathbb{P}^N$  is the Fubini–Study metric (see Theorem B in [55] or Exercise (1) under Theorem 3.4.2(A) in [25]).

Closed symplectic manifolds cannot be symplectically embedded into Euclidean space: All possible symplectic forms on Euclidean space are exact, which means that the symplectic form on the compact manifold would have to be exact. But this also means that the volume form obtained by taking a wedge power of the symplectic form must also be exact; by Stokes’ theorem, the form integrates to 0, which contradicts the volume form being everywhere non-vanishing. On open (noncompact) symplectic manifolds, however, symplectic forms can be exact, and by results from Gromov concerning the  $h$ -principle,

---

<sup>40</sup>To see how this is done, see Proposition 5 in Section 1.3 of [32] and the preceding paragraph.

open manifolds with exact symplectic forms *always* admit embeddings into Euclidean space.<sup>41</sup>

Many closed *contact* manifolds can be contact-embedded into Euclidean space. Naohiko Kasuya in [34] and [33] proved that a closed co-oriented contact manifold  $M$  of dimension  $2m + 1$  can be contact-embedded into  $\mathbb{R}^n$  in any of the following cases:

- $n = 4m + 1$ ,  $m \geq 3$  is odd,  $H_1(M; \mathbb{Z}) = 0$ ,  $\mathbb{R}^{4m+1}$
- $n = 4m + 1$ ,  $m \geq 4$  is even, and  $M$  is 2-connected
- $n = 4m + 1$ ,  $m = 2$  and  $M$  is simply connected
- $n = 5$ ,  $m = 1$ , and  $M$  has trivial first Chern class.<sup>42</sup>

For the first three cases, the contact structure on the ambient Euclidean space is the standard one; for the last case, a contact structure on  $\mathbb{R}^5$  accepting a contact-embedding is proven to exist in [33]. Kasuya also proves in [33] that any codimension two closed contact submanifold of odd dimensional Euclidean space has trivial first Chern class. Therefore, a trivial first Chern class is a necessary and sufficient condition for a dimension 3 closed contact manifold to contact-embed into  $\mathbb{R}^5$ .

Simon Donaldson and Fabian Lehmann discovered a perturbation result pertaining to the problem of realizing closed 3-forms on 5-manifolds through embeddings into Calabi-Yau 3-folds in [14]. Integral to their approach was examining strongly pseudoconvex forms, which define contact structures. It is also explained that their problem is strongly related to that of CR-embeddings of CR-manifolds, introduced in the next section.

**6.3. CR-Embeddings.** A *Cauchy–Riemann (CR) structure* on a manifold  $M$  is a integrable<sup>43</sup> real subbundle of the complexified tangent bundle  $TM_{\mathbb{C}} = TM \otimes_{\mathbb{R}} \mathbb{C} L$  such that  $L \cap \bar{L} = 0$ . A manifold with a CR-structure is referred to as a *CR-manifold*. A CR structure is a weaker form of a complex structure:

---

<sup>41</sup>In [19], see the description of the  $h$ -principle on page 54 and Theorem 7.2.3. by Gromov. One can also find the relevant concepts and results in [25].  $h$  here stands for “homotopy”; The idea is that families that satisfy the  $h$ -principle satisfy some type of nice homotopy relation that allows one to, for example, approximate solutions to certain PDE’s. Theorem 3.5 by Nash–Kuiper can be seen as a specific instance of the  $h$ -principle, for example.

<sup>42</sup>Chern classes are defined in Section 25.8 of [56].

<sup>43</sup>That is, closed under the Lie bracket

By Newlander–Nirenberg, if the dimension of  $M$  is even and  $L$  is of codimension 0, then  $M$  is a complex manifold.<sup>44</sup> As with the other geometries, a *CR-embedding* must be such that the CR-structure on the ambient space pulls back to the CR-structure on the manifold.

Global CR-embeddings into  $\mathbb{C}^N$  are a very interesting subject. By the combined work of Louis Boutet de Monvel, Joseph Kohn, Hugo Rossi, Reese Harvey, and Blaine Lawson, we have that for compact strictly pseudoconvex CR-manifolds  $X$ , the following are equivalent:

- $X$  is globally CR-embeddable into  $\mathbb{C}^N$  for some  $N$ .
- $X$  is the boundary of a normal Stein space (a generalization of Stein manifolds that allows singularities).
- The image of the Kohn Laplacian operator  $\bar{\partial}_b$  over functions is closed.<sup>45</sup>

Stein spaces are even dimensional, so CR-manifolds admitting embeddings into  $\mathbb{C}^n$  must be real odd dimensional.

Along with the above, we have by de Monvel that any compact strictly pseudoconvex CR-manifolds of dimension 5 and higher CR-embeds into  $\mathbb{C}^n$ . See [20] for more details.

For dimension 3, the Rossi spheres are counterexamples (see [1] for instance). But Sagun Chanillo, Hung-Lin Chiu, and Paul Yang proved in [10] if we additionally assume in dimension 3 that the CR Paenitz operator is non-negative and the scalar Tanaka–Webster curvature is positive, we have global CR-embeddability.<sup>46</sup>

There is a CR analogue of Kähler manifolds called *Sasakian manifolds*. Take a Riemannian manifold  $(M, g)$  and its *Riemannian cone*  $M \times \mathbb{R}^+$  with metric  $\bar{g} = t^2g + dt \otimes dt$ .  $M$  is Sasakian if  $(M \times \mathbb{R}^+, \bar{g})$  is Kähler. Sasakian

---

<sup>44</sup>Mentioned in a previous footnote was yet another definition of a complex structure involving an automorphism  $J$  on the tangent bundle with  $J^2 = -1$ —called an *almost complex structure*—that is a full complex structure when the associated Nijenhuis tensor vanishes. These concepts are related in the following way:  $J$  yields a splitting of each complexified tangent space  $T_p M_{\mathbb{C}}$  into eigenspaces:  $T_p M_{\mathbb{C}} = T_p^{1,0} M \oplus T_p^{0,1} M$  where elements of  $T_p^{1,0} M_{\mathbb{C}}$  (resp.  $T_p^{0,1} M_{\mathbb{C}}$ ) are eigenvectors with eigenvalue  $i$  (resp.  $-i$ ) under  $J$ . The distribution  $T^{1,0} M = \bigcup_{p \in M} T_p^{1,0} M$  is a subbundle of  $TM_{\mathbb{C}}$ , and  $J$  is *integrable* (i.e. a *full* complex structure) when this distribution is integrable. See Chapter 2 of [58] for more details.

<sup>45</sup>For a further introduction to the Kohn Laplacian on CR-manifolds, see [45].

<sup>46</sup>See also [52] for a survey on the CR-Paenitz operator and its relation to global CR-embeddings.

manifolds are simultaneously contact and CR, where the CR-structure is induced by the contact structure. For a further explanation, see [7].

Liviu Ornea and Misha Verbitsky proved in [44] that any compact Sasakian manifold can be CR-embedded into another Sasakian manifold diffeomorphic to a sphere.

### References

- [1] Tawfik Abbas et al. *Spectrum of the Kohn Laplacian on the Rossi sphere*. 2017. arXiv: [1708.05624](https://arxiv.org/abs/1708.05624) [[math.CV](#)]. URL: <https://arxiv.org/abs/1708.05624>.
- [2] Ben Andrews. “Notes on the isometric embedding problem and the Nash-Moser implicit function theorem”. In: *Surveys in analysis and operator theory*. Vol. 40. Australian National University, Mathematical Sciences Institute, 2002, pp. 157–209.
- [3] Michael F Atiyah and Ian Grant Macdonald. *Introduction to commutative algebra*. CRC Press, 2018.
- [4] Denis Auroux and Kartik Venkatram. *Symplectic Geometry, Lecture 18*. Lecture in Geometry of Manifolds at Massachusetts Institute of Technology. Accessed: 13 February 2026. Spring 2007.
- [5] Denis Auroux and Kartik Venkatram. *Symplectic Geometry, Lecture 19*. Lecture in Geometry of Manifolds at Massachusetts Institute of Technology. Accessed: 13 February 2026. Spring 2007.
- [6] Frederik Benirschke and Carlos A. Serván. *Isometric embeddings of Teichmüller spaces are covering constructions*. 2023. arXiv: [2305.04153](https://arxiv.org/abs/2305.04153) [[math.GT](#)]. URL: <https://arxiv.org/abs/2305.04153>.
- [7] Charles P. Boyer and Krzysztof Galicki. *Sasakian Geometry, Holonomy, and Supersymmetry*. 2007. arXiv: [math/0703231](https://arxiv.org/abs/math/0703231) [[math.DG](#)]. URL: <https://arxiv.org/abs/math/0703231>.
- [8] Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel. “Generalized multidimensional scaling: A framework for isometry-invariant partial surface matching”. In: *Proceedings of the National Academy of Sciences* 103.5 (2006), pp. 1168–1172. DOI: [10.1073/pnas.0508601103](https://doi.org/10.1073/pnas.0508601103). eprint: <https://www.pnas.org/doi/pdf/10.1073/pnas.0508601103>. URL: <https://www.pnas.org/doi/abs/10.1073/pnas.0508601103>.
- [9] Robert L. Bryant. *Calibrated embeddings in the special Lagrangian and coassociative cases*. 2000. arXiv: [math/9912246](https://arxiv.org/abs/math/9912246) [[math.DG](#)]. URL: <https://arxiv.org/abs/math/9912246>.

- [10] Sagun Chanillo, Hung-Lin Chiu, and Paul Yang. “Embeddability for 3-dimensional Cauchy–Riemann manifolds and CR Yamabe invariants”. In: *Duke Mathematical Journal* 161.15 (Dec. 2012). ISSN: 0012-7094. DOI: [10.1215/00127094-1902154](https://doi.org/10.1215/00127094-1902154). URL: <http://dx.doi.org/10.1215/00127094-1902154>.
- [11] Wei-Liang Chow. “On Compact Complex Analytic Varieties”. In: *American Journal of Mathematics* 71.4 (1949), pp. 893–914. ISSN: 00029327, 10806377. URL: <http://www.jstor.org/stable/2372375> (visited on 02/16/2026).
- [12] David Crombecque. “Nonorientable contact structures on 3-manifolds”. English. Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2023-03-02. PhD thesis. 2006, p. 49. ISBN: 978-1-109-95804-1. URL: <https://www.proquest.com/dissertations-theses/nonorientable-contact-structures-on-3-manifolds/docview/305267320/se-2>.
- [13] Manfredo P Do Carmo. *Differential geometry of curves and surfaces: revised and updated second edition*. Courier Dover Publications, 2016.
- [14] Simon Donaldson and Fabian Lehmann. *Closed 3-forms in five dimensions and embedding problems*. 2022. arXiv: [2210.16208](https://arxiv.org/abs/2210.16208) [math.DG]. URL: <https://arxiv.org/abs/2210.16208>.
- [15] Maciej Dunajski and Paul Tod. “Conformal and Isometric Embeddings of Gravitational Instantons”. In: *Geometry, Lie Theory and Applications*. Springer International Publishing, July 2021, pp. 21–48. ISBN: 9783030812966. DOI: [10.1007/978-3-030-81296-6\\_2](https://doi.org/10.1007/978-3-030-81296-6_2). URL: [http://dx.doi.org/10.1007/978-3-030-81296-6\\_2](http://dx.doi.org/10.1007/978-3-030-81296-6_2).
- [16] A. Elad and R. Kimmel. “On bending invariant signatures for surfaces”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25.10 (2003), pp. 1285–1295. DOI: [10.1109/TPAMI.2003.1233902](https://doi.org/10.1109/TPAMI.2003.1233902).
- [17] Yakov Eliashberg, Sang Seon Kim, and Leonid Polterovich. “Erratum to “Geometry of contact transformations and domains: Orderability versus squeezing””. In: *Geometry and Topology* 13.2 (Feb. 2009), pp. 1175–1176. DOI: [10.2140/gt.2009.13.1175](https://doi.org/10.2140/gt.2009.13.1175).
- [18] Yakov Eliashberg, Sang Seon Kim, and Leonid Polterovich. “Geometry of contact transformations and domains: Orderability versus squeezing”. In: *Geometry and Topology* 10.3 (Oct. 2006), pp. 1635–1747. DOI: [10.2140/gt.2006.10.1635](https://doi.org/10.2140/gt.2006.10.1635).
- [19] Yakov Eliashberg and Nikolai M Mishachev. *Introduction to the h-principle*. 48. American Mathematical Soc., 2002.
- [20] C. L. Epstein. “Bergman and Szegő Kernels, CR Analysis”. In: *Louis Boutet de Monvel, Selected Works*. Ed. by Victor W. Guillemin and Johannes Sjöstrand. Cham: Springer International Publishing, 2017, pp. 483–

583. ISBN: 978-3-319-27909-1. DOI: [10.1007/978-3-319-27909-1\\_7](https://doi.org/10.1007/978-3-319-27909-1_7). URL: [https://doi.org/10.1007/978-3-319-27909-1\\_7](https://doi.org/10.1007/978-3-319-27909-1_7).
- [21] Lawrence C Evans. *Partial differential equations*. Vol. 19. American mathematical society, 2022.
- [22] Otto Forster. *Lectures on Riemann surfaces*. Vol. 81. Springer Science & Business Media, 2012.
- [23] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order David Gilbarg; Neil S. Trudinger*. Springer, 2001.
- [24] Robert E Greene. “Metrics and isometric embeddings of the 2-sphere”. In: *Journal of Differential Geometry* 5.3-4 (1971), pp. 353–356.
- [25] Mikhail Gromov. *Partial differential relations*. Vol. 9. Springer Science & Business Media, 1986.
- [26] Mikhail Leonidovich Gromov and Vladimir Abramovich Rokhlin. “Embeddings and immersions in Riemannian geometry”. In: *Russian Mathematical Surveys* 25.5 (1970), p. 1.
- [27] Matthias Günther. “Isometric Embeddings of Riemannian Manifolds”. In: ed. by Ichiro Satake. Vol. 2. Proceedings of the International Congress of Mathematicians 1990. Springer-Verlag, 1991.
- [28] Matthias Günther. “On the perturbation problem associated to isometric embeddings of Riemannian manifolds”. In: *Annals of Global analysis and Geometry* 7 (1989), pp. 69–77.
- [29] Matthias Günther. “Zum Einbettungssatz von J. Nash”. In: *Mathematische Nachrichten* 144.1 (Jan. 1989), pp. 165–187. DOI: [10.1002/mana.19891440113](https://doi.org/10.1002/mana.19891440113).
- [30] Richard S Hamilton. “The inverse function theorem of Nash and Moser”. In: *Bulletin of the American Mathematical Society* 7.1 (1982), pp. 65–222.
- [31] Robin Hartshorne. *Algebraic geometry*. Vol. 52. Springer Science & Business Media, 2013.
- [32] Helmut Hofer and Eduard Zehnder. *Symplectic invariants and Hamiltonian Dynamics*. Birkhäuser Basel, 1994.
- [33] Naohiko Kasuya. “An obstruction for codimension two contact embeddings in the odd dimensional Euclidean spaces”. In: *Journal of the Mathematical Society of Japan* 68.2 (2016), pp. 737–743.
- [34] Naohiko Kasuya. “On contact embeddings of contact manifolds in the odd dimensional Euclidean spaces”. In: *International Journal of Mathematics* 26.07 (2015), p. 1550045. DOI: [10.1142/S0129167X15500457](https://doi.org/10.1142/S0129167X15500457).
- [35] A. Lichnerowicz. “Propagateurs, commutateurs et anticommutateurs en relativité générale”. In: (Oct. 2018). Ed. by C. DeWitt and B. DeWitt, pp. 823–864. DOI: [10.1007/s10714-018-2433-x](https://doi.org/10.1007/s10714-018-2433-x).

- [36] Chiu-Chu Melissa Liu and Shing-Tung Yau. *Positivity of quasi-local mass II*. 2005. arXiv: [math/0412292](https://arxiv.org/abs/math/0412292) [[math.DG](#)]. URL: <https://arxiv.org/abs/math/0412292>.
- [37] Jason Lotay. *Calibrated geometry and gauge theory*. Lecture notes, University of California at Berkeley. Accessed: February 14, 2026. Fall 2022. URL: <https://people.maths.ox.ac.uk/lotay/teaching.html#calib>.
- [38] Siyuan Lu. *Isometric Embedding of Riemannian Manifolds*. 2012.
- [39] Dusa McDuff and Dietmar Salamon. *Introduction to symplectic topology*. Oxford University Press, 2017.
- [40] Jürgen Moser. “A rapidly convergent iteration method and non-linear partial differential equations I”. In: *Annali della Scuola Normale Superiore di Pisa-Scienze Fisiche e Matematiche* 20.2 (1966), pp. 265–315.
- [41] Jürgen Moser. “A rapidly convergent iteration method and non-linear partial differential equations II”. In: *Annali della Scuola Normale Superiore di Pisa-Scienze Fisiche e Matematiche* 20.3 (1966), pp. 499–535.
- [42] John Nash. “The imbedding problem for Riemannian manifolds”. In: *Annals of mathematics* 63.1 (1956), pp. 20–63.
- [43] Louis Nirenberg. “The Weyl and Minkowski problems in differential geometry in the large”. In: *Communications on pure and applied mathematics* 6.3 (1953), pp. 337–394.
- [44] Liviu Ornea and Misha Verbitsky. *Embeddings of compact Sasakian manifolds*. 2007. arXiv: [math/0609617](https://arxiv.org/abs/math/0609617) [[math.DG](#)]. URL: <https://arxiv.org/abs/math/0609617>.
- [45] Marco M. Peloso. “An introduction to the analysis of the Kohn Laplacian on CR manifolds”. In: *Lecture note, Scuola Estiva di Matematica, Politecnico di Milano* (2004).
- [46] Chiakuei Peng and Zizhou Tang. “Chern-Simons invariant and conformal embedding of a 3-manifold”. In: *Acta Mathematica Sinica, English Series* 26.1 (Jan. 2010), pp. 25–28. DOI: [10.1007/s10114-010-8559-8](https://doi.org/10.1007/s10114-010-8559-8).
- [47] Peter Petersen. *Riemannian geometry*. Mar. 2019. URL: <https://link.springer.com/book/10.1007/978-3-319-26654-1>.
- [48] Alexei V. Pogorelov. “Some results on surface theory in the large”. In: *Advances in mathematics* 1.2 (1964), pp. 191–264.
- [49] Colleen Robles and Sema Salur. *Calibrated associative and Cayley embeddings*. 2009. arXiv: [0708.1286](https://arxiv.org/abs/0708.1286) [[math.DG](#)]. URL: <https://arxiv.org/abs/0708.1286>.
- [50] Jacob Schwartz. “On Nash’s implicit functional theorem”. In: *Communications on Pure and Applied Mathematics* 13.3 (1960), pp. 509–530.
- [51] Mei-Chi Shaw and Charles M. Stanton. “The Uniformization Theorem”. In: *Complex Analysis in One Variable and Riemann Surfaces*. Cham: Springer Nature Switzerland, 2025, pp. 403–430. ISBN: 978-3-031-93642-5.

- DOI: [10.1007/978-3-031-93642-5\\_11](https://doi.org/10.1007/978-3-031-93642-5_11). URL: [https://doi.org/10.1007/978-3-031-93642-5\\_11](https://doi.org/10.1007/978-3-031-93642-5_11).
- [52] Yuya Takeuchi. *CR Paneitz operator and embeddability*. 2025. arXiv: [2506.19215](https://arxiv.org/abs/2506.19215) [[math.CV](#)]. URL: <https://arxiv.org/abs/2506.19215>.
- [53] Terence Tao. “Finite-time blowup for a supercritical defocusing nonlinear wave system”. In: *Analysis & PDE* 9.8 (2016), pp. 1999–2030.
- [54] Terence Tao. *On the universality of potential well dynamics*. 2020. arXiv: [1707.02389](https://arxiv.org/abs/1707.02389) [[math.AP](#)]. URL: <https://arxiv.org/abs/1707.02389>.
- [55] David Tischler. “Closed 2-forms and an embedding theorem for symplectic manifolds”. In: *Journal of Differential Geometry* 12.2 (1977), pp. 229–235.
- [56] Loring W Tu. *Differential geometry: connections, curvature, and characteristic classes*. Springer, 2017.
- [57] Loring W. Tu. *An introduction to manifolds*. Springer, 2011.
- [58] Claire Voisin. *Hodge Theory and Complex Algebraic Geometry I: Volume 1*. Vol. 76. Cambridge University Press, 2002.
- [59] Mu-Tao Wang. *Quasilocal mass and surface Hamiltonian in spacetime*. 2012. arXiv: [1211.1407](https://arxiv.org/abs/1211.1407) [[gr-qc](#)]. URL: <https://arxiv.org/abs/1211.1407>.
- [60] Mu-Tao Wang and Shing-Tung Yau. “Isometric embeddings into the Minkowski space and new quasi-local mass”. In: *Communications in Mathematical Physics* 288.3 (2009), pp. 919–942.

SHIV YAJNIK

Department of Mathematics, Columbia University

*E-mail*: [columbiajournalofundergradmath@gmail.com](mailto:columbiajournalofundergradmath@gmail.com)

<https://journals.library.columbia.edu/index.php/cjum>