GEOMETRIC VARIATIONAL PROBLEMS IN MATHEMATICAL PHYSICS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF MATHEMATICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Christos Mantoulidis
June 2017

Abstract

We study two different geometrically flavored variational problems in mathematical physics: quasi-local mass in the initial data set approach to the general theory of relativity, and the theory of phase transitions. In the general relativity setting, we introduce a new moduli space of metrics on spheres and a new metric invariant on surfaces to help obtain a precise local understanding of the interaction of ambient scalar curvature and stable minimal surfaces in the context of three-manifolds with nonnegative scalar curvature; we use these tools to study the Bartnik and Brown-York notions quasi-local mass in general relativity. In the theory of phase transitions, we study the global behavior of two-dimensional solutions, and relate their complexity at infinity to their variational instability.

Acknowledgments

My journey in mathematics has been shaped by four people's distinctive influence. Rick Schoen, my principal Ph.D. thesis advisor, has been a truly invaluable mathematics-and-research mentor during graduate school, I am very proud to be joining his mathematical family upon the completion of my degree. Leon Simon, my undergraduate thesis advisor, has been a source of constant encouragement over the past ten years, and his teaching and approach to mathematics unquestionably drew me to study the subject in the first place. Eleni Mitsiou, my high school math coach, is the person that has single-handedly laid the entirety of my mathematical foundations. Christos Papadopoulos, my first math-physics-and-chemistry teacher, fueled me with enough support and enough enthusiasm for being a "panepistimon" during his lifetime to last me for all of my lifetime.

I would be remiss not to express great gratitude for the long-lasting influence that Brian White, Michael Eichmair, Pengzi Miao, Lenya Ryzhik, Rafe Mazzeo, and András Vasy have all had on me. I am also grateful for all my colleagues whose age is close enough to mine for me to have found the courage to ask all my basic math questions to: Alessandro Carlotto, Daren Cheng, Otis Chodosh, Nick Edelen, Or Hershkovits, Chao Li, Jesse Madnick, Davi Maximo, and Xin Zhou.

For their support and encouragement during graduate school, I would like to thank Zinovia Chatzidimitriadou, Dimitris Gerothanasis, Désirée Gréverath, Vasso Katsarou, Evita Nestoridi, Manolis Papadakis, Babis Papadopoulos, Niccolò Ronchetti, Mihalis Savvas, Evan Scouros, Jay Shah, Paris Syminelakis, and Sara Timtim. During the past three years, I have also been fortunate to have spent time at the University of California, Irvine, and I would like to thank Jess Boling, Jacque Gahari, Casey Kelleher, Alex Mramor, Shoo Seto, Heather Warta, and Amir Vig for making my visits to UCI so great. I am indebted to Gretchen Lantz, who made sure my degree got successfully conferred despite various odds, twice.

I am extremely grateful to: my aunt, a constant source of love and support despite our very infrequent communication; Anna, who has been a pillar of encouragement, love, and direction, as well as a source of amazing days worth celebrating; and—last but not least—my mother, who has always been there for me in every conceivable way, particularly since the passing of my father. Thank you all.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Overview

One of the most important geometric functionals is the area (or volume) functional for hypersurfaces Σ^{n-1} inside Riemannian manifolds (M^n, g). Suppose Σ carries a unit normal vector field ν. The **first variation** of area along a normal perturbation $f\nu$, $f \in C_c^\infty(\Sigma)$, is given by

$$\int_\Sigma fH \, dv_\Sigma;$$

$H = \operatorname{div}_\Sigma \nu$ is the **mean curvature** of $\Sigma^{n-1} \subset (M^n, g)$. Critical points for the area functional are called **minimal hypersurfaces**.

Not all critical points have the same variational behavior. The **Morse index** of a critical point measures the number of linearly independent directions for instability. This is captured by the **second variation** of area, which, along the same normal perturbation, is given by

$$\int_\Sigma \left[\|\nabla^\Sigma f\|^2 - \| \mathrm{I\!I}_\Sigma \|^2 f^2 - \operatorname{Ric}_M(\nu, \nu) f^2 \right] \, dv_\Sigma,$$

where ∇^Σ is the tangential gradient on Σ, $\mathrm{I\!I}_\Sigma$ is the second fundamental form of Σ with respect to ν, and Ric_M is the ambient Ricci curvature. A minimal hypersurface with nonnegative definite second variation is called **stable**. From a physical perspective, unstable critical points are a lot less likely to be observed than stable ones.

In this thesis we study the critical behavior of certain geometric functionals, including the area functional, which appear in two different problems in mathematical physics: the initial data set approach to the general theory of relativity, and the theory of phase transitions.

Several problems in the initial data set approach to general relativity reduce entirely to variational problems in Riemannian geometry. Specifically, under certain symmetry assumptions, energy is modeled by scalar curvature, and boundaries of quasi-local black hole regions ("apparent horizons") by stable minimal hypersurfaces.

In Chapter 2 we obtain a precise local understanding of the interaction of ambient scalar curvature and stable minimal surfaces, in the context of three-dimensional manifolds with nonnegative scalar curvature. In the process of doing so, we define a new moduli space of metrics on surfaces, $\mathscr{M}_+$, and introduce a new metric invariant, the Λ-invariant.

In Chapter 3 we discuss the initial data set approach to the general theory of relativity, and apply to it the variational tools developed in Chapter 2 regarding the moduli space "$\mathscr{M}_+$" and the "Λ-invariant." Among other results: we compute the exact value of the Bartnik mass of apparent horizons, a previously uncomputable quantity; we show the static minimization conjecture cannot hold for apparent horizons; we construct a large class of new examples of apparent horizos; we disprove G. Gibbons' formulation of Thorne's hoop conjecture in relativity; we formulate a geometric generalization of the Brown-York quasi-local mass; and, we derive a "cut-and-fill" operator for three-dimensional manifolds of positive scalar curvature, which, with respect to the Λ-invariant, is a dual operation to the Schoen-Yau [61] and Gromov-Lawson [28] connect-sum construction.

In Chapter 4 we turn our attention to the theory of phase transitions, as modeled by the scalar-valued Allen-Cahn equation. De Giorgi conjectured that solutions of this equation behave like minimal hypersurfaces. We discuss $2k$-ended solutions of the equation in $\mathbf{R}^2$ and confirm a conjecture regarding a lower bound for the Morse index of $2k$-ended solutions.

1.2 Notation and conventions

Table 1.1: Summary of notation

Notation	**Explanation**
$\approx$	diffeomorphism, diffeomorphic manifolds
$\cong$	isometry, isometric manifolds
$\triangleq$	definition of variable, function, etc.
$\nabla^{M,g}$	Levi-civita connection, gradient on Riemannian manifold (M,g)
$\nabla^k_{M,g}$	k-th covariant derivative on Riemannian manifold (M,g)
$\Delta_{M,g}$	Laplace-Beltrami operator, $\operatorname{Tr}\nabla^2$, on Riemannian manifold (M,g)

Table 1.1: Summary of notation

Notation	**Explanation**
$\Gamma(E)$	smooth sections of a vector bundle E over a smooth manifold
Λ^{2k}	$2k$-tuple of oriented rays in $\mathbf{R}^2$ (Section 4.2.1)
$\Lambda_{(\Sigma,\gamma)}$	Λ-invariant of Riemannian 2-manifold (Σ, γ) (Section 2.3)
C^k	Banach space of k-times continuously differentiable functions
C^k_{loc}	vector space of locally C^k functions
C^k_c	vector space of compactly supported C^k functions
$C^{k,\alpha}$	Banach space of C^k functions with α-Hölder k-th derivatives
$C^{k,\alpha}_{\mathrm{loc}}$	vector space of locally $C^{k,\alpha}$ functions
H_Σ, $\mathbf{H}_\Sigma$	mean curvature scalar and vector of a hypersurface Σ (Section A.2)
$\mathrm{I\!I}_\Sigma$	second fundamental form of a hypersurface Σ (Section A.2)
$\mathrm{ind}(\cdot)$	Morse index for a critical point of a variational problem (Section 4.3.1)
L^p	Banach space of functions with $\int \lvert f\rvert^p < \infty$
L^p_{loc}	vector space of locally L^p functions in
$\mathrm{Met}(M)$	bundle of metrics on a smooth manifold M
$\mathcal{M}_{2k}$	moduli space of $2k$-ended Allen-Cahn solutions in $\mathbf{R}^2$ (Section 4.2.1)
$\mathscr{M}_+$	moduli space of special metrics on $\mathbf{S}^2$ (Section 2.2)
Rm, Ric, R, K	Riemann, Ricci, scalar, Gauss curvature (Section A.1)
$W^{k,p}$	Banach space of functions with $f, \nabla f, \ldots, \nabla^k f \in L^p$ ("Sobolev space")
$W^{k,p}_{\mathrm{loc}}$	locally $W^{k,p}$ functions
$W^{k,p}_0$	$W^{k,p}$-closure of C^∞_c
$\mathbf{X}$, $\mathbf{Y}$, $\mathbf{e}$, $\mathbf{f}$, etc.	vectors, vector fields
υ, $d\upsilon$, σ, $d\sigma$, ℓ, $d\ell$	volume, surface, length forms or measures

Additionally:

- we simplify notation by dropping subscripts and superscripts, if those are clear from the context; e.g., the Levi-Civita connection $\nabla^{M,g}$ of a Riemannian manifold (M, g) can be ∇^M if the metric is clear but the manifold isn't, ∇^g if the manifold is clear but the metric isn't, or ∇ if everything is clear from the context;

- completeness and compactness for Riemannian manifolds will always refer to metric completeness and compactness; i.e., complete and compact Riemannian manifolds may have boundary; closed Riemannian manifolds are compact and without boundary.

Chapter 2

Scalar Curvature 3-manifold Geometry

2.1 Introduction

One of the most fundamental questions in differential geometry is:

Question. *Which closed manifolds M^n, $n \geq 2$, can carry Riemannian metrics with (everywhere) positive scalar curvature?*

We briefly recall the various curvature tensors of an n-dimensional Riemannian manifold (M^n, g). Appendix A lists basic facts about Riemannian geometry—including some about curvature tensors—that are needed in this thesis. For a more thorough treatment, see [55, 48].

The Riemann curvature 4-tensor is:

$$\mathrm{Rm}(\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}) \triangleq \langle \nabla_{\mathbf{X}} \nabla_{\mathbf{Y}} \mathbf{Z} - \nabla_{\mathbf{Y}} \nabla_{\mathbf{X}} \mathbf{Z} - \nabla_{[\mathbf{X}, \mathbf{Y}]} \mathbf{Z}, \mathbf{W} \rangle, \tag{2.1.1}$$

the Ricci curvature 2-tensor is:

$$\mathrm{Ric}(\mathbf{X}, \mathbf{Y}) \triangleq \mathrm{Tr}\,\mathrm{Ric}(\mathbf{X}, \cdot, \cdot, \mathbf{Y}), \tag{2.1.2}$$

and the scalar curvature function is:

$$R = \mathrm{Tr}\,\mathrm{Ric}(\cdot, \cdot). \tag{2.1.3}$$

When $n = 2$, the situation is completely understood via elementary considerations. First, the Ricci and scalar curvature are given, respectively, by Kg and $2K$, where K denotes the Gauss curvature of the two-dimensional Riemannian manifold (M^2, g). Therefore, from the Gauss-Bonnet theorem we know that for every metric g on M^2, the scalar curvature integrates to a topological invariant, namely,

$$\int_M R \, d\sigma = 4\pi\chi(M),$$

where $\chi(M)$ is the Euler-characteristic of M. In particular, if M carries a metric with $R > 0$, then M will have $\chi(M) > 0$. By the classification of two-dimensional surfaces, we know that, up to diffeomorphism, the only closed ones with positive Euler characteristic are $\mathbf{S}^2$ and $\mathbf{RP}^2$. Conversely, it is obviously true that $\mathbf{S}^2$ and $\mathbf{RP}^2$ carry metrics of positive scalar curvature; e.g., the round metric on $\mathbf{S}^2$ and its corresponding quotient on $\mathbf{RP}^2$. In other words:

Fact. *A closed two-dimensional manifold carries a metric of positive scalar curvature if and only if it is diffeomorphic to* $\mathbf{S}^2$ *or* $\mathbf{RP}^2$*.*

When $n = 3$, the situation is now also completely understood, but the methods involved are far from elementary; this case was settled given the combined work of Schoen-Yau [61, 59], Gromov-Lawson [28], and Perelman's proof of the geometrization theorem [52, 54, 53].

Fact. *A closed three-dimensional manifold carries a metric of positive scalar curvature if and only if it is diffeomorphic to a connected sum of finitely many* $\mathbf{S}^2 \times \mathbf{S}^1$*s together with finitely many spherical three-dimensional manifolds* $\mathbf{S}^3/G$ *(*G *a finite subgroup of* $\mathrm{SO}(4)$*, acting freely by rotations).*

The methods in this chapter are in the spirit of those in the Schoen-Yau approach to the classification of three-dimensional manifolds with positive scalar curvature. Namely, we make extensive use of the theory of minimal surfaces. Appendix A.2 contains a brief summary of minimal surface results we are going need.

2.2 Moduli space $\mathscr{M}_+$ of metrics on $\mathbf{S}^2$

2.2.1 Definitions

We start this section with a lemma that motivates the moduli space of metrics that we will study.

Lemma 2.2.1 (M.). *Let* (Σ^2, γ) *be a closed two-dimensional Riemannian manifold. The following are equivalent:*

(1) *For every* $f \in C^\infty(\Sigma^2)$, $f \not\equiv 0$,

$$\int_\Sigma \left[\|\nabla^\Sigma f\|^2 + K_\Sigma f^2\right] \, d\sigma > 0,$$

where K_Σ *is the Gauss curvature;*

(2) *There exists* $u \in C^\infty(\Sigma^2)$, $u > 0$, *such that* $(\Sigma^2 \times \mathbf{S}^1, \gamma + u^2 d\theta^2)$ *has positive scalar curvature;*

(3) (Σ^2, γ) *occurs as a two-sided stable minimal surface in a three-dimensional Riemannian manifold with positive scalar curvature.*

Proof. (1) $\implies$ (2). From Propositions B.2.2, B.2.4, the first eigenfunction $u \in C^\infty(\Sigma^2)$ of the operator $-\Delta_\Sigma + K_\Sigma$ on (Σ^2, γ) satisfies $u > 0$ and

$$\lambda_1(L_\gamma) \int_\Sigma u^2 \, d\sigma = \int_\Sigma \left[\|\nabla^\Sigma u\|^2 + K_\Sigma u^2\right] \, d\sigma > 0,$$

where $\lambda_1(L_\gamma) \in \mathbf{R}$ is the first eigenvalue of L_γ; i.e., $\lambda_1(L_\gamma) > 0$. Then, we simply take $M \triangleq \Sigma^2 \times \mathbf{S}^1$ with the metric $g \triangleq \gamma + u^2 d\theta^2$. From Lemma A.1.6 we see that the scalar curvature of (M^3, g) is

$$R_M = \frac{2}{u} \left(-\Delta_\Sigma u + K_\Sigma u\right) \equiv 2\lambda_1(L_\gamma) > 0.$$

(2) $\implies$ (3). We claim that we can simply take $M^3 \triangleq \Sigma^2 \times \mathbf{S}^1$ endowed with $g \triangleq \gamma + u^2 d\theta^2$, which we already know has positive scalar curvature. We need to show that $\Sigma^2 \times \{\theta_0\} \subset M^3$, $\theta_0 \in \mathbf{S}^1$, is stable. Recall that, by Lemma A.1.6,

$$R_M = \frac{2}{u} \left(-\Delta_\Sigma u + K_\Sigma u\right) \iff R_M - 2K_\Sigma = -2\frac{\Delta_\Sigma u}{u} = -2\Delta_\Sigma \log u - 2\|\nabla^\Sigma \log u\|^2. \tag{2.2.1}$$

Moreover, from the same lemma, $\Sigma^2 \times \{\theta_0\}$ is totally geodesic in M^3. Therefore, for every $f \in C^\infty(\Sigma^2)$, $f \not\equiv 0$, we have

$$\begin{aligned}
&\int_\Sigma \left[\|\nabla^\Sigma f\|^2 - \frac{1}{2}(R_M - R_\Sigma + \| \mathrm{I\!I}_\Sigma \|^2) f^2\right] d\sigma \\
&\quad = \int_\Sigma \left[\|\nabla^\Sigma f\|^2 - \frac{1}{2}\left(R_M - 2K_\Sigma\right) f^2\right] d\sigma \\
&\quad = \int_\Sigma \left[\|\nabla^\Sigma f\|^2 + (\Delta_\Sigma \log u) f^2 + \|\nabla^\Sigma \log u\|^2 f^2\right] d\sigma \\
&\quad = \int_\Sigma \left[\|\nabla^\Sigma f\|^2 - 2f\langle \nabla^\Sigma \log u, \nabla^\Sigma f\rangle + \|\nabla^\Sigma \log u\|^2 f^2\right] d\sigma,
\end{aligned}$$

which is nonnegative after using Cauchy-Schwarz to estimate $2f\langle \nabla^\Sigma \log u, \nabla^\Sigma f\rangle \geq -\|\nabla^\Sigma f\|^2 -$

$\|\nabla^\Sigma \log u\|^2 f^2$. In particular, Σ^2 is stable due to Proposition A.2.6.

(3) $\implies$ (1). Since Σ^2 is a two-sided stable minimal surface in (M^3, g), by Proposition A.2.6 we know that, for every $f \in C^\infty(\Sigma^2)$,

$$\int_\Sigma \|\nabla^\Sigma f\|^2 - \frac{1}{2}\left(R_M - 2K_\Sigma + \|\,\mathrm{I\!I}_\Sigma\,\|^2\right) f^2 d\sigma \geq 0.$$

The result follows because $R_M > 0$, $\|\,\mathrm{I\!I}_\Sigma\,\|^2 \geq 0$. $\square$

In particular, we see that the intrinsic operator $-\Delta + K$ on two-dimensional Riemannian manifolds captures the ability of this manifold to occur as a two-sided stable minimal surface in a three-dimensional Riemannian manifold with positive scalar curvature. This naturally leads to the following moduli space of two-dimensional Riemannian manifolds.

Definition 2.2.2 (M.-Schoen [40]). *Let*

$$\mathscr{M}_+ \triangleq \{(\mathbf{S}^2, \gamma) : \lambda_1(-\Delta_\gamma + K_\gamma) > 0\}$$

be the moduli space of Riemannian 2-spheres $(\mathbf{S}^2, \gamma)$ *on which the second order elliptic operator* $L_\gamma = -\Delta_\gamma + K_\gamma$ *is positive definite. Endow* $\mathscr{M}_+$ *with the subspace topology of* $\mathrm{Met}(\mathbf{S}^2)$*. Here* $\lambda_1(\cdot)$ *represents the first eigenvalue of an elliptic operator.*

Remark 2.2.3. *Note that, by Proposition A.2.7 and Lemma 2.2.1,* $\mathbf{S}^2$ *is the only closed orientable surface that can carry metrics whose associated operator* $-\Delta + K$ *is positive definite.*

2.2.2 Results

The following proposition guarantees that $\mathscr{M}_+$ is nontrivial.

Proposition 2.2.4 (M.)**.** $\mathscr{M}_+$ *contains all* $(\mathbf{S}^2, \gamma)$ *with* $K_\gamma \geq 0$.

Proof. Applying Proposition B.2.4, let $f \in C^\infty(\mathbf{S}^2)$, $f > 0$, be a first eigenfunction of L_γ normalized to have unit L^2 norm. Then

$$\lambda_1(L_\gamma) = \int_{\mathbf{S}^2} \|\nabla^\gamma \left[f\|_\gamma^2 + K_\gamma f^2\right] d\sigma_\gamma \geq 0,$$

and this is a strict inequality unless $K_\gamma \equiv 0$, which is impossible in view of the Gauss-Bonnet theorem on 2-spheres:

$$\int_{\mathbf{S}^2} K_\gamma \, d\sigma_\gamma = 2\pi\chi(\mathbf{S}^2) = 4\pi.$$

The result follows. $\square$

Proposition 2.2.5 (M.)**.** *$\mathscr{M}_+$ is closed under dilations and diffeomorphisms.*

Proof. This follows from $\lambda_1(L_{t^2\gamma}) = t^{-2}\lambda_1(L_\gamma)$, $t > 0$, and $\lambda_1(L_{\varphi^*\gamma}) = \lambda_1(L_\gamma)$, $\varphi \in \mathrm{Diff}(\mathbf{S}^2)$. $\square$

In fact, $\mathscr{M}_+$ is a much larger moduli space than Proposition 2.2.4 suggests. According to the theorem below, it contains elements with arbitrarily large negative integral curvature.

Theorem 2.2.6 (M.-Schoen [40, Theorem 3.1])**.** *$\mathscr{M}_+$ is relatively open in every pointwise conformal class of metrics on $\mathbf{S}^2$, with respect to the (weaker) C^1 topology on $\mathrm{Met}(\mathbf{S}^2)$. Furthermore, for every $c > 0$, the subset*

$$\left\{(\mathbf{S}^2, \gamma) \in \mathscr{M}_+ : \int_{\mathbf{S}^2} (K_\gamma)_- \, d\sigma_\gamma \geq c\right\} \subset \mathscr{M}_+$$

is C^1 dense in $\mathscr{M}_+$; here, $(K_\gamma)_- = \max\{0, -K_\gamma\}$.

This theorem requires that we gain further insight on $\mathscr{M}_+$. We will use the following auxiliary differential operator

$$\mathcal{Q}_{w,\gamma_*}\phi = -\Delta_*\phi - \|\nabla^*(\phi - w)\|_*^2 + K_*,$$

where Δ_*, ∇^*, K_*, $\|\cdot\|_*$, are computed with respect to some round metric γ_* on $\mathbf{S}^2$. The following proposition relates $\mathscr{M}_+$ with $\mathcal{Q}_{w,\gamma_*}$ and strengthens Proposition 2.2.4. above.

Lemma 2.2.7 (M.-Schoen [40, Proposition 3.2])**.** *Suppose $w \in C^\infty(\mathbf{S}^2)$ and γ_* is a round metric on $\mathbf{S}^2$. Then*

$$(\mathbf{S}^2, e^{2w}\gamma_*) \in \mathscr{M}_+ \iff \textit{there exists } \phi \in C^\infty(\mathbf{S}^2) \textit{ with } \mathcal{Q}_{w,\gamma_*}\phi > 0.$$

Proof. We first claim that

$$(\mathbf{S}^2, \gamma) \in \mathscr{M}_+ \iff \text{there exists } f \in C^\infty(\mathbf{S}^2) \text{ with } f > 0 \text{ and } L_\gamma f > 0. \tag{2.2.2}$$

The ($\Longrightarrow$) direction of (2.2.2) follows by choosing f to be the first eigenfunction of L_γ, by Proposition B.2.4 and the definition of $\mathscr{M}_+$.

The ($\Longleftarrow$) direction of (2.2.2) goes by contradiction. Suppose $f > 0$ and $L_\gamma f > 0$, but that $(\mathbf{S}^2, \gamma) \notin \mathscr{M}_+$; i.e., $\lambda \triangleq \lambda_1(L_\gamma) \leq 0$. Choose $h > 0$ to be the first eigenfunction of L_γ by Proposition B.2.4, so that $L_\gamma h = \lambda h$ with $\lambda \leq 0$. Since $\mathbf{S}^2$ is compact and neither one of f, h vanishes, there must exist a constant $t > 0$ and a point $p \in \mathbf{S}^2$ such that $f \geq th$ and $f(p) = th(p)$. The auxiliary map $f - th$ attains a global minimum, zero, at p, so

$$\begin{aligned} L_\gamma(f-th)(p) &= -\Delta_\gamma(f-th)(p) + K_\gamma(f-th)(p) = -\Delta_\gamma(f-th)(p) \le 0 \\ \iff L_\gamma f(p) &\le t L_\gamma h(p). \end{aligned}$$

This contradicts that $L_\gamma f > 0$ and $L_\gamma h = \lambda h \le 0$ on $\mathbf{S}^2$. The claim follows.

Next, since every $f > 0$ can be written as e^ζ, and $L_\gamma e^\zeta = e^\zeta\big(-\Delta_\gamma\zeta - \|\nabla^\gamma\zeta\|_\gamma^2 + K_\gamma\big)$, the previous claim can be restated as:

$$(\mathbf{S}^2,\gamma) \in \mathscr{M}_+ \iff \text{there exists } \zeta \in C^\infty(\mathbf{S}^2) \text{ with } -\Delta_\gamma\zeta - \|\nabla^\gamma\zeta\|_\gamma^2 + K_\gamma > 0. \tag{2.2.3}$$

Finally, take $\gamma = e^{2w}\gamma_*$ as in the statement. Using Lemma A.1.7 to conformally change the metric from $\gamma = e^{2w}\gamma_*$ to γ_* we have

$$\begin{aligned} &-\Delta_\gamma\zeta - \|\nabla^\gamma\zeta\|_\gamma^2 + K_\gamma \\ &\quad = e^{-2w}\big(-\Delta_*\zeta - \|\nabla^*\zeta\|_*^2 + K_* - \Delta_* w\big) \\ &\quad = e^{-2w}\big(-\Delta_*(\zeta+w) - \|\nabla^*\zeta\|_*^2 + K_*\big) = \mathcal{Q}_{w,\gamma_*}(\zeta+w). \end{aligned}$$

The conclusion follows with $\phi = \zeta + w$. □

Proof of Theorem 2.2.6. First let us prove C^1 openness. Fix any element of $\mathscr{M}_+$. By the uniformization theorem, it can be written as $e^{2v}\gamma_*$ for some round metric γ_* that generates the pointwise conformal class, and some $v \in C^\infty(\mathbf{S}^2)$. By Lemma 2.2.7 there must exist $\phi \in C^\infty(\mathbf{S}^2)$ with $\mathcal{Q}_{v,\gamma_*}\phi > 0$. For any $w \in C^\infty(\mathbf{S}^2)$ and the same round metric γ_*,

$$\begin{aligned} \mathcal{Q}_{w,\gamma_*}\phi - \mathcal{Q}_{v,\gamma_*}\phi &= \big(-\Delta_*\phi - \|\nabla^*(\phi-w)\|_*^2 + K_*\big) - \big(-\Delta_*\phi - \|\nabla^*(\phi-v)\|_*^2 + K_*\big) \\ &= \|\nabla^*(\phi-v)\|_*^2 - \|\nabla^*(\phi-w)\|_*^2 \\ &= \langle\nabla^*(2\phi-w-v), \nabla^*(w-v)\rangle_* \\ &\ge -\|\nabla^*(2\phi-w-v)\|_*\|\nabla^*(w-v)\|_*, \end{aligned}$$

so

$$\mathcal{Q}_{w,\gamma_*}\phi \ge \mathcal{Q}_{v,\gamma_*}\phi - \|\nabla^*(2\phi-w-v)\|_*\|\nabla^*(w-v)\|_*; \tag{2.2.4}$$

i.e., $\mathcal{Q}_{w,\gamma_*}\phi > 0$ if $\|\nabla^*(w-v)\|_*$ is small enough. The claim follows from Lemma 2.2.7.

Next we need to establish the density claim. Let v, ϕ be as before, and $n \in \{1,2,\ldots\}$, $\alpha \in (0,1)$ be parameters that are to be determined. Without loss of generality, write $\gamma_* = d\theta^2 + \sin^2\theta\, d\phi^2$. Define $h = -n^{-1}\cos(n\theta)$, which satisfies $\nabla^* h = \sin(n\theta)\frac{\partial}{\partial\theta}$ and $\Delta_* h = n\cos(n\theta) + \sin(n\theta)\cot\theta$ and extends smoothly across the north ($\theta = 0$) and south ($\theta = \pi$) poles. We now choose $\alpha \in (0,1)$ by the C^1-openness in conformal classes to be small enough that $(\mathbf{S}^2, e^{2(v+\alpha h)}\gamma_*) \in \mathscr{M}_+$. This can be done independently of n. The conformal metric $\gamma = e^{2(v+\alpha h)}\gamma_*$ satisfies, in view of Lemma A.1.7,

$$\begin{aligned}\int_{\mathbf{S}^2} (K_\gamma)_- \, d\sigma_\gamma &= \int_{\mathbf{S}^2} \big(1 - \Delta_* v - \alpha\, \Delta_* h\big)_- \, d\sigma_* \\ &= \int_{\mathbf{S}^2} \Big[\big(1 - \Delta_* v - \alpha\, n\, \cos(n\theta) - \alpha\, \sin(n\theta)\, \cot\theta\big)_-\Big] \, d\sigma_*.\end{aligned}$$

If $D = \{\pi/3 \leq \theta \leq 2\pi/3\}$ then $1 - \Delta_{\gamma_*} v - \alpha\, \sin(n\theta)\, \cot(\theta) \leq \Lambda = \Lambda(v, \alpha)$ on D and $\cos(n\theta) \geq 1/2$ on a set $D_n \subset D$ with $\sigma_*(D_n) \geq \mu > 0$ independently of n. Altogether this gives

$$\int_{\mathbf{S}^2} (K_\gamma)_- \, d\sigma_\gamma \geq \int_{D_n} \big(\Lambda - \frac{1}{2}\, \alpha\, n\big)_- \, d\sigma_* \geq \big(\Lambda - \frac{1}{2}\, \alpha\, n\big)_- \mu,$$

which can be made arbitrarily large by sending $n \uparrow \infty$. The density claim follows, since this can be done for all small enough $\alpha \in (0,1)$. $\square$

There is an immediate corollary to the computation performed to check for relative openness above.

Corollary 2.2.8 (M.-Schoen [40, Corollary 3.3])**.** *If γ_* is round, $w \in C^\infty(\mathbf{S}^2)$, and $\|\nabla^* w\|_*^2 < K_*$, then $(\mathbf{S}^2, e^{2w}\gamma_*) \in \mathscr{M}_+$.*

Proof. Choosing $v = \phi = 0$ in (2.2.4), we get

$$\mathcal{Q}_{w,\gamma_*} 0 \geq \mathcal{Q}_{0,\gamma_*} 0 - \|\nabla^* w\|_*^2 = K_* - \|\nabla^* w\|_*^2,$$

which is positive when $\|\nabla^* w\|_*^2 < K_*$. $\square$

Next, we establish an important topological property of $\mathscr{M}_+$, which essentially says that any element of $\mathscr{M}_+$ is cobordant to a round metric, along a special bridge of positive scalar curvature. We will refer to this property as "strong PSC cobordance."

Proposition 2.2.9 (Strong PSC cobordance, M., cf. [40, Lemma 1.2])**.** *Let $(\mathbf{S}^2, \gamma) \in \mathscr{M}_+$ be given. There exists a Riemannian manifold $(M^3, g) \cong (\mathbf{S}^2 \times [0,1], g)$ foliated by 2-spheres $(\Sigma_t, \gamma(t))$, $t \in [0,1]$, such that*

(1) $(\Sigma_t, \gamma(t)) \in \mathscr{M}_+$ for all $t \in [0,1]$;

(2) $(\Sigma_0, \gamma(0)) \cong (\mathbf{S}^2, \gamma)$, $(\Sigma_1, \gamma(1))$ is round, and all Σ_t, $t \in [0,1/3] \cup [2/3,1]$, are totally geodesic in (M^3, g);

(3) all Σ_t, $t \in [0,1]$, are minimal 2-spheres in (M^3, g);

(4) (M^3, g) *has positive scalar curvature.*

Specifically,

$$g = \gamma(t) + v(\cdot, t)^2 dt^2,$$

with $v(\cdot, t) > 0$ *a first eigenfunction of* $L_t \triangleq L_{\gamma(t)}$*, for all* $t \in [0, 1]$.

Proof. We will construct the metric g in steps, ensuring one conclusion at a time. By the uniformization theorem, we may write $\gamma = e^{2w}\gamma_*$ for some round metric γ_* of the same area. Indicate covariant derivatives, the Laplace-Beltrami operator, the area form, and norms taken with respect to γ_* by ∇^*, Δ_*, σ_*, $\|\cdot\|_*$. Likewise, we will later encounter families of $\gamma^k(t)$, for which we will respectively use the notation $\nabla^{k,t}$, $\Delta_{k,t}$, $\sigma_{k,t}$, $\|\cdot\|_{k,t}$.

Fix a smooth, decreasing $\zeta : [0, 1] \to [0, 1]$ with $\zeta(0) = 1$, $\zeta(1) = 0$, and $\zeta' \equiv 0$ on $[0, 1/3] \cup [2/3, 1]$.

Claim. Let $\gamma^1(t) \triangleq e^{2\zeta(t)w}\gamma_*$. The metric $\gamma^1(t) + dt^2$, with $\Sigma_t = \mathbf{S}^2 \times \{t\}$, satisfies (1).

Proof of claim. Let $f \in C^\infty(\mathbf{S}^2) \setminus \{0\}$. Recall that

$$\int_{\mathbf{S}^2} f L_{1,t} f \, d\sigma_{1,t} = \int_{\mathbf{S}^2} \left[\|\nabla^{1,t} f\|_{1,t}^2 + K_{1,t} f^2\right] \, d\sigma_{1,t} = \int_{\mathbf{S}^2} \left[\|\nabla^* f\|_*^2 + (K_* - \zeta(t)\Delta_* w) f^2\right] \, d\sigma_*,$$

where the second equality follows from the conformal invariance of Dirichlet energy in dimension 2 and the change of Gauss curvature under conformal metric change (Lemma A.1.7). The right hand side above is positive when $t = 0$ (since $\gamma^1(0) = \gamma$ and $(\mathbf{S}^2, \gamma) \in \mathscr{M}_+$) and positive when $t = 1$ (since $\gamma^1(1) = \gamma_*$ and $(\mathbf{S}^2, \gamma_*) \in \mathscr{M}_+$ by Proposition 2.2.4), so it must be positive for all $t \in [0, 1]$, seeing as to how it is also an affine function of t. In other words,

$$\int_{\mathbf{S}^2} f L_{1,t} f \, d\sigma_{1,t} > 0$$

for all $f \in C^\infty(\mathbf{S}^2) \setminus \{0\}$, $t \in [0, 1]$, so $(\mathbf{S}^2, \gamma^1(t)) \in \mathscr{M}_+$. $\square$

Recalling that $\mathscr{M}_+$ is invariant under dilations (Proposition 2.2.5), we see that there exists a smooth $\lambda : [0, 1] \to \mathbf{R}$ with $\lambda(0) = \lambda(1) = 1$, so that $\gamma^2(t) \triangleq e^{2\zeta(t)w + 2\lambda(t)}\gamma_*$ all have the same area; namely,

$$\lambda'(t) = -\zeta'(t) \fint_{\mathbf{S}^2} w \, d\sigma_{1,t}, \tag{2.2.5}$$

where the integral above indicates a mean-value integral. Replacing $\gamma^1(t)$ by $\gamma^2(t)$ and recalling $\zeta' = 0$ on $[0, 1/3] \cup [2/3, 1]$, Lemma A.1.6 shows that (1), (2) are both true for $\gamma^2(t) + dt^2$, with $\Sigma_t = \mathbf{S}^2 \times \{t\}$.

Claim. There exists a smooth path $t \mapsto \mathbf{X}_t \in \Gamma(T\mathbf{S}^2)$, $t \in [0,1]$, such that

$$\operatorname{div}_{2,t} \mathbf{X}_t = -2(\zeta'(t)w + \lambda'(t)) \text{ for all } t \in [0,1].$$

Proof of claim. From the maximum principle (Theorem B.1.1), $\ker \Delta_{2,t} = \{\text{constants}\} \subset C^\infty(\mathbf{S}^2)$ for every $t \in [0,1]$. Additionally, from (2.2.5) and the fact that $\gamma^2(t)$ is a dilation of $\gamma^1(t)$,

$$\int_{\mathbf{S}^2} -2(\zeta'(t)w + \lambda'(t))\, d\sigma_{2,t} = 0.$$

From Lemma B.3.1, there exists a smooth map $t \mapsto \psi(\cdot,t)$ of solutions to $\Delta_{2,t}\psi(\cdot,t) = -2(\zeta'(t)w + \lambda'(t))$. The claim follows for $\mathbf{X}_t \triangleq \nabla^{2,t}\psi(\cdot,t)$. Note that $\mathbf{X}_t = \mathbf{0}$ when $t \in [0,1/3] \cup [2/3,1]$. ☐

Let $\{\varphi_t\}_{t\in[0,1]} \in \operatorname{Diff}(\mathbf{S}^2)$ denote the integral flow generated by $\{\mathbf{X}_t\}_{t\in[0,1]}$, and $\gamma^3(t) \triangleq {\varphi_t}^*\gamma^2(t)$.

Claim. The metric $\gamma^3(t) + dt^2$ with $\Sigma_t = \mathbf{S}^2 \times \{t\}$ satisfies (1)-(3).

Proof. By Proposition 2.2.5, (1)-(2) remain true after replacing $\gamma^2(t)$ with $\gamma^3(t)$, keeping in mind that $\mathbf{X}_t = \mathbf{0}$ when $t \in [0,1/3] \cup [2/3,1]$. Moreover, differentiating the smooth family of 2-forms $d\sigma_{3,t}$ in t gives

$$\frac{d}{dt} d\sigma_{3,t} = \frac{d}{dt}\left[{\varphi_t}^* d\sigma_{2,t}\right] = {\varphi_t}^*\left[\frac{d}{dt} d\sigma_{2,t} + \mathcal{L}_{\mathbf{X}_t} d\sigma_{2,t}\right]. \tag{2.2.6}$$

Computing the evolution of the area forms $d\sigma_{2,t}$ as in Lemma A.1.6, and also $\mathcal{L}_{\mathbf{X}_t} d\sigma_{2,t} = d(\iota_{\mathbf{X}_t} d\sigma_{2,t}) = \operatorname{div}_{2,t} \mathbf{X}_t\, d\sigma_{2,t}$, (2.2.6) becomes

$$\begin{aligned}\frac{d}{dt} d\sigma_{3,t} &= {\varphi_t}^*\left[\frac{1}{2}\operatorname{Tr}_{2,t}\left(\frac{d}{dt}\gamma_t^2\right) d\sigma_{2,t} + \operatorname{div}_{2,t}\mathbf{X}_t\, d\sigma_{2,t}\right] \\ &= {\varphi_t}^*\left[\left(2(\zeta'(t)w + \lambda'(t)) + \operatorname{div}_{2,t}\mathbf{X}_t\right) d\sigma_{2,t}\right] = 0.\end{aligned}$$

The claim follows from Lemma A.1.6 again, since $\frac{d}{dt} d\sigma_{3,t}$ is the mean curvature. ☐

Finally, we seek to fulfill all of (1)-(4). Using Lemma B.3.2 we may define a smooth function $u : \mathbf{S}^2 \times [0,1] \to \mathbf{R}$ so that, for every $t \in [0,1]$, $u(\cdot,t)$ is a first eigenfunction for the operator $L_{3,t} = -\Delta_{3,t} + K_{3,t}$. Let $\alpha > 0$ be a constant that is yet to be determined, and consider the metric

$$g_\alpha = \gamma^3(t) + \alpha^{-2}u(\cdot,t)^2\, dt^2,$$

on $\mathbf{S}^2 \times [0,1]$. By the scalar curvature expression in Lemma A.1.6 we see that

$$R_{g_\alpha} = 2\lambda_1(L_{3,t}) - \frac{\alpha^2}{4}\left\|\frac{\partial}{\partial t}\gamma_t^3\right\|_{3,t}^2 - \frac{\alpha^2}{u}\frac{\partial}{\partial t}\left(\frac{1}{u}\operatorname{Tr}_{3,t}\frac{\partial}{\partial t}\gamma_t^3\right),$$

which is positive for sufficiently small $\alpha > 0$, since $\lambda_1(L_{3,t}) > 0$. $\square$

We will often combine the previous lemma with:

Lemma 2.2.10 (Gluing lemma, M., cf. [40, Lemma 2.2])**.** *Suppose $f_i \in C^\infty(I_i)$, $I_i \triangleq [a_i, b_i]$, $i = 1, 2$, are such that*

(1) $f_i > 0$, $f_i' \geq 0$, $f_i'' \geq 0$ on I_i,

(2) $f_i'' < \frac{1}{2} f_i(1 - f_i^{-2}(f_i')^2)$ on I_i,

(3) $f_1(b_1) \leq f_2(a_2)$, and

(4) $f_1'(b_1) = f_2'(a_2)$.

Translating the intervals I_1, I_2, so that $f_2(a_2) - f_1(b_1) = (a_2 - b_1) f_1'(b_1)$, there exists $f \in C^\infty([a_1, b_2])$ such that

(1) $f > 0$, $f' \geq 0$, $f'' \geq 0$,

(2) $f \equiv f_1$ on $[a_1, \frac{a_1+b_1}{2}]$ and $f \equiv f_2$ on $[\frac{a_2+b_2}{2}, b_2]$, and

(3) if (M^n, g) has scalar curvature bounded from below per

$$R_g \geq n(n-1)\rho_0 > 0,$$

then $f(t)^2 g + dt^2$ is a positive scalar curvature metric on the product space $M \times [a_1, b_2]$.

If $f_1'(b_1)$, $f_2'(a_2) \neq 0$, then no additional critical points are introduced in f.

Remark 2.2.11. *Hypothesis (3) is not used in the proof, but in translating I_1, I_2 as in the statement.*

Proof of Lemma 2.2.10. Define

$$F(t) = \begin{cases} f_1(t) & \text{for } t \in I_1 \\ \frac{a_2 - t}{a_2 - b_1} \cdot f_1(b_1) + \frac{t - b_1}{a_2 - b_1} \cdot f_2(a_2) & \text{for } t \in (b_1, a_2) \\ f_2(t) & \text{for } t \in I_2 \end{cases}$$

to be a linear interpolation joining f_1, f_2. In view of the provided hypotheses, $F \in C^{1,1}([a_1, b_2])$, and is nondecreasing. Define $F_\sigma \in C^\infty[a_1, b_2]$ by mollifying F, with $\sigma > 0$ small and a standard symmetric mollifier decreasing in the radial direction, and so that $F_\sigma \equiv F$ on $[a_1, \frac{a_1+b_1}{2}] \cup [\frac{a_2+b_2}{2}, b_2]$. Likewise, F_σ is nondecreasing.

Claim. F_σ satisfies the conclusion of the lemma when $\sigma > 0$ is small enough.

Proof of claim. Conclusions (1), (2), and the non-introduction of critical points when $f_1'(b_1)$, $f_2'(a_2) \neq 0$ are easy consequences of mollification. To check (3) we proceed as follows. Note from Lemma A.1.6 that

$$f'' < \frac{\rho_0 - (f')^2}{f} \implies f(t)^2 g + dt^2 \text{ has positive scalar curvature.} \tag{2.2.7}$$

For a function $\varphi : \operatorname{dmn} \varphi \to (0, \infty)$, write

$$\mathcal{C}[\varphi] \triangleq \left\{ \left(t, \frac{\rho_0 - \varphi'(t)^2}{\varphi(t)} \right) : t \in \operatorname{dmn} \varphi' \right\} \subseteq \operatorname{dmn} \varphi \times \mathbf{R},$$

and

$$\mathcal{T}[\varphi] \triangleq \{(t, \varphi''(t)) : t \in \operatorname{dmn} \varphi''\} \subseteq \operatorname{dmn} \varphi \times \mathbf{R}.$$

Conclusion (3) follows from (2.2.7) provided, for small $\sigma > 0$, the curve $\mathcal{T}[F_\sigma]$ lies strictly below the curve $\mathcal{C}[F_\sigma]$ on the strip $[a_1, b_2] \times \mathbf{R}$. (Note: "$\mathcal{C}$" indicates positive scalar curvature "constraints" on second derivatives, while "$\mathcal{T}$" denotes "true" second derivatives.) By hypothesis (2),

$$\mathcal{C}[f_1] \cup \mathcal{C}[f_2] \text{ lies strictly above } \mathcal{T}[f_1] \cup \mathcal{T}[f_2]. \tag{2.2.8}$$

By (H4), the curve $\Gamma_{\mathcal{C}}$ traced by the graph of

$$[b_1, a_2] \ni x \mapsto \frac{\rho_0 - f_1'(b_1)^2}{F(x)}$$

joins the right endpoint of $\mathcal{C}[f_1]$ to the left endpoint of $\mathcal{C}[f_2]$; notice that it is convex decreasing on this interval and, therefore, does not dip below its lowest point at $x = a_2$. Thus, by (2.2.8) and (H1), $\mathcal{C}[F] = \mathcal{C}[f_1] \cup \Gamma_{\mathcal{C}} \cup \mathcal{C}[f_2]$ is a connected curve,

$$\widetilde{\mathcal{T}}[F] \triangleq \mathcal{T}[F] \cup (\{b_1\} \times [0, f_1''(b_1)]) \cup (\{a_2\} \times [0, f_2''(a_2)])$$

is, too, and

$$\mathcal{C}[F] \text{ lies strictly above } \widetilde{\mathcal{T}}[F] \text{ in the strip } [a_1, b_2] \times \mathbf{R}. \tag{2.2.9}$$

Using $F \in C^{1,1}[a_1, b_2] \cap C^2([a_1, b_2] \setminus \{b_1, a_2\})$ and the fact that mollification commutes with weak derivatives, it's easy to see that

$$\lim_{\sigma \downarrow 0} \mathcal{C}[F_\sigma] = \mathcal{C}[F] \text{ and } \lim_{\sigma \downarrow 0} \mathcal{T}[F_\sigma] = \mathcal{T}[F] \cup (\{b_1\} \times [0, f_1''(b_1)]) \cup (\{a_2\} \times [0, f_2''(a_2)]) \tag{2.2.10}$$

in the Hausdorff topology. From (2.2.9) and (2.2.10) we therefore see that $\mathcal{C}[F_\sigma]$ lies strictly above $\mathcal{T}[F_\sigma]$ for small $\sigma > 0$, and conclusion (3) follows. $\square$

The previous claim concludes the proof of the lemma. $\square$

Proposition 2.2.9 and Lemma 2.2.10 give two important corollaries:

Corollary 2.2.12 (PSC cobordance, M.)**.** *For every $(\mathbf{S}^2, \gamma_i) \in \mathscr{M}_+$, $i \in \{1, 2\}$, there exists $(\mathbf{S}^2 \times [0,1], g)$ with positive scalar curvature and totally geodesic boundary $\partial(\mathbf{S}^2 \times [0,1], g) \cong \sqcup_{i=1,2}(\mathbf{S}^2, \gamma_i)$.*

Corollary 2.2.13 (PSC null-cobordance, M.)**.** *For every $(\mathbf{S}^2, \gamma) \in \mathscr{M}_+$ there exists a disk $(\mathbf{D}^3, g)$ with positive scalar curvature and totally geodesic boundary $\partial(\mathbf{D}^3, g) \cong (\mathbf{S}^2, \gamma)$.*

2.3 The Λ-invariant of two-dimensional surfaces

2.3.1 Definitions

The following lemma, in various forms, is well-known to scalar curvature experts. We reproduce it here because it is not readily available in the literature in a convenient form.

Lemma 2.3.1 (M.)**.** *Let (Ω^n, g), $n \geq 3$, be a compact Riemannian n-manifold, with nonnegative scalar curvature. Suppose $\partial\Omega = \partial_D\Omega \sqcup \partial_N\Omega$ is a partition of the connected components of $\partial\Omega$ into two disjoint families. Suppose that u is a nonconstant harmonic function on Ω, $u \equiv \beta_j > 0$ on every component $\Sigma_j \subseteq \partial_D\Omega$, and $\nabla_\nu u \equiv 0$ on $\partial_N\Omega$, where ν denotes the outward pointing unit normal to $\partial\Omega$. Then the following are true:*

(1) $u > 0$ in Ω and $\widetilde{g} = u^{\frac{4}{n-2}} g$ is a smooth Riemannian metric with nonnegative scalar curvature,

(2) if $\Sigma \subseteq \partial_D\Omega$ is such that $u|_\Sigma = \beta \equiv \max_{\partial_D\Omega} u$, then $\beta^{\frac{2}{n-2}} H_{\Sigma,\widetilde{g}} - H_{\Sigma,g} = c(n)\beta^{\frac{4}{n-2}} \nabla^g_\nu u > 0$.

(3) if $\Sigma \subseteq \partial_D\Omega$ is such that $u|_\Sigma = \beta \equiv \min_{\partial_D\Omega} u$, then $\beta^{\frac{2}{n-2}} H_{\Sigma,\widetilde{g}} - H_{\Sigma,g} = c(n)\beta^{\frac{4}{n-2}} \nabla^g_\nu u < 0$.

(4) if $\Sigma \subseteq \partial_N\Omega$, then $H_{\Sigma,\widetilde{g}} = u^{-\frac{2}{n-2}} H_{\Sigma,g}$.

If, instead of being harmonic, u is a nonconstant solution of

$$\Delta_g u = \frac{n-2}{4(n-1)} R_g u \text{ in } \Omega,$$

then conclusions (2) and (4) continue to hold, while (1) is sharpened to

(1') $u > 0$ and $\widetilde{g} = u^{\frac{4}{n-2}}$ is a smooth scalar-flat Riemannian metric.

Proof. If $u > 0$ were false, then u would attain a nonpositive minimum at some $p \in \Omega$. By the maximum principle (Theorem B.1.1), $p \in \partial\Omega$. On the other hand, $u > 0$ on $\partial_D\Omega$ by hypothesis

(2), so $p \in \partial_N \Omega$. By the Hopf boundary point lemma (Lemma B.1.2), $\nabla^g_\nu u(p) < 0$. This, however, contradicts hypothesis (3). This proves $u > 0$. That scalar curvature is nonnegative is an immediate consequence of Lemma A.1.7, which describes changes of scalar curvature under conformal transformations.

The identity in (2) is also an immediate consequence of Lemma A.1.7. Let's prove the inequality. By the maximum principle, u must attain a maximum on $\partial\Omega$. By the Hopf boundary point lemma (Lemma B.1.2) and hypothesis (3), this maximum cannot occur on $\partial_N \Omega$, so $\max_\Omega u = \max_{\partial_D \Omega} u$. Therefore, the inequality in (2) is a consequence of the Hopf boundary point lemma.

(3) follows analogously.

(4) is an immediate consequence of Lemma A.1.7, which describes changes of scalar curvature under conformal transformations. ☐

One interpretation of this lemma is that scalar-curvature-aware conformal deformations in compact manifolds with boundary have inverse effects on the interior scalar curvature and on the boundary mean curvature.

Question. *If we fix a lower bound for scalar curvature in the interior, can we get an a priori L^1 estimate on the boundary mean curvature, which depends only on the intrinsic geometry of the boundary?*

P. Miao and I have begun the study of this question for compact, mean-convex 3-manifolds with nonnegative scalar curvature and a prescribed boundary metric in [39]. We will show in Section 3.5 how all this occurs naturally in a relativistic setting.

We first introduce the relevant definitions before stating the main results.

Definition 2.3.2 (Fill-ins, M.-Miao [39, Definition 1.1]). *Let (Σ_j^2, γ_j), $j = 1, \ldots, k$, $k \geq 1$, be closed, orientable Riemannian manifolds. Write $\mathcal{F}_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}}$ for the set of compact Riemannian manifolds (Ω^3, g) such that:*

(1) $(\partial\Omega, g|_{\partial\Omega}) \cong \sqcup_{j=1}^k (\Sigma_j, \gamma_j)$,

(2) $\partial\Omega^3$ is mean-convex; i.e., the mean curvature vector of $\partial\Omega^3$ points inward, and

(3) the scalar curvature of (Ω^3, g) is nonnegative.

We also define the Λ-invariant of this family of two-dimensional manifolds as

$$\Lambda_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}} \triangleq \sup \left\{ \int_{\partial\Omega} H_g \, d\sigma : (\Omega, g) \in \mathcal{F}_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}} \right\}.$$

When $k = 1$, we will simply write $\mathcal{F}_{(\Sigma,\gamma)}$, $\Lambda_{(\Sigma,\gamma)}$ instead of $\mathcal{F}_{\{(\Sigma,\gamma)\}}$, $\Lambda_{\{(\Sigma,\gamma)\}}$.

Below we collect known properties of $\Lambda_{(\Sigma,\gamma)}$ in the relativistically important case of spheres that will be proven later.

Theorem 2.3.3 (M.-Miao [38, Proposition 2.1])**.** *For any (Σ^2, γ), with Σ^2 diffeomorphic to $\mathbf{S}^2$, the following are true:*

(1) $\Lambda_{(\Sigma,\gamma)}$ is finite whenever $\mathcal{F}_{(\Sigma,\gamma)}$ is nonempty.

(2) If $\Lambda_{(\Sigma,\gamma)}$ is attained by some $(\Omega^3, g) \in \mathcal{F}_{(\Sigma,\gamma)}$, i.e., if

$$\Lambda_{(\Sigma,\gamma)} = \int_{\partial\Omega} H_g \, d\sigma,$$

then Ω^3 is a 3-ball and g is flat.

(3) If (Σ^2, γ) is isometric to a strictly convex surface Σ_0^2 in $\mathbf{R}^3$ with mean curvature H_0, then

$$\Lambda_{(\Sigma,\gamma)} = \int_{\iota(\Sigma)} H_0 \, d\sigma,$$

where $\iota : \Sigma \to \Sigma_0 \subset \mathbf{R}^3$ is the isometric embedding.

(4) The value $\Lambda_{(\Sigma,\gamma)}$ can be estimated from below if (Σ^2, γ) is isometric to a mean-convex, outer-minimizing sphere in $\mathbf{R}^3$. In this case,

$$\Lambda_{(\Sigma,\gamma)} \geq \sqrt{16\pi\sigma_\gamma(\Sigma)}.$$

Moreover, equality holds if and only if (Σ^2, γ) is isometric to a round sphere.

Remark 2.3.4. *The supremum of an empty set is conventionally $-\infty$. In the case of 2-spheres, it is easily seen that $\mathcal{F}_{(\Sigma,\gamma)}$ is nonempty if (Σ, γ) has positive Gauss curvature. In fact, it is also the case that $\mathcal{F}_{(\Sigma,\gamma)}$ is nonempty for certain metrics γ on 2-spheres Σ with arbitrarily negative portions (in the L^1 sense) of curvature; indeed, this is a consequence of Theorem 2.2.6 and Corollary 2.2.13. It is not known if there exist 2-spheres with $\mathcal{F}_{(\Sigma^2,\gamma)} = \emptyset$.*

Note that (1) is an a priori L^1 bound on the boundary mean curvature of mean-convex (Ω, g) with spherical boundary and nonnegative scalar curvature; it depends only on the induced intrinsic boundary geometry. Such an a priori estimate is crucial to the definition and study of $\mathfrak{m}(\partial\Omega, \Omega, g)$ in Section 3.5. Its proof is particularly interesting, and perhaps unconventional in the study of nonnegative scalar curvature, in that it requires the study of asymptotically hyperbolic—not Euclidean—manifolds. It requires deep results in mathematical general relativity due to Wang-Yau [72] and Shi-Tam [64, 66], who sought to understand the boundary mean curvature of compact manifolds whose scalar curvature has a negative lower bound. Using that work, the L^1 estimate was derived in [39] for Σ consisting of one or more 2-spheres. The same estimate, for a single 2-sphere, was derived independently by Lu [37]; in Lu's work it serves as a geometric inequality that allows for a priori estimates of geometric quantities via elliptic partial differential equation techniques.

Let us now return to the motivation of studying $\mathcal{F}$ and the Λ-invariant, which was the inverse relationship between nonnegative interior scalar curvature and boundary mean curvature. We begin by showing that, on any fill-in $(\Omega^3, g) \in \mathcal{F}_{(\Sigma^2,\gamma)}$, each component of (Σ^2, γ) contributes independently to maximizing $\|H\|_{L^1}$, agnostically of other components.

Theorem 2.3.5 (Λ-additivity and rigidity, M.-Miao [39, Theorem 1.2])**.** *Let (Σ_j^2, γ_j), $j = 1, \ldots, k$, $k \geq 1$, be closed, orientable. For all $(\Omega^3, g) \in \mathcal{F}_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}}$ and for every $j = 1, \ldots, k$,*

$$\int_{\Sigma_j} H_{g,j} \, d\sigma_j \leq \Lambda_{(\Sigma_j,\gamma_j)}. \tag{2.3.1}$$

If there exists $(\Omega, g) \in \mathcal{F}_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}}$ for which (2.3.1) is an equality for some $j \in \{1, \ldots, k\}$, then $\partial\Omega$ has a unique boundary component (i.e., $k = 1$) and (Ω^3, g) is isometric to a mean-convex handlebody with flat interior whose genus is that of Σ_1 (since $k = 1$). In particular, if $\operatorname{genus}(\Sigma_1) = 0$ *then (Ω^3, g) is an Alexandrov embedded mean-convex ball in $\mathbf{R}^3$.*

Moreover, the Λ-invariant is additive across components in the sense that

$$\Lambda_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}} = \sum_{j=1}^{k} \Lambda_{(\Sigma_j,\gamma_j)}, \tag{2.3.2}$$

provided $\mathcal{F}_{(\Sigma_j,\gamma_j)}$, $j = 1, \ldots, k$, and $\mathcal{F}_{(\Sigma_1,\gamma_1),\ldots,(\Sigma_k,\gamma_k)}$ are all nonempty.

For this theorem we had to introduce a geometric "cut-and-fill" operation which plays a dual role to the Schoen-Yau [61] and Gromov-Lawson [28] positive scalar curvature connect-sum: cut along essential, stable 2-spheres and fill in the holes with positive scalar curvature disks using techniques from the work in Section 2.2 on the moduli space $\mathscr{M}_+$.

Remark 2.3.6. *In the course of the proof of Theorem 2.3.5 we will see that*

$$\mathcal{F}_{(\Sigma_1,\gamma_1),\ldots,(\Sigma_k,\gamma_k)} \text{ is empty} \iff \text{some } \mathcal{F}_{(\Sigma_j,\gamma_j)},\ j = 1, \ldots, k, \text{ is empty.}$$

Notice that, certainly, (2.3.1) is an equality for all $j = 1, \ldots, k$ if

$$\int_{\partial\Omega} H_g \, d\sigma = \Lambda_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}},$$

i.e., if (Ω^3, g) attains the supremum $\Lambda_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}}$. In other words, Theorem 2.3.5 confirms that a disconnected boundary cannot support a maximizing configuration.

2.3.2 Results

The class $\mathcal{F}$ of fill-ins and the Λ-invariant were introduced in [39] alongside another two objects:

(1) a class $\mathring{\mathcal{F}}$ of fill-ins for a single prescribed boundary component that allows for the existence of additional, unprescribed minimal boundary components of arbitrary topology, and

(2) the corresponding $\mathring{\Lambda}$-invariant.

Elements in $\mathring{\mathcal{F}}_{(\Sigma,\gamma)}$ are not only more suitable for discussion on quasi-local mass, but they can also serve as building blocks, effectively allowing to deduce $\mathcal{F}_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}}$-related results from them via a cut-and-fill technique: cutting composite manifolds across carefully chosen minimal surfaces, and, when necessary, fill in holes with 3-disks of positive scalar curvature.

We present a similar approach—but not identical—to [39] here. We will succinctly define and use $\mathring{\mathcal{F}}$ below, as it helps simplify the statements of the technical lemmas, but we will not introduce or study $\mathring{\Lambda}$ until Section 3.5, where we discuss quasi-local mass.

Definition 2.3.7 (Fill-ins, II, M.-Miao [39, Definition 1.2])**.** *Let (Σ^2, γ) be a closed, orientable Riemannian manifold. Denote by $\mathring{\mathcal{F}}_{(\Sigma,\gamma)}$ the set of all compact Riemannian manifolds (Ω^3, g) such that:*

(1) $\partial\Omega^3$ has a connected component Σ_O, which, with the induced metric, is isometric to (Σ^2, γ),

(2) Σ_O is mean-convex,

(3) $\Sigma_H \triangleq \partial\Omega \setminus \Sigma_O$, if nonempty, is a minimal surface, possibly disconnected, and

(4) the scalar curvature of (Ω, g) is nonnegative.

We now start building the technical toolkit that will go into the proof of Theorems 2.3.3 and 2.3.5.

Lemma 2.3.8 (Cutting, M.-Miao [39, Lemma 2.1])**.** *Let $\{(\Sigma_j^2, \gamma_j)\}_{j=1,\ldots,k}$ be $k \geq 2$ closed and orientable. Let $(\Omega^3, g) \in \mathcal{F}_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}}$. For every $j \in \{1, \ldots, k\}$ there exists an orientable $(\Omega_j^3, g_j) \in \mathring{\mathcal{F}}_{(\Sigma_j,\gamma_j)}$ such that*

(1) the mean curvature of Σ_j in (Ω, g) is the same as that of $\Sigma_O^{(j)}$ in (Ω_j, g_j),

(2) $\Sigma_H^{(j)} \triangleq \partial\Omega_j \setminus S_O^{(j)}$ is nonempty and consists of stable, orientable minimal surfaces,

(3) (Ω_j^3, g_j) has nonnegative scalar curvature, which is positive if that of (Ω^3, g) is.

Proof. First suppose Ω is orientable. Write S_j for the boundary component of Ω corresponding to our fixed Σ_j. Since the mean curvature vector points inward on $\partial\Omega$, there exists a smooth, oriented minimal surface S in the interior of Ω (possibly covered with multiplicity) that minimizes area in the integral homology class of S_j. (Specifically, we minimize area in the sense of integer multiplicity currents.) Denote (Ω_j^3, g_j) the metric completion of the component of $\Omega \setminus S$ containing S_j. Note that Ω_j is, by construction, orientable, and $(\Omega_j^3, g_j) \in \mathring{\mathcal{F}}_{(\Sigma_j,\gamma_j)}$ manifestly satisfies (1), (2), (3).

If Ω is nonorientable, consider its orientation double cover $\pi : \widetilde{\Omega} \implies \Omega$. Using the fact that S_j itself is orientable, one can show that $\pi^{-1}(S_j)$ is a disjoint union of two copies of S_j. Denote one of them by $\widetilde{S}_j$. The lemma follows by repeating the argument in the previous paragraph with (Ω, g) and S_j replaced by $(\widetilde{\Omega}, \pi^*(g))$ and $\widetilde{S}_j$, respectively. $\square$

Lemma 2.3.9 (Filling, M.-Miao [39, Lemma 2.2]). *Let (Σ^2, γ) be closed and orientable. If $(\Omega^3, g) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)}$ is such that*

(1) $\Sigma_H \triangleq \partial\Omega \setminus \Sigma_O$ is nonempty and consists of stable minimal $\mathbf{S}^2$s, and

(2) the ambient scalar curvature is strictly positive along $\Sigma_H \subseteq \partial\Omega$.

For every $\eta > 0$, there exists $(D^3, h) \in \mathcal{F}_{(\Sigma,\gamma)}$, orientable if Ω is, with $H_{\partial D,h} > H_{\Sigma_O,g} - \eta$.

Proof. Recall Lemma A.1.6. From hypotheses (1), (2), and Lemma 2.2.1, we see directly that on every stable minimal 2-sphere $T_\ell \subseteq \Sigma_H$:

$$\lambda_1(-\Delta_\ell + K_\ell) > 0, \text{ i.e., } T_\ell \in \mathscr{M}_+. \tag{2.3.3}$$

Here Δ_ℓ and K_ℓ denote the Laplace-Beltrami operator and the Gauss curvature on the 2-sphere T_ℓ. By Corollary 2.2.13, we can glue 3-disks of positive scalar curvature onto each T_ℓ to obtain a compact $C^{0,1}$ manifold D^3 with $\partial D \cong \Sigma_O$, and which is smooth away from $\cup_\ell T_\ell$. If the resulting metric $\widetilde{g}$ on D^3 were smooth across $\cup_\ell T_\ell$, then $(D^3, \widetilde{g}) \in \mathcal{F}_{(\Sigma,\gamma)}$. In general $\widetilde{g}$ will not be smooth across $\cup_\ell T_\ell$, so we apply [42, Proposition 3.1] to $(D^3, \widetilde{g})$ followed by a small conformal deformation to obtain another metric h on D^3 such that $(D^3, h) \in \mathcal{F}_{(\Sigma,\gamma)}$ and $H_{\partial D,h} \geq H_{\partial D,\widetilde{g}} - \eta/2 \geq H_{\Sigma_O,g} - \eta$. $\square$

Lemma 2.3.10 (Doubling, M.-Miao [39, Lemma 2.3]). *Let (Σ^2, γ) be closed and orientable, and $(\Omega^3, g) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)}$ be such that $\Sigma_H \triangleq \partial\Omega \setminus \Sigma_O$ is nonempty. Let D^3 denote the doubling of Ω across Σ_H, with $\partial D = S_O \cup S_O'$, where S_O corresponds to $\Sigma_O \subseteq \partial\Omega$ and S_O' corresponds to its mirror image under the doubling. For every $\eta > 0$ there exists a Riemannian metric h on D^3 such that:*

(1) S_O, with the induced metric from h, is isometric to (Σ^2, γ),

(2) $H_{S_O,h} > H_{\Sigma_O,g}$,

(3) $H_{S'_O,h} > H_{\Sigma_O,g} - \eta$, *and*

(4) (D^3, h) *is scalar-flat.*

Proof. Let's first reduce to the case of $R_g \equiv 0$. If $R_g \equiv 0$ to begin with, set $\widetilde{g} = g$. Else, if $R_g \geq 0$ but $R_g \not\equiv 0$, employ Lemma 2.3.1 with a solution u of

$$\Delta_g u - \frac{1}{8} R_g u = 0 \text{ in } \Omega,\ u = 1 \text{ on } \Sigma_O,\ \nabla^g_\nu u = 0 \text{ on } \Sigma_H,$$

to obtain $\widetilde{g}$ with $H_{\Sigma_H,\widetilde{g}} = 0$ and $H_{\Sigma_O,\widetilde{g}} > H_{\Sigma_O,g}$. Now replace (Ω^3, g) with $(\Omega^3, \widetilde{g})$.

In any case, $H_{\Sigma_O,\widetilde{g}} \geq H_{\Sigma_O,g}$ and $H_{\Sigma_H,\widetilde{g}} = 0$. Given a small $\varepsilon \in (0,1)$, let $\phi_1 \in C^\infty(\Omega)$ be the unique function such that

$$\Delta_{\widetilde{g}} \phi_1 = 0,\ \phi = 1 \text{ on } \Sigma_O,\ \phi_1 = 1 - \frac{1}{2}\varepsilon \text{ on } \Sigma_H.$$

Let $\phi_2 = (2-\varepsilon) - \phi_1$. Employ Lemma 2.3.1 and consider $g_1 = \phi_1^4 \widetilde{g}$ and $g_2 = \phi_2^4 \widetilde{g}$. These two metrics satisfy the following properties:

(P1) the induced metrics on Σ_H from g_1 and g_2 agree,

(P2) the mean curvature of Σ_H in (Ω, g_1) with respect to the inward normal agrees with the mean curvature of Σ_H in (Ω, g_2) with respect to the outward normal,

(P3) $H_{\Sigma_O,g_1} > H_{\Sigma_O,\widetilde{g}}$ on Σ_O by conclusion (2) of the lemma, and

(P4) as long as $\varepsilon > 0$ is small, $H_{\Sigma_O,g_2} > H_{\Sigma_O,\widetilde{g}} - \eta/2 > 0$.

Now attach (Ω, g_1) and (Ω, g_2) along Σ_H, and call the resulting manifold (D, h). In (D, h), denote the Σ_O coming from (Ω, g_1) by S_O and the one coming from (Ω, g_2) by S'_O. If the metric h were smooth across Σ_H, then it satisfies all the properties required. In general, we can replace h with another metric that is obtained by applying [42, Proposition 3.1] to (D, h) at Σ_H followed by a small conformal deformation that fixes the boundary and decreases the boundary mean curvature by a small amount, so that $H_{S_O,h} > H_{\Sigma_O,\widetilde{g}}$ and $H_{S'_O,h} > 0$. The result follows. □

We are now in a position to prove Theorem 2.3.5.

Proof of Theorem 2.3.5 (Λ-additivity and rigidity). We will first prove inequality (2.3.1) and identity (2.3.2), as those are obviously only nontrivial when $k \geq 2$. We will prove the rigidity case of (2.3.1) at the end.

Let $(\Omega, g) \in \mathcal{F}_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}}$, with $k \geq 2$. For each $j = 1, \ldots, k$, denote the boundary component in $\partial\Omega$ corresponding to Σ_j by S_j.

Let $\phi < 0$ be a fixed function on Ω and let w be the unique solution to

$$\Delta_g w - \frac{1}{8} R_g w = \phi \text{ in } \Omega \text{ and } w = 0 \text{ on } \partial\Omega.$$

Let $\varepsilon > 0$ be given. For small $\tau > 0$, consider the metrics $g^{(\tau)} \triangleq (1+\tau w)^4 g$, all of which have strictly positive scalar curvature in Ω by Lemma A.1.7. Moreover, from the same lemma, the boundary $\partial\Omega$ has mean curvature

$$H_{g^{(\tau)}} = H_g + 4\nabla^g(1+\tau w) = H_g + 4\tau\nabla^g w;$$

which is positive and

$$\int_{S_j} H_{g^{(\tau)},j}\, d\sigma_j \geq \int_{S_j} H_{g,j} d\sigma_j - \varepsilon$$

if τ is small enough. Pick one such τ and write $\widetilde{g}$ for this $g^{(\tau)}$.

Employing Lemma 2.3.8, cut $(\Omega, \widetilde{g})$ to isolate the boundary component S_j and obtain $(\Omega_j, \widetilde{g}_j)$ as in the statement of the lemma, with $R_{\widetilde{g}_j} > 0$, $H_{\widetilde{g}_j} = H_{\widetilde{g},j}$ on S_j and $\partial\Omega_j = S_j \cup T_j$, with T_j a nonempty union of smooth, stable, oriented minimal surfaces. Moreover it follows from $R_{\widetilde{g}_j} > 0$ and Proposition A.2.7 that T_j consists of 2-spheres. Next, employ the filling lemma (Lemma 2.3.9) once for every sphere in T_j to replace $(\Omega_j, \widetilde{g}_j)$ with $(M_j, h_j) \in \mathcal{F}_{(\Sigma_j,\gamma_j)}$, with the mean curvature H_{h_j} of S_j in (M_j, h_j) satisfying

$$\int_{S_j} H_{h_j}\, d\sigma_j \geq \int_{S_j} H_{\widetilde{g}_j}\, d\sigma_j - \varepsilon \geq \int_{S_j} H_{g,j} d\sigma_j - 2\varepsilon.$$

Since $(M_j, h_j) \in \mathcal{F}_{(\Sigma_j,\gamma_j)}$, the left hand side is bounded from above by $\Lambda_{(\Sigma_j,\gamma_j)}$. Rearranging,

$$\int_{S_j} H_{g,j}\, d\sigma_j \leq 2\varepsilon + \Lambda_{(\Sigma_j,\gamma_j)}.$$

Letting $\varepsilon \downarrow 0$, we obtain (2.3.1).

Notice that if we had not let $\varepsilon \downarrow 0$, and instead carried out the procedure above for all $j = 1, \ldots, k$, and summed over j, then

$$\int_{\partial\Omega} H_g\, d\sigma \leq 2k\varepsilon + \sum_{j=1}^{k} \Lambda_{(\Sigma_j,\gamma_j)}.$$

Letting $\varepsilon \downarrow 0$ now and recalling that $(\Omega, g) \in \mathcal{F}_{\{(\Sigma_i,\gamma_i)\}_{i=1,\ldots,k}}$ was arbitrary, we conclude

$$\Lambda_{\{(\Sigma_j,\gamma_j)\}_{j=1,\ldots,k}} \leq \sum_{j=1}^{k} \Lambda_{(\Sigma_j,\gamma_j)},$$

i.e., the "$\leq$" direction in (2.3.2).

It remains to check the reverse direction "$\geq$" in (2.3.2). We will assume that all quantities on the right are finite, as a similar argument carries through to the general case. Let $\varepsilon > 0$ be given. For each $j = 1, \dots, k$, let $(\Omega_j, g_j) \in \mathcal{F}_{(\Sigma_j, \gamma_j)}$ be such that

$$\int_{\partial\Omega_j} H_{g_j}\, d\sigma_j \geq \Lambda_{(\Sigma_j,\gamma_j)} - \varepsilon.$$

On each Ω_j, let $\phi_j < 0$ be a fixed function and let w_j be the unique solution to

$$\Delta_{g_j} w_j - \frac{1}{8} R_{g_j} w_j = \phi_j \text{ in } \Omega_j \text{ and } w_j = 0 \text{ on } \partial\Omega_j.$$

For small $\tau > 0$, consider the metrics $g_j^{(\tau)} = (1 + \tau w_j)^4 g_j$. Then $g_j^{(\tau)}$ has strictly positive scalar curvature on Ω_j, $\partial\Omega_j$ has positive mean curvature $H_{g_j^{(\tau)}}$ in $(\Omega_j, g_j^{(\tau)})$, and

$$\int_{\partial\Omega_j} H_{g_j^{(\tau)}}\, d\sigma_j \geq \Lambda_{(\Sigma_j,\gamma_j)} - 2\varepsilon$$

if τ is small enough. Applying the connect-sum construction for positive scalar curvature manifolds (cf. [61] or [28]), we obtain $\Omega = \Omega_1 \# \dots \# \Omega_k$ endowed with a metric g of positive scalar curvature that coincides with $g_j^{(\tau)}$ near each $\partial\Omega_j$. In particular, $(\Omega, g) \in \mathcal{F}_{\{(\Sigma_j,\gamma_j)\}_{j=1,\dots,k}}$ and it satisfies

$$\int_{\partial\Omega} H_g\, d\sigma \geq \sum_{j=1}^{k} \Lambda_{(\Sigma_j,\gamma_j)} - 2k\varepsilon.$$

Identity (2.3.2) follows by sending $\varepsilon \downarrow 0$.

Now, we turn to prove the rigidity case of (2.3.1).

First let us deal with the base case $k = 1$, which is the most difficult.

We claim $R_g \equiv 0$. If $R_g \not\equiv 0$, applying Lemma 2.3.1 with the unique solution to

$$\Delta_g u - \frac{1}{8} R_g u = 0 \text{ in } \Omega \text{ and } u = 1 \text{ on } \partial\Omega.$$

and considering the conformally deformed metric $\hat{g} = u^4 g$, we see that $(\Omega, \hat{g}) \in \mathcal{F}_{(\Sigma,\gamma)}$ and, by conclusion (2) of the lemma,

$$\Lambda_{(\Sigma,\gamma)} = \int_{\partial\Omega} H_g\, d\sigma < \int_{\partial\Omega} H_{\hat{g}}\, d\sigma \leq \Lambda_{(\Sigma,\gamma)},$$

a contradiction. Therefore, u is constant, which shows $R_g \equiv 0$ in Ω.

Now consider the space of metrics $\bar{g}$ on Ω given by

$$\mathrm{Met}_g^0(\Sigma) \triangleq \{\bar{g} \in \mathrm{Met}(\Sigma) : R_{\bar{g}} = 0 \text{ and } \bar{g}|_{T\partial\Omega} = g|_{T\partial\Omega}\},$$

where $\bar{g}|_{T\partial\Omega}$ denotes the induced metric on $\partial\Omega$ from $\bar{g}$. Since $g \in \mathrm{Met}_g^0(\Sigma)$ maximizes the total boundary mean curvature in $\mathcal{F}$, one knows that g is a critical point of the smooth functional

$$\mathrm{Met}_g^0(\Sigma) \ni \bar{g} \mapsto \int_{\partial\Omega} H_{\bar{g}} \, d\sigma.$$

From [44, Corollary 2.1] it follows that g must be Ricci flat, and hence flat (Ω is three-dimensional).

It remains for us to check the topological conclusion. [41, Theorem 1, Proposition 1] tells us that in the absence of interior minimal surfaces, Ω is necessarily a handlebody with mean-convex boundary. Therefore, it remains to show that there are no closed embedded minimal surfaces (oriented or not) in the interior of Ω. We proceed by contradiction.

If there were such a minimal surface T then by compactness there would exist an interior smooth geodesic $\Gamma : [0, \ell] \to \Omega$ joining a pair of closest points between T and $\partial\Omega$; $\Gamma(0) \in T$, $\Gamma'(0) \perp T$, $\Gamma(\ell) \in \partial\Omega$, and $\Gamma'(\ell) \perp \partial\Omega$. The second variation of the length of this geodesic summed among a basis of two unit normal variations V_i, $i = 1, 2$, is

$$\begin{aligned}\sum_{i=1}^{2} \delta^2\Gamma(V, V) &= -\int_0^\ell \mathrm{Ric}_g(\Gamma'(s), \Gamma'(s)) ds - H_{\partial\Omega, g}(\Gamma(\ell)) - H_{T,g}(\Gamma(0)) \\ &= -H_{\partial\Omega, g}(\Gamma(\ell)) < 0,\end{aligned}$$

since g is flat and T is minimal. This means Γ is unstable, in contradiction with its minimizing nature. The topological claim follows.

Finally, when $\mathrm{genus}(\Sigma) = 0$ then we know Ω a genus-0 handlebody, i.e., a 3-ball. We've shown its metric g is flat, so we can locally (and therefore globally since it is simply connected) immerse Ω in $\mathbf{R}^3$. This completes the case $k = 1$ of the theorem.

Now suppose $k \geq 2$. Our assumption is that

$$\int_{S_j} H_{g,j} \, d\sigma_j = \Lambda_{(\Sigma_j, \gamma_j)}$$

holds for some fixed $j \in \{1, \ldots, k\}$, where S_j represents the boundary component corresponding to the Σ_j on which we have equality. Employ the cutting lemma (Lemma 2.3.8) to isolate the boundary component S_j and obtain (Ω_j, g_j) as in the lemma, followed the doubling lemma (Lemma 2.3.10) to double Ω_j across $\partial\Omega_j \setminus S_j$ and obtain (D_j, h_j). Writing S_j' for the mirror image of S_j

under the doubling, we have $\partial D_j = S_j \cup S_j'$. By the conclusions of the doubling lemma, $(D_j, h_j) \in \mathcal{F}_{\{(\Sigma_j,\gamma_j),(\Sigma_j,\gamma_j')\}}$ for some metric γ_j', as induced by h, so by inequality (2.3.1),

$$\int_{S_j} H_{h_j}\, d\sigma_j \leq \Lambda_{(\Sigma_j,\gamma_j)}.$$

By conclusion (2) of the doubling lemma, the mean curvature H_{h_j} of S_j in (D_j, h_j) exceeds the original mean curvature $H_{g,j}$, so,

$$\Lambda_{(\Sigma_j,\gamma_j)} = \int_{S_j} H_{g,j}\, d\sigma_j < \int_{S_j} H_{h_j}\, d\sigma_j \leq \Lambda_{(\Sigma_j,\gamma_j)},$$

a contradiction. □

Finally, we set to prove Theorem 2.3.3 regarding the Λ-invariant on 2-spheres.

Proof of Theorem 2.3.3. (1) We establish an a priori L^1 estimate on the boundary mean curvature for all elements $(\Omega^3, g) \in \mathcal{F}_{(\Sigma,\gamma)}$. We make use of the following result of Shi and Tam in [66], which is built on the work on Wang and Yau in [72]:

Theorem (Shi-Tam [66, Theorem 3.8])**.** *Let (Ω^3, g) be compact, orientable, with boundary Σ, and scalar curvature $R_g \geq -6\kappa^2$ for some constant $\kappa > 0$. Suppose Σ^2 is a 2-sphere with Gauss curvature $K > -\kappa^2$ and positive mean curvature H. Let $\iota : \Sigma^2 \hookrightarrow \mathbf{H}^3_{-\kappa^2}$ be an isometric embedding and denote $\Sigma_0 \triangleq \iota(\Sigma)$, which is a convex surface in $\mathbf{H}^3_{-\kappa^2}$. Let $D \subset \mathbf{H}^3_{-\kappa^2}$ be the bounded region enclosed by Σ_0. Then for any $p \in D$,*

$$\int_{\Sigma} H \cosh \kappa r(p, \iota(z))\, d\sigma(z) \leq \int_{\Sigma_0} H_0 \cosh \kappa r(p, y)\, d\sigma_0(y), \tag{2.3.4}$$

where $r(p, \cdot)$ denotes the distance to p in $\mathbf{H}^3_{-\kappa^2}$, $d\sigma_0$ is the area element on Σ_0, and H_0 is the mean curvature of Σ_0 in $\mathbf{H}^3_{-\kappa^2}$. Moreover, equality in (2.3.4) holds if and only if (Ω, g) is isometric to the domain D bounded by Σ_0 in $\mathbf{H}^3_{-\kappa^2}$.

Remark 2.3.11. *The existence of an isometric embedding $\iota : \Sigma \to \mathbf{H}^3_{-\kappa^2}$, when $K > -\kappa^2$, is guaranteed by Pogorelov's embedding theorem [56, Theorem 1].*

Even though not explicitly stated, the proof of [66, Theorem 3.8] assumes Ω is orientable since the proof uses the fact that Ω is a spin manifold.

Let us assume, for now, that Ω is orientable and, therefore, spin (see [33, Chapter VII, Theorem 1]). Choose $\kappa > 0$ so that the Gauss curvature K_γ of (Σ^2, γ) satisfies $K_\gamma > -\kappa^2$. Clearly, $R_g \geq 0 \geq -6\kappa^2$

on Ω^3. Therefore, [66, Theorem 3.8] quoted above applies to (Ω^3, g) to assert

$$\int_{\partial\Omega} H_g \, d\sigma \leq \int_{\partial\Omega} H_g \cosh \kappa r \, d\sigma \leq \int_{\iota(\Sigma)} H_0 \cosh \kappa r \, d\sigma,$$

where H_0 is the mean curvature of the isometric embedding of $\iota : (\Sigma, \gamma) \hookrightarrow \mathbf{H}^3_{-\kappa^2}$ and r is the geodesic distance in $\mathbf{H}^3_{-\kappa^2}$ to a fixed point contained inside the region enclosed by $\iota(\Sigma) \subset \mathbf{H}^3_{-\kappa^2}$. The right hand side is intrinsic to (Σ^2, γ) and, therefore

$$\sup \left\{ \int_{\partial\Omega} H_g \, d\sigma : (\Omega^3, g) \in \mathcal{F}_{(\Sigma,\gamma)} \text{ is orientable} \right\} < \infty. \tag{2.3.5}$$

Next, let us lift the orientability hypothesis. Suppose $(\Omega^3, g) \in \mathcal{F}_{(\Sigma,\gamma)}$ is nonorientable. Let $\widetilde{\Omega}$ be the orientation double cover of Ω and let $\pi : \widetilde{\Omega} \implies \Omega$ be the corresponding covering map. It is easily seen that $\partial\widetilde{\Omega} = \pi^{-1}(\partial\Omega)$ and $\partial\widetilde{\Omega}$ doubly covers $\partial\Omega$. Let $\widetilde{\Sigma} = \pi^{-1}(\Sigma)$. Since Σ is a 2-sphere and $\widetilde{\Sigma}$ doubly covers Σ, $\widetilde{\Sigma}$ is the disjoint union of two 2-spheres, which we denote by $\widetilde{\Sigma}^{(1)}$ and $\widetilde{\Sigma}^{(2)}$. Note that $\widetilde{\Sigma}^{(1)}$ and $\widetilde{\Sigma}^{(2)}$, with their induced metric, are both isometric to (Σ, γ).

Let $\eta > 0$ be arbitrary. Employ the cutting lemma (Lemma 2.3.8) to isolate $\widetilde{\Sigma}^{(1)}$ and produce a manifold on which you employ the filling lemma (Lemma 2.3.9) to produce $(D^3, h) \in \mathcal{F}_{(\Sigma,\gamma)}$ with $H_{\partial D,h} > H_{\Sigma^{(1)},g} - \eta$. Note that the cutting lemma is guaranteed to produce orientable three-dimensional manifolds, so the three-dimensional manifold (D^3, h) produced by the filling lemma is also orientable. Therefore,

$$\int_{\widetilde{\Sigma}^{(1)}} (H_g - \eta) \, d\sigma < \int_{\partial D} H_h \, d\sigma \leq \sup \left\{ \int_{\partial\Omega} H \, d\sigma : (\Omega^3, g) \in \mathcal{F}_{(\Sigma,\gamma)} \text{ is orientable} \right\}.$$

Rearranging and recalling that $\widetilde{\Sigma}^{(1)} \cong \partial\Omega$ and $H_{\partial\Omega,g} = H_{\Sigma^{(1)},g}$, we see that

$$\int_{\partial\Omega} H_g \, d\sigma \leq \eta |\Sigma|_\gamma + \sup \left\{ \int_{\partial\Omega} H \, d\sigma : (\Omega^3, g) \in \mathcal{F}_{(\Sigma,\gamma)} \text{ is orientable} \right\}, \tag{2.3.6}$$

whose right hand side is intrinsic to (Σ^2, γ). Thus, by taking the supremum over all $(\Omega^3, g) \in \mathcal{F}_{(\Sigma,\gamma)}$, orientable or not, we conclude from (2.3.5) and (2.3.6) that

$$\Lambda_{(\Sigma,\gamma)} = \sup \left\{ \int_{\partial\Omega} H_g \, d\sigma : (\Omega^3, g) \in \mathcal{F}_{(\Sigma,\gamma)} \right\} < \infty.$$

(2) is a corollary of Theorem 2.3.5.

(3) is a direct consequence of [64, Theorem 1] (also see Theorem 3.5.2 in Section 3.5).

(4) Suppose Σ_0 is a mean-convex, outer area-minimizing 2-sphere in $\mathbf{R}^3$, which bounds a region

$\Omega_0^3 \subset \mathbf{R}^3$. By definition, $(\Omega_0^3, \delta) \in \mathcal{F}_{(\Sigma,\gamma)}$ where δ denotes the Euclidean metric on $\mathbf{R}^3$. Therefore

$$\Lambda_{(\Sigma,\gamma)} \geq \int_{\Sigma_0} H_0 \, d\sigma. \tag{2.3.7}$$

On the other hand, if H_0 is the mean curvature of Σ_0, then by the Aleksandrov-Fenchel inequality for outer-minimizing surfaces (see Definition 3.2.7, and [23, Theorem 6] due to Huisken),

$$\int_{\Sigma_0} H_0 \, d\sigma \geq \sqrt{16\pi\sigma_\delta(\Sigma_0)} = \sqrt{16\pi\sigma_\gamma(\Sigma)}; \tag{2.3.8}$$

equality holds if and only if Σ_0 is round. The desired inequality follows by combining (2.3.7) and (2.3.8). The equality case follows from the equality case in (2.3.8) and (3). $\square$

Remark 2.3.12. *It would be interesting to know—and is the subject of further research—whether conclusion (1) in Theorem 2.3.3 continues to hold if we replace the boundary 2-sphere with a surface of higher genus. Note that Theorem 2.3.5 does not make any genus assumptions.*

Chapter 3

General Theory of Relativity

3.1 Introduction

3.1.1 Spacetime

Lorentzian manifolds (L^n, h) differ from their more standard counterparts, Riemannian manifolds, in that their metric tensor is not necessarily positive definite. It is natural, then, to consider at any point $p \in L$ the following nontrivial decomposition of the tangent space T_pL:

$$T_pL = \{\mathbf{X} \in T_pL : \langle \mathbf{X}, \mathbf{X} \rangle_h = 0\} \cup \{\mathbf{X} \in T_pL : \langle \mathbf{X}, \mathbf{X} \rangle_h < 0\} \cup \{\mathbf{X} \in T_pL : \langle \mathbf{X}, \mathbf{X} \rangle_h > 0\}. \quad (3.1.1)$$

Vectors satisfying $\langle \mathbf{X}, \mathbf{X} \rangle_h = 0$ are called null, and altogether they form a cone (i.e., a set closed under dilations) at every point p. That cone is called a lightcone. Vectors with $\langle \mathbf{X}, \mathbf{X} \rangle_h < 0$ are called timelike. Finally, vectors with $\langle \mathbf{X}, \mathbf{X} \rangle_h > 0$ are called spacelike. Likewise, submanifolds N^k of a Lorentzian manifold (L^n, h) on which h restricts to a null, negative definite, or positive definite 2-tensor are called null, timelike, or spacelike.

Lorentzian geometry arises in Einstein's general theory of relativity, where spacetime is a four-dimensional Lorentzian manifold (L^4, h). Photons travel along null curves, massive particles travel along timelike curves, i.e., they travel slower than photons, and spacelike hypersurfaces ("slices") roughly represent a "$t = t_0$" snapshot of spacetime for a particular observer. Einstein asserted that matter curves spacetime, and expressed this through the equation

$$\mathrm{Ric}_L - \frac{1}{2} R_L h = 8\pi T, \quad (3.1.2)$$

that is referred to as **Einstein's equation**. In (3.1.2), Ric_L is the Ricci curvature of (L^4, h), $R_L = \mathrm{Tr}_L \mathrm{Ric}_L$ is the scalar curvature of (L^4, h), and T is the **stress-energy 2-tensor** that represents the matter field. Mathematical general relativity, broadly speaking, concerns itself with the study of solutions of Einstein's equation.

For a thorough mathematical treatment of general relativity, we refer the reader to [68].

3.1.2 Initial data sets

The **initial data set approach** to general relativity is motivated by the fact that, in suitable gauge, Einstein's equation (3.1.2) is a quasilinear wave equation. In analogy with the linear wave equation, one demands the existence of a spacelike ("Cauchy") hypersurface $M^3 \overset{\iota}{\hookrightarrow} L^4$ that intersects every inextendible timelike curve exactly once. Such a spacetime is called **globally hyperbolic**, and M serves as a "$t = t_0$" initial hypersurface, dividing spacetime into a "future" and a "past" portion. It carries a future-pointing timelike unit normal vector field η that corresponds to the velocity of a stationary observer at "$t = t_0$." (See [6] for a proof of this.)

The spacelike hypersurface M^3, its Riemannian metric, and its second fundamental form with respect to η serve as Cauchy data. The Cauchy formulation of Einstein's equation (3.1.2) is, therefore:

$$\left\{\begin{matrix} \mathrm{Ric}_L - \frac{1}{2} R_L h = 8\pi T \text{ in } L^4 \\ \iota^* h = g \text{ on } M^3 \\ \iota^* \nabla^L \eta = p \text{ on } M^3; \end{matrix}\right\} \tag{3.1.3}$$

$\iota^* h$ is the metric on M^3 induced by (L^4, h), and is prescribed to equal g; $\iota^* \nabla^L \eta$ is the induced second fundamental form $\mathrm{II}_M(\mathbf{X}, \mathbf{Y}) = \langle \nabla^L_{\mathbf{X}} \eta, \mathbf{Y} \rangle_h$, $\mathbf{X}$, $\mathbf{Y} \in TM$, and is prescribed to equal p.

We note, in passing, that not all tuples (M^3, g, p) are valid initial data. Since (3.1.3) expresses T in terms of Ric_L, R_L, the Gauss and Codazzi equations on M^3 (Proposition A.1.5) force implicit relationships among: the positive definite symmetric 2-tensor g; the symmetric 2-tensor p; the scalar field $\mu \triangleq T(\eta, \eta)$ on M^3; and 1-form $J \triangleq \iota^*(T(\eta, \cdot))$ on M^3. Specifically:

$$16\pi\mu = R_M - \|p\|_M^2 + (\mathrm{Tr}_M p)^2 \tag{3.1.4}$$

$$8\pi J = \mathrm{div}_M p - d \, \mathrm{Tr}_M p, \tag{3.1.5}$$

These are called the **Einstein constraint equations**.

Definition 3.1.1 (Initial data set, cf. Schoen-Yau [62, Section 1]). *A tuple* (M^3, g, p) *is called an initial data set if it satisfies the Einstein constraint equations* (3.1.4)-(3.1.5).

The stress-energy tensor $T(\mathbf{X},\mathbf{Y})$ measures the flux in the $\mathbf{Y}$ direction of energy measured by an observer moving with velocity $\mathbf{X}$. Said otherwise, the vector $-T(\mathbf{X},\cdot)^\sharp$ is the energy flow velocity as experienced by an observer moving with velocity $\mathbf{X}$. T generalizes the classical stress tensor from Newtonian mechanics that measures spacial pressure and shear. In (Lorentzian) orthonormal coordinates with first basis vector $\mathbf{X} \triangleq \eta$ being an observer's spacetime velocity, T is given in matrix form by

$$T = \left[\begin{array}{c:c} \mu & J \\ \hdashline J & \text{stress} \end{array}\right]; \tag{3.1.6}$$

$\mu = T(\eta,\eta)$ and $J = \iota^*(T(\eta,\cdot))$, which have already appeared in the Einstein constraint equations, are called the is called the **energy density** and the **energy flux density**, respectively.

Remark 3.1.2. *Unlike Newtonian mechanics, energy density and energy flux density are dependent on the observer's spacetime velocity. This complicates attempts to formulate a notion of energy. As a matter of fact, in the general theory of relativity there is a plethora of competing notions of energy which are not, a priori, related with one another. Section 3.2 discusses this matter in more detail.*

3.1.3 Assumptions

We now list our assumptions on the initial data sets (M^3, g, p) studied in this thesis.

Assumption 1. We assume that

$$(M^3, g, p) \text{ is time-symmetric, i.e., } p = 0. \tag{TS}$$

From now on, thus, we will ignore p. We will see that this assumption puts the study of initial data sets in direct correspondence with the geometry of three-dimensional Riemannian manifolds, scalar curvature, and minimal surfaces. Physically, (TS) encodes the requirement that M^3 be totally geodesic in (L^4, h).

Assumption 2. We assume that

$$(M^3, g) \text{ is asymptotically flat. (See Definition 3.1.3.)} \tag{AF}$$

Physically, (AF) encodes the requirement that (M^3, g) be a slice of a spacetime modeling isolated gravitating systems.

Assumption 3. We assume that

$$T \text{ satisfies the dominant energy condition on } (M^3, g). \text{ (See Definition 3.1.4.)} \tag{DEC}$$

Physically, (DEC) requires that energy be prohibited from traveling faster than light.

Now follow all relevant definitions:

Definition 3.1.3 (cf. Schoen-Yau [62, Section 1]). *A Riemannian manifold (M^3, g) is asymptotically flat with q ends at infinity if:*

(1) there exists a compact $K \subset\subset M$ such that

$$M \setminus K \approx \bigsqcup_{\ell=1}^{q} E_\ell, \text{ with } E_\ell \approx \mathbf{R}^3 \setminus \{\mathbf{x} \in \mathbf{R}^3 : |\mathbf{x}| \leq 1\} \text{ for } \ell \in \{1, \ldots, q\};$$

(2) for every $\ell \in \{1, \ldots, q\}$, there exists a diffeomorphism $\Phi_\ell : E_\ell \to \mathbf{R}^3 \setminus \{\mathbf{x} \in \mathbf{R}^3 : |\mathbf{x}| \leq 1\}$ and a constant C_ℓ such that, using Φ_ℓ as a coordinate chart,

$$\sum_{i,j=1}^{3} \left(|g_{ij} - \delta_{ij}| + |\mathbf{x}| \sum_{k=1}^{3} |\partial_k g_{ij}| + |\mathbf{x}|^2 \sum_{k,m=1}^{3} |\partial_k \partial_m g_{ij}| \right) \leq C_\ell |\mathbf{x}|^{-1}.$$

(3) the scalar curvature of (M^3, g) is integrable; i.e.,

$$\int_M |R_M| \, dv_M < \infty.$$

Definition 3.1.4 (Dominant energy condition, cf. Wald [68, Section 9.2]). *A stress-energy tensor T on a globally hyperbolic Lorentzian manifold (L^4, h) is said to satisfy the dominant energy condition if, for every $p \in L$ and every future-pointing timelike $\mathbf{X} \in T_pL$, $-T(\mathbf{X}, \cdot)^\sharp$ is a future-pointing timelike or null vector.*

On the initial datum (M^3, g), the dominant energy condition requires that $-T(\eta, \cdot)^\sharp$ be future-pointing timelike or null. From Einstein's equation (3.1.2) we see that

$$8\pi T(\eta, \cdot) = \operatorname{Ric}_L(\eta, \cdot) - \frac{1}{2} R_L \langle \eta, \cdot \rangle_h. \tag{3.1.7}$$

Recall that, by Assumption (TS), M^3 is totally geodesic in L^4. Tracing the Gauss equation (Proposition A.1.5) twice over M^3 we find that

$$R_M = R_L + 2 \operatorname{Ric}_L(\eta, \eta). \tag{3.1.8}$$

(The sign change relative to the Riemannian Gauss equation is due to $\langle \eta, \eta \rangle_h = -1$.) Likewise, tracing the Codazzi equation (Proposition A.1.5) of M over M we find that

$$\iota^*(\mathrm{Ric}_L(\eta, \cdot)) = 0. \tag{3.1.9}$$

Plug (3.1.8), (3.1.9) into (3.1.7). Then, recalling $\langle \eta, \eta \rangle_h = -1$, $8\pi T(\eta, \cdot) = -8\pi T(\eta, \eta)\eta^\flat = -\frac{1}{2} R_M \eta^\flat$, so

$$-T(\eta, \cdot)^\sharp = \frac{1}{16\pi} R_M \eta \text{ is future pointing timelike or null } \iff R_M \geq 0.$$

Summarizing, (AF), (DEC), and (TS), reduce the study of initial data sets to the study of

asymptotically flat 3-manifolds (M^3, g) *with nonnegative scalar curvature.*

We conclude this section by describing the most important asymptotically flat initial data sets that occur as time-symmetric slices of spacetimes that satisfy the Einstein equations.

Example 3.1.5 (Euclidean space)**.** $(\mathbf{R}^3, \delta)$ *is asymptotically flat with one end and flat. It appears as a time-symmetric* $\{t = t_0\}$ *slice of Minkowski space* $(\mathbf{R}^3 \times \mathbf{R}, \delta - dt^2)$*, with* $T \equiv 0$.

Example 3.1.6 (Mass-m Schwarzschild)**.**

$$\left(\mathbf{S}^2 \times (2m, \infty), r^2 \gamma_{\mathbf{S}^2} + \left(1 - \frac{2m}{r}\right)^{-1} dr^2 \right) \tag{3.1.10}$$

is asymptotically flat with one end and scalar-flat; $\gamma_{\mathbf{S}^2}$ *denotes the round metric on* $\mathbf{S}^2$ *with area* 4π*. This manifold appears as a time-symmetric* $\{t = t_0\}$ *slice of (Lorentzian) mass-m Scharzschild*

$$\left(\mathbf{S}^2 \times (2m, \infty) \times \mathbf{R}, r^2 \gamma_{\mathbf{S}^2} + \left(1 - \frac{2m}{r}\right)^{-1} dr^2 - \left(1 - \frac{2m}{r}\right) dt^2 \right),$$

the first nontrivial solution of (3.1.2)*, with* $T \equiv 0$*. Notice that* (3.1.10) *is incomplete. In fact, its metric completion up to* $\{r = 2m\}$ *is a complete manifold with boundary. Two copies of* (3.1.10) *can be found isometrically embedded in*

$$\left(\mathbf{R}^3 \setminus \{\mathbf{0}\}, \left(1 + \frac{2m}{|\mathbf{x}|}\right)^4 \delta \right). \tag{3.1.11}$$

This manifold is complete, asymptotically flat with two ends, conformally flat, and scalar-flat. The two isometric copies of (3.1.10) *can be found in the regions* $\{0 < |\mathbf{x}| < 2m\}$ *and* $\{|\mathbf{x}| > 2m\}$.

3.2 Mass

3.2.1 ADM mass

Recall from Remark 3.1.2 that there is no observer-independent pointwise notion of energy density in general relativity. One of the key advantages of studying isolated gravitating systems is that they admit a meaningful notion of total energy that can be read off their slices (M^3, g). This notion is called the ADM energy and was formulated by Arnowitt, Deser, and Misner in 1959.

Definition 3.2.1 (ADM energy, cf. Arnowitt-Deser-Misner [2]). *Let (M^3, g) be asymptotically flat with q ends $E_1, \ldots, E_q$. We define*

$$\mathfrak{m}_{\mathrm{ADM}}(M^3, g) \triangleq \sum_{\ell=1}^{q} \mathfrak{m}_{\mathrm{ADM}}(E_\ell, g), \tag{3.2.1}$$

where each $\mathfrak{m}_{\mathrm{ADM}}(E_\ell, g)$ is defined by choosing an asymptotically flat chart $\Phi_\ell : E_\ell \to \mathbf{R}^3 \setminus \{\mathbf{x} = (x^1, x^2, x^3) \in \mathbf{R}^3 : |\mathbf{x}| \leq 1\}$ and setting

$$\mathfrak{m}_{\mathrm{ADM}}(E_\ell, g) \triangleq \frac{1}{16\pi} \lim_{r \uparrow \infty} \int_{\{|\mathbf{x}|=r\}} (\partial_i g_{ij} - \partial_j g_{ii}) \frac{x^j}{r} \, d\sigma; \tag{3.2.2}$$

here $d\sigma$ is the area form induced by the Euclidean metric on the 2-sphere $\{|\mathbf{x}| = r\} \subset \Phi_\ell^{-1}(E_\ell)$.

Remark 3.2.2. *For consistency with the literature, we will use the term* **ADM mass** *(not energy) throughout this thesis. Strictly speaking, ADM mass is a different quantity that also takes into account the angular momentum of (M^3, g). However, in the time-symmetric setting* (TS)*, which we are assuming, the two notions coincide.*

Remark 3.2.3. *The right hand side of* (3.2.2) *may, a priori, depend on one's choice of a coordinate chart $\Phi_\ell : E_\ell \to \{\mathbf{x} \in \mathbf{R}^3 : |\mathbf{x}| > 1\}$. Robert Bartnik showed [4, Theorem 4.2] that $\mathfrak{m}_{\mathrm{ADM}}(E_\ell, g)$ (and thus $\mathfrak{m}_{\mathrm{ADM}}(M, g)$) is independent of one's choice of a coordinate chart provided natural decay conditions are imposed in defining asymptotic flatness—our Definition 3.1.3 fulfills them.*

Example 3.2.4. *The ADM mass of $(\mathbf{R}^3, \delta)$ is $\mathfrak{m}_{\mathrm{ADM}}(\mathbf{R}^3, \delta) = 0$.*

Example 3.2.5. *The ADM mass of mass-m Schwarzschild* (3.1.10) *is m.*

The limiting flux integral in (3.2.2) is not, a priori, positive. This made the establishment of the following theorem of Schoen-Yau [60] (later also proven by Witten [73]) a considerable milestone in the field:

Theorem 3.2.6 (Positive mass theorem, cf. Schoen-Yau [60, 62], Witten [73], Parker-Taubes [50]).

Let (M^3, g) be asymptotically flat with q ends $E_1, \ldots, E_q$, nonnegative scalar curvature and complete without boundary, or with boundary consisting of closed surfaces whose mean curvature vector does not point outside M. Then,

$$\mathfrak{m}_{\mathrm{ADM}}(E_\ell, g) \geq 0 \text{ for all } \ell = 1, \ldots, q.$$

Equality $\mathfrak{m}_{\mathrm{ADM}}(E_\ell, g) = 0$ holds for some $\ell \in \{1, \ldots, k\}$ if and only if $(M^3, g) \cong (\mathbf{R}^3, \delta)$.

3.2.2 Apparent horizons

R. Penrose conjectured ([51]) that spacetime settles down to a Kerr black hole and, moreover, that ADM mass can be estimated from below by the surface area of a cross section of the event horizon.

Proven (in different forms) by Bray [9] and Huisken-Ilmanen [32], the Penrose inequality quantifies the fact that, in the time symmetric setting, $\mathfrak{m}_{\mathrm{ADM}}$ can be further estimated from below by a special choice of closed interior surfaces that are visible from those ends. We need to make two definitions first:

Definition 3.2.7 (cf. Bray [9, Definitions 5, 6], Huisken-Ilmanen [32])**.** *Let (M^3, g) be complete and asymptotically flat, possibly with boundary. A surface $\Sigma \subset M$ is said to be* **outer-minimizing** *with respect to an end E if there exists no surface $\Sigma' \subset M$ separating Σ from E's asymptotic infinity with less area. A minimal surface $T \subset M$ is said to be* **outermost** *with respect to an end E if there is no minimal surface $T' \subset M$ separating T from E's asymptotic infinity. Outermost minimal surfaces are also called* **apparent horizons**.

Closed minimal surfaces in (M^3, g) are called **horizons**. Horizons need not be subsets of apparent horizons. Physically, apparent horizons represent quasi-local black hole boundaries that are observable from exterior regions. Mathematically, apparent horizons are extremely important in the initial data set approach to general relativity, as well as in this thesis.

Remark 3.2.8. *It is not hard to see that every mean-convex outermost surface is outer-minimizing.*

Example 3.2.9. *Mass-m Schwarzschild has exactly one apparent horizon; it is a round 2-sphere with area $16\pi m^2$. In* (3.1.10) *it corresponds to $\{r = 2m\}$ and in* (3.1.11) *to $\{|\mathbf{x}| = m/2\}$.*

Hawking's theorem on the topology of black holes asserts that cross sections of the event horizon in four-dimensional asymptotically flat stationary black hole spacetimes obeying the dominant energy condition are topologically 2-spheres [31]. The result extends to apparent horizons in black hole spacetimes that are not necessarily stationary. In the time symmetric initial data set setting, apparent horizons are stable critical points for area.

Lemma 3.2.10 (Huisken-Ilmanen [32, Lemma 4.1]). *Let (M^3, g) be a complete, asymptotically flat manifold with q ends, nonnegative scalar curvature, with boundary consisting of compact minimal surfaces. The set*

$$K_1 \triangleq \overline{\bigcup \{\Sigma^2 \subset M \ \textit{is a closed immersed minimal surface}\}}$$

is compact. The set $M \setminus K_1$ has exactly q unbounded connected components, one corresponding to each of the q ends. For each such unbounded component, its metric completion E is:

(1) diffeomorphic to $\mathbf{R}^3$ minus a finite number of open balls with disjoint closures,

(2) complete with minimal boundary ∂E that minimizes area in its homology class, and

(3) contains no closed, immersed minimal surfaces in its interior.

With all this notation in place, we may return to the Penrose conjecture. There are two related—but inequivalent—versions of the Penrose conjectured that have been proven to date:

Theorem 3.2.11 (Riemannian Penrose inequality, I , Bray [9, Theorem 4]). *Let (M^3, g) be complete, asymptotically flat with nonnegative scalar curvature and one end, and an outer-minimizing (possibly empty) minimal boundary. Then,*

$$\mathfrak{m}_{\mathrm{ADM}}(M^3, g) \geq \sqrt{\frac{\sigma(\partial M^3)}{16\pi}}. \tag{3.2.3}$$

Equality holds if and only if (M, g) is isometric to an exterior region of mass-$\mathfrak{m}_{\mathrm{ADM}}(M^3, g)$ Schwarzschild

$$\left(\mathbf{R}^3 \setminus B_{m/2}(\mathbf{0}), \left(1 + \frac{\mathfrak{m}_{\mathrm{ADM}}(M^3, g)}{2|\mathbf{x}|}\right)^4 \delta\right)$$

outside their respective apparent horizons.

Theorem 3.2.12 (Riemannian Penrose inequality, II, Huisken-Ilmanen [32]). *Let (M^3, g) be complete, asymptotically flat with one end and nonnegative scalar curvature, with compact, outer-minimizing (possibly empty) boundary ∂M, and no closed minimal surfaces in its interior. If Σ is a component of $\partial\Omega$ and $\partial\Omega \setminus \Sigma$ is minimal, then*

$$\mathfrak{m}_{\mathrm{ADM}}(M^3, g) \geq \sqrt{\frac{\sigma(\Sigma^2)}{16\pi}} \left(1 - \frac{1}{16\pi} \int_{\Sigma} H^2 \, d\sigma\right). \tag{3.2.4}$$

Equality holds if and only if (M^3, g) embeds isometrically into mass-$\mathfrak{m}_{\mathrm{ADM}}(M^3, g)$ Schwarzschild,

$$\left(\mathbf{R}^3 \setminus \{\mathbf{0}\}, \left(1 + \frac{\mathfrak{m}_{\mathrm{ADM}}(M)}{2|\mathbf{x}|}\right)^4 \delta\right).$$

Remark 3.2.13. *Theorem 3.2.11 allows the mass to be compared to the area of a possibly disconnected surface, but that surface must be a horizon. On the other hand, Theorem 3.2.12 requires the comparison to be with a connected surface, but the surface need not be a horizon.*

3.3 Bartnik quasi-local mass of apparent horizons

3.3.1 Definitions

In 1989, Robert Bartnik gave a definition of quasi-local mass for compact spacelike domains, motivated by the concept of electrostatic capacity. His quasi-local mass notion, adapted per Huisken-Ilmanen, is:

Definition 3.3.1 (Bartnik quasi-local mass, cf. Bartnik [5], Huisken-Ilmanen [32, Section 9])**.** *Let* (Ω^3, g) *be compact, connected, have nonnegative scalar curvature and nonempty boundary consisting of a mean-convex surface* Σ *and a (possibly empty) union of minimal surfaces* T*. Define* $\mathcal{E}_{(\Omega,g)}$ *to consist of all complete, asymptotically flat* (M^3, g) *with nonnegative scalar curvature, one end, containing an isometric copy of* (Ω, g)*, no closed interior minimal surfaces. When* $\mathcal{E}_{(\Omega,g)}$ *is nonempty, define*

$$\mathfrak{m}_{\mathrm{B}}(\Omega^3, g) \triangleq \inf\{\mathfrak{m}_{\mathrm{ADM}}(M^3, g) : (M^3, g) \in \mathcal{E}_{(\Omega,g)}\}.$$

The Bartnik mass satisfies three properties that one would require of a quasi-local mass functional. (We refer to [32, Section 9] for proofs of these facts.)

(1) Monotonicity: if (Ω, g) and (Ω', g') are as in Definition 3.3.1 and (Ω, g) embeds isometrically inside (Ω', g'), and $\mathcal{E}_{(\Omega',g')}$ is nonempty, then $\mathfrak{m}_{\mathrm{B}}(\Omega', g') \geq \mathfrak{m}_{\mathrm{B}}(\Omega, g)$.

(2) Nonnegativity: if (Ω, g) is as in Definition 3.3.1, $\mathcal{E}_{(\Omega,g)}$ is nonempty, and $\mathfrak{m}_{\mathrm{B}}(\Omega, g) = 0$, then (Ω, g) is flat.

(3) Exhaustion: if (M^3, g) is complete, asymptotically flat with one end, with nonnegative scalar curvature and a (perhaps empty) boundary consisting of minimal surfaces, $\{\Omega_k\}_{k=1,2,\ldots}$ is a sequence of bounded sets as in Definition 3.3.1 whose indicator functions $\mathbf{1}_{\Omega_k}$ converge to 1 in $C^0_{\mathrm{loc}}(M)$, then $\lim_k \mathfrak{m}_{\mathrm{B}}(\Omega_k, g) = \mathfrak{m}_{\mathrm{ADM}}(M^3, g)$.

Despite its apparent simplicity, Bartnik's quasi-local mass has proven to be an elusive quantity to compute and understand. The issue is twofold:

(1) it is difficult to construct admissible extensions, and

(2) $\mathfrak{m}_{\mathrm{ADM}}$ is difficult to compute, so computing its infimum over said class of extensions is practically intractable.

One of the major open questions in the initial data set study of general relativity is Bartnik's static minimization conjecture:

Conjecture 3.3.2 (Bartnik's static minimization conjecture, cf. Huisken-Ilmanen [32, Section 9])**.** *Let (Ω^3, g) be as in Definition 3.3.1, with nonempty $\mathcal{E}_{(\Omega,g)}$. Then $\mathfrak{m}_{\mathrm{B}}(\Omega, g)$ is uniquely attained by an extension $(M^3, \hat{g})$ satisfying the following properties:*

(1) $(M^3, \hat{g})$ is smooth on $M \setminus \partial\Omega$ and $C^{0,1}$ on $\partial\Omega$;

(2) $(M \setminus \Omega, \hat{g})$ has nonnegative scalar curvature, and induces the same metric and mean curvature on $\partial\Omega$ as (Ω, g);

(3) $(M \setminus \Omega, \hat{g})$ does not contain minimal surfaces in its interior;

(4) $(M \setminus \Omega, \hat{g})$ is static, i.e., there exists a bounded, positive function f on $M \setminus \Omega$ with $\Delta_{\hat{g}} f = 0$, $f \operatorname{Ric}_{\hat{g}} = \nabla^2_{\hat{g}} f$, and $f \to 1$ at infinity.

In view of the first listed difficulty above, as well as the conjectured regularity of the minimizing extension, there are alternate definitions of Bartnik's mass that only require extensions to satisfy the properties asserted in Conjecture 3.3.2. In this context, one speaks of the quasilocal mass function $m_B(\partial\Omega, g, H)$.

Definition 3.3.3 (Bartnik mass of apparent horizons, M., cf. M.-Schoen [40])**.** *Let (Σ^2, γ) be given. Define $\mathcal{E}(\Sigma, \gamma, H = 0)$ to consist of all complete, asymptotically flat (M^3, g) with no closed interior minimal surfaces and minimal boundary $\partial(M^3, g) \cong (\Sigma^2, \gamma)$. If $\mathcal{E}(\Sigma, \gamma, H = 0)$ is nonempty, define*

$$\mathfrak{m}_{\mathrm{B}}(\Sigma, \gamma, H = 0) \triangleq \inf\{\mathfrak{m}_{\mathrm{ADM}}(M^3, g) : (M^3, g) \in \mathcal{E}_{(\Sigma,\gamma,H=0)}\}.$$

Note that, by the Riemannian Penrose inequality (Theorems 3.2.11, 3.2.12),

$$\mathfrak{m}_{\mathrm{B}}(\Sigma^2, \gamma, H = 0) \geq \sqrt{\frac{\sigma_\gamma(\Sigma^2)}{16\pi}}. \tag{3.3.1}$$

Remark 3.3.4. *Apparent horizons were not part of the Bartnik conjecture, nor were they part of the classical notion of Bartnik quasi-local mass.*

3.3.2 Results

We begin by characterizing the intrinsic geometry of apparent horizons; that is, the Riemannian metrics which can arise. We then compute the Bartnik quasi-local mass of all "nondegenerate" apparent horizons—the first explicit computation of any nontrivial Bartnik mass—and, because of the sharpness of the computation, we show that the the static minimization conjecture is **false** for apparent horizon boundaries. In other words, the Bartnik mass of apparent horizons is **never** achieved by nondegenerate initial data sets, except for the trivial case of round spheres that occur as Schwarzschild horizons. Results in this section appear in a similar form in [40], published jointly with R. Schoen.

Theorem 3.3.5 (M., cf. M.-Schoen [40])**.** *For time symmetric asymptotically flat initial data sets with the dominant energy condition,*

$$\mathscr{M}_+ \subseteq \{ \textit{ connected components of apparent horizons } \} \subseteq \mathscr{M}_+ \cup \partial \mathscr{M}_+. \tag{3.3.2}$$

Apparent horizons $(\mathbf{S}^2, \gamma) \in \mathscr{M}_+$ *all satisfy*

$$\mathfrak{m}_{\mathrm{B}}(\mathbf{S}^2, \gamma, H = 0) = \sqrt{\sigma_\gamma(\mathbf{S}^2)/16\pi}, \tag{3.3.3}$$

and $\mathfrak{m}_{\mathrm{B}}(\mathbf{S}^2, \gamma, H = 0)$ *is not attained by any element of* $\mathcal{E}_{(\mathbf{S}^2,\gamma,H=0)}$ *unless* $(\mathbf{S}^2, \gamma)$ *is round.*

Components of apparent horizons that are not in $\mathscr{M}_+$ *("degenerate") are totally geodesic 2-spheres in ambient scalar-flat regions. Apparent horizons with degenerate components can be perturbed into ones whose components are in* $\mathscr{M}_+$ *without affecting the ADM mass of their initial data set.*

Remark 3.3.6. *The chain of inclusions* (3.3.2) *bring all tools of Section 2.3 to our disposal. For example, notice that as a corollary of Theorem 3.3.5 and Corollary 2.2.6, there exists a very large class of examples of apparent horizons with arbitrarily large* $\int_{S^2} \max\{0, -K_\gamma\}\, d\sigma_\gamma$. *All known examples of apparent horizons prior to our constructions had* $K_\gamma > 0$.

We use the methods in Chapter 2 to construct a collar extension of γ, with positive scalar curvature, and derive a gluing procedure to attach this collar extension to an exterior mass-m Schwarzschild region. The extension procedure is explicit and allows us to keep track of the mass parameter m accurately and, in fact, have it be arbitrarily close to the optimal value. The result that apparent horizons with a nondegeneracy condition are in bijective correspondence with a related moduli space of metrics was also obtained by B. Smith in [67], who was also able to specify the tracefree part of the second fundamental of the horizon. The methods in [67], however, do not give control on the ADM mass of the extension as ours do.

The main theorem of this section, Theorem 3.3.7, constructs a minimizing sequence of smooth

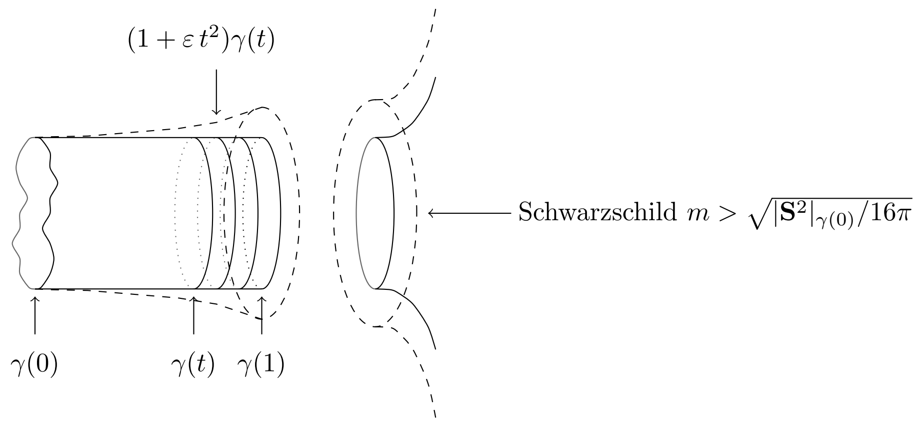

Figure 3.1: Gluing $\mathscr{M}_+$ cylinder to Schwarzschild

extensions of the Bartnik data $(\Sigma^2, \gamma, H = 0)$.

Theorem 3.3.7 (M.-Schoen [40, Theorem 2.1])**.** *For all* $(\mathbf{S}^2, \gamma) \in \mathscr{M}_+$*,* $m > \sqrt{\sigma_\gamma(\mathbf{S}^2)/16\pi}$*, there exists* $(M^3, g) \in \mathcal{E}_{(\mathbf{S}^2,\gamma,H=0)}$ *which is isometric to a mass-m Schwarzschild metric outside a compact set.*

Remark 3.3.8. *The manifold* $(M^3, g) \in \mathcal{E}_{(\mathbf{S}^2,\gamma,H=0)}$ *constructed in Theorem 3.3.7 has the following additional properties:*

(1) It is foliated by mean-convex 2-spheres that eventually coincide with Schwarzschild coordinate spheres.

(2) Its boundary $\partial(M, g) \cong (\mathbf{S}^2, \gamma)$ *is totally geodesic and, in fact,* (M^3, g) *can be doubled across its boundary to obtain a smooth, complete, asymptotically flat manifold with nonnegative scalar curvature, two ends, and a unique apparent horizon.*

Remark 3.3.9. *While we only construct initial data sets with horizon boundary, we can construct a local spacetime containing this data. First, we double across the horizon to construct a complete initial data set with no boundary and the property that there is a single apparent horizon (see Remark 3.3.8 above). We can then set up a Cauchy problem for the Einstein equations* (3.1.3) *coupled with a matter field (e.g., a perfect fluid with vanishing initial 3-velocity and appropriate state function) and use the existence theory for the Cauchy problem to construct a maximal spacetime extension which satisfies the dominant energy condition (see [57]). We expect that, after a relatively short time, there will be an apparent horizon which is near a round metric on* $\mathbf{S}^2$ *since the initial data set contains approximately round 2-spheres which are nearly apparent horizons. We do not know if there is a modification of our construction which produces vacuum initial data, i.e., with vanishing stress-energy tensor in* (3.1.3).

Let us first prove Theorem 3.3.5 assuming that Theorem 3.3.7 holds true.

Proof of Theorem 3.3.5. Fix $(\mathbf{S}^2, \gamma) \in \mathscr{M}_+$, let $m > \sqrt{\sigma_\gamma(\mathbf{S}^2)/16\pi}$, and let (M^3, g) be constructed in Theorem 3.3.7 with $\mathfrak{m}_{\mathrm{ADM}}(M, g) < m$. By the maximum principle applied to the mean-convex spherical foliation of (M^3, g) (see Remark 3.3.8), we see that (M^3, g) contains no interior closed minimal surfaces, and so its boundary $(\mathbf{S}^2, \gamma)$ is outer-minimizing. In other words, $(M^3, g) \in \mathcal{E}_{(\mathbf{S}^2, \gamma, H=0)}$, so $\mathfrak{m}_{\mathrm{B}}(\mathbf{S}^2, \gamma, H = 0) \leq \mathfrak{m}_{\mathrm{ADM}}(M, g) < m$. Since $m > \sqrt{\sigma_\gamma(\mathbf{S}^2)/16\pi}$ was arbitrary:

$$\mathfrak{m}_{\mathrm{B}}(\mathbf{S}^2, \gamma, H = 0) = \inf\{\mathfrak{m}_{\mathrm{ADM}}(M^3, g) : (M^3, g) \in \mathcal{E}_{(\mathbf{S}^2, \gamma, H=0)}\} \leq \sqrt{\frac{\sigma_\gamma(\mathbf{S}^2)}{16\pi}}. \tag{3.3.4}$$

Conversely, by the Riemannian Penrose inequality (Theorems 3.2.11, 3.2.12),

$$\mathfrak{m}_{\mathrm{B}}(\mathbf{S}^2, \gamma, H = 0) = \inf\{\mathfrak{m}_{\mathrm{ADM}}(M^3, g) : (M^3, g) \in \mathcal{E}_{(\mathbf{S}^2, \gamma, H=0)}\} \geq \sqrt{\frac{\sigma_\gamma(\mathbf{S}^2)}{16\pi}}. \tag{3.3.5}$$

Putting (3.3.4) and (3.3.5) gives (3.3.3).

Along the same note, the extension procedure in Theorem 3.3.7 guarantees that every element of $\mathscr{M}_+$ appears as a component of an apparent horizon for an initial data set. Therefore,

$$\mathscr{M}_+ \subseteq \{\text{ connected components of apparent horizons }\}. \tag{3.3.6}$$

Conversely, Lemma 2.2.10 characterizes the topology of connected components of apparent horizons as being spherical, and that every component (Σ^2, γ) bounds a 3-ball, so it is two-sided. Being stable two-sided surfaces, from Proposition A.2.6 they satisfy

$$\frac{1}{2}\int_\Sigma \left(R_M - 2K_\Sigma + \| \mathrm{II}_\Sigma \|^2\right) \varphi^2 \, d\sigma \leq \int_\Sigma \|\nabla^\Sigma \varphi\|^2 \, d\sigma. \tag{3.3.7}$$

Since R_M, $\| \mathrm{II}_\Sigma \|^2 \geq 0$, it follows that the second order linear elliptic operator $-\Delta + K$ on (Σ^2, γ) is a nonnegative definite operator; i.e.,

$$\{\text{ connected components of apparent horizons }\} \subseteq \mathscr{M}_+ \cup \partial \mathscr{M}_+. \tag{3.3.8}$$

Combining (3.3.6) and (3.3.8) gives (3.3.2).

Next, suppose that $(\mathbf{S}^2, \gamma) \in \mathscr{M}_+$, $(M^3, g) \in \mathcal{E}_{(\mathbf{S}^2, \gamma, H=0)}$ are such that $\mathfrak{m}_{\mathrm{ADM}}(M^3, g) = \mathfrak{m}_{\mathrm{B}}(\mathbf{S}^2, \gamma, H = 0) = \sqrt{\sigma_\gamma(\mathbf{S}^2)/16\pi}$. The rigidity case of the Riemannian Penrose inequality (Theorem 3.2.11) holds, so we can conclude that (M^3, g) is Schwarzschild and, therefore, its boundary $(\mathbf{S}^2, \gamma) \cong \partial(M^3, g)$ is round. This proves that no nonround elements of $\mathscr{M}_+$ can have minimizing extensions.

Finally, suppose (M^3, g) is a complete, asymptotically flat manifold without boundary, whose apparent horizon contains a $(\Sigma^2, \gamma) \notin \mathscr{M}_+$. Recalling (3.3.7), $R_M|_\Sigma \equiv 0$ and $\mathrm{II}_\Sigma = 0$, so

$$\min_\Sigma \left\| \mathrm{Ric}_M - \frac{1}{3} R_M g \right\| \geq \delta > 0. \tag{3.3.9}$$

Now, since (M^3, g) is complete and asymptotically flat, there exists a Ricci flow (M^3, g_t), $t \in [0, \tau_0)$, with constant ADM mass ([7]). By the evolution equation

$$\left(\frac{\partial}{\partial t} - \Delta_{M,g(t)} \right) R_{M,g(t)} = \frac{2}{3} R^2_{M,g(t)} + 2 \left\| \mathrm{Ric}_{M,g(t)} - \frac{1}{3} R_{M,g(t)} g(t) \right\|^2_{M,g(t)}$$

and the parabolic maximum principle, we know that either $R_M > 0$ for $t > 0$, or $\mathrm{Ric}_M - \frac{1}{3} R_M g \equiv 0$ at $t = 0$. The latter is impossible by virtue of (3.3.9), so the former must be true. It remains to check that the apparent horizon Σ has not disappeared abruptly in (M^3, g_t), $t > 0$.

Claim. Let $\nu \in \Gamma(\Sigma; N\Sigma)$ be a choice of unit normal to Σ. There exists $\varepsilon > 0$ and a local foliation of M near Σ by disjoint surfaces Σ_s, $s \in (-\varepsilon, \varepsilon)$, with constant mean curvature s with respect to their normal vector pointing in the same direction as ν.

Proof of claim. Let $\alpha \in (0, 1)$. Consider the Banach space map $\mathcal{H} : C^{2,\alpha}(\Sigma) \times \mathbf{R} \to C^\alpha(\Sigma) \times \mathbf{R}$ given by

$$\mathcal{H}(u, t) \triangleq \left(H(u) - t, \int_\Sigma u \, d\sigma \right),$$

where $H(u)$ denotes the mean curvature of the surface $\Sigma(u) \triangleq \{\exp_x u(x)\nu(x) : x \in \Sigma\}$. By virtue of $R_M|_\Sigma \equiv 0$ and $\| \mathrm{II}_\Sigma \| = 0$, the linearization of $\mathcal{H}$ at $(u, t) = (0, 0)$ is

$$\delta\mathcal{H}(0, 0)\{v, \tau\} \triangleq \left(-\Delta_\Sigma v + K_\Sigma v - \tau, \int_\Sigma v \, d\sigma \right).$$

We claim that $\delta\mathcal{H}(0, 0)\{v, \tau\}$ is an isomorphism of Banach spaces. We recall that, in the case that we are studying, $\lambda_1(-\Delta_\Sigma + K_\Sigma) = 0$.

This linear map is clearly injective, so let us prove surjectivity. By Proposition B.2.4, $\ker(-\Delta_\Sigma + K_\Sigma) = \{\varphi\}$ for a smooth $\varphi > 0$. Suppose $(w, s) \in C^\alpha \times \mathbf{R}$ are given. Choose

$$\tau = -\frac{\langle w, \varphi \rangle_{L^2(\Sigma)}}{\langle 1, \varphi \rangle_{L^2(\Sigma)}}$$

so that $w + \tau \perp \varphi$ with respect to $L^2(\Sigma)$. By the Fredholm alternative [27, Theorem 5.11], there exists $v \in C^{2,\alpha}(\Sigma)$, uniquely determined mod $\mathrm{span}\{\varphi\}$, such that $-\Delta_\Sigma v + K_\Sigma v = w + \tau$. The condition $\int_\Sigma v \, d\sigma = s$ may be accounted for precisely by adding or subtracting a multiple of φ. $\square$

Invoke the claim above, supposing that $s = \varepsilon$ denotes the surface separating Σ from the asymptotically flat end. For small $t > 0$, the neighborhood U of Σ bounded by Σ_ε and $\Sigma_{-\varepsilon}$ remains mean-convex. By Lemma A.2.8, there exists an area minimizing oriented surface $\Sigma'_t \subset U$ that, by a standard compactness argument from geometric measure theory, converges in the C^∞ graphical topology to Σ as $t \downarrow 0$. □

Before we present the proof of the main theorem, we need to make a few remarks on Schwarzschild metrics. Recall from Example 3.1.6 that mass-m Schwarzschild is a scalar-flat metric on $\mathbf{S}^2 \times (2m, \infty)$ given by

$$g_m = t^2\gamma_* + \Big(1 - \frac{2m}{t}\Big)^{-1} dt^2,$$

with γ_* a round metric on $\mathbf{S}^2$ with area 4π. It will be more convenient for us to view this in the form

$$(\mathbf{S}^2 \times [0, \infty), g_m),\ g_m = u_m(s)^2\gamma_* + ds^2, \tag{3.3.10}$$

where s is the radial geodesic arclength and $s = 0$ corresponds to the horizon $\{t = 2m\}$ in the previous coordinates. The function u_m is easily seen to satisfy

$$u_m(0) = 2m,\ u'_m(0) = 0,\ u'_m(s) = \Big(1 - \frac{2m}{u_m(s)}\Big)^{1/2},\ u''_m(s) = \frac{m}{u_m(s)^2} \text{ for } s \geq 0. \tag{3.3.11}$$

The end goal is to glue g_m near some $\{s = s_0\}$, $0 < s_0 \ll 1$, to a scalar positive collar $\mathbf{S}^2 \times [0, 1]$ that connects $(\mathbf{S}^2, \gamma)$ to a round $\mathbf{S}^2$ through a mean-convex foliation with spheres. In order to do this we will first need to "bend" g_m near $\{s = s_0\}$ and make it scalar positive on an annular region. This positivity will afford us certain freedom in gluing; specifically, it will help provide the $\rho_0 > 0$ for conclusion (4) of Lemma 2.2.10 to apply.

Lemma 3.3.10 (Schwarzschild bending, M.-Schoen [40, Lemma 2.3])**.** *Let $s_0 \in (0, \infty)$. There exists a $\delta \in (0, s_0)$ such that, For every $\delta \in (0, \delta_0]$, there exists a smooth function $\sigma : [s_0 - \delta, \infty) \to (0, \infty)$ such that:*

(1) $\sigma(s) = s$ for all $s \geq s_0$,

(2) σ is monotonically increasing, and

(3) $u_m(\sigma(s))^2\gamma_* + ds^2$ has positive scalar curvature on $s \in [s_0 - \delta, s_0)$; here, γ_* is a round metric on $\mathbf{S}^2$ with area 4π.

Proof. Recalling, from Lemma A.1.6, that the scalar curvature of a metric $f(s)^2\gamma_* + ds^2$ is $\frac{1-(f')^2-2f''f}{f^2}$, what we need to ensure is that, for $s < s_0$,

$$
\begin{aligned}
1-\left(\frac{d}{ds}u_m(\sigma(s))\right)^2-2\left(\frac{d^2}{ds^2}u_m(\sigma(s))\right)u_m(\sigma(s))>0 \\
\iff 1-(u_m')^2(\sigma')^2-2u_m''u_m(\sigma')^2-2u_m'u_m\sigma''>0 \\
\iff 1-(\sigma')^2+(\sigma')^2\big[1-(u_m')^2-2u_m''u_m\big]-2u_m'u_m\sigma''>0.
\end{aligned}
$$

The term inside the bracket vanishes in view of Schwarzschild being scalar flat, so the inequality collapses to $1-(\sigma')^2-2u_m'u_m\sigma''>0$. It will suffice to construct a smooth $\theta : [s_0-\delta_0, s_0] \to \mathbf{R}$ to play the role of σ', provided it satisfies

(1) $\theta(s_0)=1$, $\theta'(s_0)=\theta''(s_0)=\ldots=0$;

(2) $\theta>0$ on $[s_0-\delta_0, s_0]$;

(3) $1-\theta^2-2u_m'u_m\theta'>0$ on $[s_0-\delta_0, s_0)$.

Having done so, we simply integrate to recover σ to the left of s_0 and, by (1) above, it can be smoothly extended via $\sigma(s)\equiv s$ to the right of s_0. The lemma is concluded by observing that

$$
\theta(s) \triangleq \begin{cases} 1+\exp\big(-(s-s_0)^{-2}\big) & \text{for } s<s_0 \\ 1 & \text{for } s=s_0 \end{cases}
$$

works as desired. □

We are now in a position to prove Theorem 3.3.7.

Proof of Theorem 3.3.7. Construct a cylinder $\mathbf{S}^2\times[0,1]$ with positive scalar curvature as in Proposition 2.2.9, where $\mathbf{S}^2\times[0,1]$ is endowed with $g\triangleq\gamma(t)+u(x,t)^2dt^2$, with $u(\cdot,t)>0$ an eigenfunction of $L_t\triangleq-\Delta_{\gamma(t)}+K_{\gamma(t)}$ with eigenvalue $\lambda_1(L_t)>0$.

Consider, now, the perturbed metric $g_\varepsilon\triangleq(1+\varepsilon t^2)\gamma(t)+u(x,t)^2dt^2$. Recall from Lemma A.1.6 that

$$
\begin{aligned}
R_{g_\varepsilon}&=2\lambda_1(L_t)-\left(\frac{1}{2u}\operatorname{Tr}_{(1+\varepsilon t^2)\gamma(t)}\left(\frac{\partial}{\partial t}(1+\varepsilon t^2)\gamma(t)\right)\right)^2\\
&\quad-\left\|\frac{1}{2u}\frac{\partial}{\partial t}(1+\varepsilon t^2)\gamma(t)\right\|^2_{(1+\varepsilon t^2)\gamma(t)}-\frac{2}{u}\frac{\partial}{\partial t}\left(\operatorname{Tr}_{(1+\varepsilon t^2)\gamma(t)}(1+\varepsilon t^2)\gamma(t)\right),
\end{aligned}
$$

and from the conclusion of Proposition 2.2.9, that $R_{g_\varepsilon}>0$ for $\varepsilon=0$. Since the minimum of R_{g_ε} on $\mathbf{S}^2\times[0,1]$ varies continuously with ε, there exists $\varepsilon_0>0$ such that $R_{g_\varepsilon}>0$ for all $\varepsilon\in(0,\varepsilon_0]$. Recall, moreover, that in the original metric $g=\gamma(t)+u(x,t)^2dt^2$ all the spheres $\mathbf{S}^2\times\{t\}$ are minimal.

Thus, in the new metric, we will have that

$$H_{\mathbf{S}^2\times\{t\},g_\varepsilon} = \frac{1}{2u}\operatorname{Tr}_{(1+\varepsilon t^2)\gamma(t)}\frac{\partial}{\partial t}(1+\varepsilon t^2)\gamma(t) = H_{\mathbf{S}^2\times\{t\},g} + \frac{2\varepsilon t}{(1+\varepsilon t^2)u} = \frac{2\varepsilon t}{(1+\varepsilon t^2)u},$$

i.e., $\mathbf{S}^2 \times \{0\}$ is minimal, and all remaining spheres are mean-convex.

By the construction in Proposition 2.2.9, there exists a round metric γ_* with area 4π and such that $\gamma(t) = \rho^2\gamma_*$ for all $t \in [2/3, 1]$. Moreover, since the first eigenfunctions of $L_* = -\Delta_{\gamma_*} + K_{\gamma_*}$ are constants, we may change variables

$$\mathbf{S}^2 \times [0,1] \ni (x,t) \mapsto (x,s) \in \mathbf{S}^2 \times [0,T],$$

so that, the region $t \in [2/3, 1]$ gets mapped to

$$g_\varepsilon = (1+\varepsilon T^{-2}s^2)\rho^2\gamma_* + ds^2,\ s \in [2T/3, T].$$

Write $f_\varepsilon(s) \triangleq (1+\varepsilon T^{-2}s^2)^{1/2}\rho$, $s \in [2T/3, T]$. Clearly,

$$f'_\varepsilon(s) = (1+\varepsilon T^{-2}s^2)^{-1/2}\varepsilon T^{-2}\rho s \tag{3.3.12}$$

and

$$\begin{aligned} f''_\varepsilon(s) &= -(1+\varepsilon T^{-2}s^2)^{-3/2}\varepsilon^2T^{-4}\rho s^2 + (1+\varepsilon T^{-2}s^2)^{-1/2}\varepsilon T^{-2}\rho \\ &= \frac{(1+\varepsilon T^{-2}s^2)\varepsilon T^{-2}\rho - \varepsilon^2 T^{-4}\rho s^2}{(1+\varepsilon T^{-2}s^2)^{3/2}} = (1+\varepsilon T^{-2}s^2)^{-3/2}\varepsilon T^{-2}\rho \end{aligned} \tag{3.3.13}$$

and

$$\frac{f'_\varepsilon(T)}{f_\varepsilon(T)} = \frac{(1+\varepsilon)^{-1/2}\varepsilon T^{-1}\rho}{(1+\varepsilon)^{1/2}\rho} = \frac{1}{T}\left(1 - \frac{1}{1+\varepsilon}\right). \tag{3.3.14}$$

In particular, from (3.3.14) we see that

$$\frac{\partial}{\partial\varepsilon}\arctan\left[\frac{1}{T}\left(1-\frac{1}{1+\varepsilon}\right)\right] > 0,$$

i.e., the slope traced out by the curve $(0,\varepsilon_0] \ni \varepsilon \mapsto \Gamma(\varepsilon) \triangleq (f_\varepsilon(T), f'_\varepsilon(T)) \in \mathbf{R}^2$, which lies on the first quadrant, is decreasing as $\varepsilon \downarrow 0$. The curve converges, as $\varepsilon \downarrow 0$, to the point

$$\lim_{\varepsilon\downarrow 0}\Gamma(\varepsilon) = (\rho, 0) = \left(\sqrt{\frac{\sigma_\gamma(\mathbf{S}^2)}{4\pi}}, 0\right) \in \mathbf{R}^2.$$

Now, let m be as in the statement of the theorem, i.e.,

$$m > \sqrt{\frac{\sigma_\gamma(\mathbf{S}^2)}{16\pi}}. \tag{3.3.15}$$

Notice that the curve $0 < s \mapsto \Delta(s) \mapsto (u_m(s), u_m'(s))$, which also lies on the first quadrant, converges as $s \downarrow 0$ to

$$\lim_{s\downarrow 0} \Delta(s) = (2m, 0) \in \mathbf{R}^2.$$

This point lies to the right of the smooth curve Γ by (3.3.15). For small enough $s_0 > 0$, the smooth compact curve $[0, s_0] \ni s \mapsto \Delta(s)$ lies strictly below and to the right of the smooth compact curve $[0, \varepsilon_0] \ni \varepsilon \mapsto \Gamma(\varepsilon)$. Applying Lemma 3.3.10 above, we modify the curve $\Delta(s)$ on a small time interval $[s_0 - \delta, s_0]$ and call the new curve $[s_0 - \delta, s_0] \ni s \mapsto \widetilde{\Delta}(s) = (u_m(\sigma(s)), (u_m(\sigma(s)))')$. For small enough $\delta > 0$, by continuity of dependence, the curve $\widetilde{\Delta}$ will continue to be below and to the right of Γ.

Finally, we apply Lemma 2.2.10 to form a bridge of positive scalar curvature between g_ε, $s \in [2T/3, T]$, for some $\varepsilon \in (0, \varepsilon_0]$, and the bent Schwarzschild metric above. We note that:

(1) The role of f_2 in the lemma will be played by $[s_0 - \delta, s_0] \ni s \mapsto u_m(\sigma(s))$. Condition (1) of the lemma is satisfied by virtue of (3.3.11), after possibly shrinking $\delta > 0$ to retain strict convexity. Condition (2) is precisely guaranteed to hold by Lemma 3.3.10, which bends thin annular regions in Schwarzschild metrics to positive scalar curvature.

(2) The role of f_1 will be played by $[2T/3, T] \ni s \mapsto f_\varepsilon(s)$, for an $\varepsilon \in (0, \varepsilon_0]$ that is to be determined. Condition (1) is satisfied by virtue of (3.3.12)-(3.3.13). Condition (2) is guaranteed to hold since $R_{g_\varepsilon} > 0$.

Having already chosen δ, we choose $\varepsilon \in (0, \varepsilon_0]$ so that $\Gamma(\varepsilon)$ and $\widetilde{\Delta}(s_0 - \delta)$ have the same vertical coordinate in $\mathbf{R}^2$. There is a unique such choice in view of (3.3.14). This fulfills the matching slope portion of hypothesis (3) in Lemma 2.2.10. The remaining portion of that hypothesis is true because Γ lies to the left of $\widetilde{\Delta}$.

The gluing lemma, then, creates a bridge $f(s)^2\gamma_* + ds^2$ with positive scalar curvature that preserves the $t \in [0, 2/3]$ region and the $s \in [s_0, \infty)$ exterior Schwarzschild region. It also guarantees a mean-convex foliation, since Lemma 2.2.10 introduced no new critical points. The result follows. $\square$

3.4 Thorne's hoop conjecture

3.4.1 Formulation

Conjecture 3.4.1 (Thorne's hoop conjecture, cf. Gibbons [26, Section 1])**.** *Horizons form if and only if a mass m gets compacted into a region whose circumference in every direction is $\leq 4\pi m$.*

This conjecture has been invoked in numerical relativity and studies of hole scattering in four and higher dimensions, and it has been suggested that it may provide a route to a precise formulation of the idea that there is a minimal length in quantum gravity. ([14])

The notion of "circumference" in this conjecture is open to interpretation. Reformulating this conjecture in a mathematically precise manner is an issue of mathematical and physical significance. Gibbons posed the following mathematically precise formulation of the conjecture in [26] with the Birkhoff invariant of the horizon playing the role of circumference.

Definition 3.4.2 (Birkhoff invariant, cf. Gibbons [26, Section 2])**.** *We define the Birkhoff invariant of $(\mathbf{S}^2, \gamma)$ to be*

$$\beta(\mathbf{S}^2, \gamma) = \inf \left\{ \sup_{c \in \mathbf{R}} \ell_\gamma(f^{-1}(c)) : f \in C^\infty(\mathbf{S}^2) \textit{ with two critical points} \right\},$$

where ℓ_γ denotes the one-dimensional Hausdorff measure induced by γ.

Given Definition 3.4.2, Gibbons's formulation of the hoop conjecture takes the form:

Conjecture 3.4.3 (Gibbons [26, Section 2])**.** *Every outermost marginally trapped $(\mathbf{S}^2, \gamma)$ lying in a Cauchy hypersurface satisfying the dominant energy condition must have $\beta(\mathbf{S}^2, \gamma) \leq 4\pi \mathfrak{m}_{\mathrm{ADM}}$.*

In addition to giving a mathematically precise formulation of the conjecture, [26] shows that Conjecture 3.4.3 holds on general charged rotating Kerr-Newman black holes. More generally, 3.4.3 has been checked to hold true on: the horizons of all four-charged rotating black hole solutions of ungauged supergravity theories, allowing for the presence of a negative cosmological constant; for multi-charged rotating solutions in gauged supergravity; and for the Ernst-Wild static black holes immersed in a magnetic field which are asymptotic to the Melvin solution [14].

Let $\lambda(\mathbf{S}^2, \gamma)$ denote the length of the shortest nontrivial geodesic of $(\mathbf{S}^2, \gamma)$. It follows from [8] that $\beta(\mathbf{S}^2, \gamma) \geq \lambda(\mathbf{S}^2, \gamma)$. In particular, if Conjecture 3.4.3 is true, it will be true with $\beta(\mathbf{S}^2, \gamma)$ replaced by $\lambda(\mathbf{S}^2, \gamma)$. It was shown in [26, Section 6.1], by using Pu's systolic inequality for $\mathbf{RP}^2$, that Conjecture 3.4.3 holds with the Birkhoff invariant $\beta(\mathbf{S}^2, \gamma)$ replaced by $\lambda(\mathbf{S}^2, \gamma)$, as long as $(\mathbf{S}^2, \gamma)$ has an antipodal $\mathbf{Z}_2$ symmetry:

Theorem 3.4.4 ([26, Section 6.1])**.** *Let (M^3, g) be complete, asymptotically flat, with nonnegative scalar curvature, one end, and outermost minimal boundary $\partial(M^3, g) \cong (\mathbf{S}^2, \gamma)$, which is a Riemannian cover of an $(\mathbf{RP}^2, \gamma)$. Then $\lambda(\mathbf{S}^2, \gamma) \leq 4\pi \mathfrak{m}_{\mathrm{ADM}}(M^3, g)$.*

3.4.2 Results

Using our optimal computation of $\mathfrak{m}_{\mathrm{B}}(\mathbf{S}^2, \gamma, H = 0)$, we disprove Gibbons' formulation of Thorne's hoop conjecture, by reducing it to a false statement in systolic geometry. Results in this section appear in joint work with R. Schoen published in [40].

Theorem 3.4.5 (M.-Schoen [40, Theorem 4.2])**.** *There exists a complete, asymptotically flat (M^3, g) with nonnegative scalar curvature, one end, totally geodesic boundary $\partial(M^3, g) \cong (\mathbf{S}^2, \gamma)$, that is precisely Schwarzschild outside a compact set, but such that $\lambda(\mathbf{S}^2, \gamma) > 4\pi \mathfrak{m}_{\mathrm{ADM}}(M, g)$.*

Our counterexample is motivated by a construction that is originally due to Croke [13] and relies on our optimal Bartnik mass computation for apparent horizons in Theorem 3.3.5. As was mentioned in Remark 3.3.9, while we only construct initial data sets here, one can easily extend those to spacetimes with Cauchy hypersurfaces that satisfy the same conclusions.

The proof of Theorem 3.4.5 requires:

Lemma 3.4.6 (M.-Schoen [40, Lemma 4.1])**.** *For any $\delta > 0$, there exists $(\mathbf{S}^2, \gamma)$ with $\lambda(\mathbf{S}^2, \gamma) \geq 2\sqrt{3} - \delta$ and $\sigma_\gamma(\mathbf{S}^2) \leq 2\sqrt{3} + \delta$.*

We first describe the metric. Consider the regular hexagon with vertices $\{2e^{\frac{k\pi i}{3}}\}_{k=0,\ldots,5} \subset \mathbf{C}$. It is a fundamental domain for the torus $\Sigma_0 \triangleq \mathbf{C}/\Lambda$, where Λ is the lattice generated by $\omega_1 = 2\sqrt{3}e^{-\frac{\pi i}{6}}$ and $\omega_2 = 2\sqrt{3}e^{\frac{\pi i}{6}}$. Now, the rotation $R(z) \triangleq e^{\frac{2\pi i}{3}} z$ is well-defined on the torus and fixes the equivalence classes mod Λ of three points: 0, 2, and $2e^{\frac{\pi i}{3}}$. The group generated by R is $\cong \mathbf{Z}_3$, so we denote it as $\mathbf{Z}_3$. Let $\Sigma = \Sigma_0/\mathbf{Z}_3$.

From the construction, we can see that Σ may be thought of as two equilateral triangles of side length 2, adjacent along a common edge, with the two sides from each of the common vertices identified in the direction from the vertex. We see that $\mathbf{S}^2$ is a topological 2-sphere with a singular metric which is smooth and flat away from three points, each of which is a cone point with cone angle $\frac{2\pi}{3}$. Note that the area of Σ is equal to $2\sqrt{3}$.

Proof of Lemma 3.4.6. We first observe that a cone point of angle $\frac{2\pi}{3}$ can be approximated by a smooth metric with nonnegative curvature, which agrees with the original cone metric outside any chosen ball centered at the vertex. Furthermore, we can take the approximating metric to be rotationally symmetric about a point and such that the square of the distance function from that point is a strictly convex function. This can be seen, for example, by isometrically embedding the cone into $\mathbf{R}^3$ as the graph of the function

$$(x, y) \mapsto \alpha\sqrt{x^2 + y^2},$$

with $\alpha = 2\sqrt{2}$ chosen so that the cone angle coincides with $\frac{2\pi}{3}$. We can then smooth the function, while keeping it convex and rotationally symmetric. By the Gauss-Bonnet theorem, the total Gauss curvature of a ball whose boundary is in the flat region is $\frac{4\pi}{3}$, and since smaller balls have smaller total curvature, the Gauss-Bonnet theorem implies that the geodesic curvature of the boundary of any rotationally symmetric ball is a positive constant. Thus, the square of the distance function from the center of symmetry is strictly convex.

Using the construction above, we may approximate the metric defining Σ by a metric γ on $\mathbf{S}^2$ which is smooth, has nonnegative curvature, and agrees with the metric of Σ outside a small neighborhood U of the cone points. It remains to show that, if the U is chosen small enough, we will have

$$\sigma_\gamma(\mathbf{S}^2) \leq 2\sqrt{3} + \delta \tag{3.4.1}$$

and

$$\lambda(\mathbf{S}^2, \gamma) \geq 2\sqrt{3} - \delta. \tag{3.4.2}$$

The first inequality, (3.4.1), is clear since the area of the singular surface Σ is $2\sqrt{3}$. We now show that the length of any geodesic of γ is at least $2\sqrt{3} - \delta$ if U is chosen small enough.

First, we observe that there can be no closed geodesic lying entirely in the nonflat region near a cone point. This follows from the convexity of the distance function and the maximum principle. The lemma will then follow from the following:

Claim. Let Γ be a smooth closed curve which intersects the flat portion of γ and is a geodesic there. Then $\ell_\gamma(\Gamma) \geq 2\sqrt{3} - \delta$, if the neighborhood U is chosen small enough.

Proof of claim. We consider three cases:

(1) Suppose Γ lies entirely in the flat region.

In this case, we lift Γ beginning at some initial point p to the flat torus Σ_0. The lifted curve $\widetilde{\Gamma}$ is then a geodesic extending from the chosen lift $\widetilde{p}$ to a point which is the image of $\widetilde{p}$ under an element of the rotation group $\mathbf{Z}_3$. Since $\widetilde{\Gamma}$ is a geodesic, its tangent vector field $\widetilde{\Gamma}'$ extends as a parallel vector field on Σ_0. Since the nonidentity elements of $\mathbf{Z}_3$ do not fix this parallel vector field, it follows that the curve $\widetilde{\Gamma}$ is a closed geodesic of Σ_0 and, therefore, has length $\geq 2\sqrt{3}$, the shortest translation distance in the lattice $\Lambda \subset \mathbf{C}$.

(2) The second case we consider is when Γ intersects precisely one component of the neighborhood U of the cone points.

In this case, we again lift Γ beginning at the neighborhood of the cone point to a curve $\widetilde{\Gamma}$ in Σ_0. Since there is only one lift of the cone point to Σ_0, it follows that the initial and final endpoint of $\widetilde{\Gamma}$ lie in a small neighborhood of the same point. Since the minimum translation distance in the lattice $\Lambda \subset \mathbf{C}$ is $2\sqrt{3}$, we again have $\ell_\gamma(\widetilde{\Gamma}) \geq 2\sqrt{3} - \delta$ if U is chosen small enough.

(3) Finally, we consider the case in which Γ intersects more than one cone neighborhood.

In this case, the length of Γ is $\geq 4 - \delta$ if U is chosen small enough. This is because the distance between cone points on Σ is equal to 2 and the curve must have at least two segments extending from one cone neighborhood to another, since it is a closed curve.

$\square$

The claim concludes the proof of Lemma 3.4.6. $\square$

Proof of Theorem 3.4.5. Let δ be given. Suppose $(\mathbf{S}^2, \gamma)$ is as in Lemma 3.4.6, with $\sigma_\gamma(\mathbf{S}^2) \leq 2\sqrt{3}+\delta$ and $\lambda(\mathbf{S}^2, \gamma) \geq 2\sqrt{3} - \delta$. Note that

$$\frac{\lambda(\mathbf{S}^2, \gamma)^2}{\sigma_\gamma(\mathbf{S}^2)} \geq \frac{(2\sqrt{3} - \delta)^2}{2\sqrt{3} + \delta} > \pi,$$

as long as $\delta > 0$ is small enough.

Since $K_\gamma \geq 0$, it follows from Proposition 2.2.4 that $(\mathbf{S}^2, \gamma) \in \mathscr{M}_+$. Therefore, by (3.3.3) in Theorem 3.3.5, there exists a$(M^3, g) \in \mathcal{E}_{(\mathbf{S}^2, \gamma, H=0)}$ in which $(\mathbf{S}^2, \gamma)$ is a totally geodesic apparent horizon (see Remark 3.3.8), and such that

$$\frac{\lambda(\mathbf{S}^2, \gamma)^2}{16\pi \mathfrak{m}_{\mathrm{ADM}}(M, g)^2} > \pi.$$

The result follows by simply rearranging the inequality above and taking square roots. $\square$

3.5 Variational analog of Brown-York quasi-local mass

3.5.1 Definitions

In 1993, Brown and York ([10]) formulated an alternative notion of quasi-local mass in general relativity by employing a Hamilton-Jacobi analysis of the Einstein-Hilbert action.

Definition 3.5.1 (Brown-York quasi-local mass, Brown-York [10])**.** *Given a compact spacelike hypersurface Ω in spacetime, assuming its boundary $\partial\Omega$ is a 2-sphere with positive Gauss curvature, the Brown-York mass of $\partial\Omega$ is given by*

$$\mathfrak{m}_{\mathrm{BY}}(\Omega, g) = \frac{1}{8\pi} \int_{\partial\Omega} (H_0 - H_g)\, d\sigma. \tag{3.5.1}$$

Here, H_0 is the mean curvature of the isometric embedding of $\partial\Omega$ into flat Euclidean space, $(\mathbf{R}^3, \delta)$.

The existence and uniqueness of such an embedding of $\partial\Omega$ is guaranteed by the solution to Weyl's embedding problem ([47, 56]), because $\partial\Omega$ has been assumed to have positive Gauss curvature.

The Brown-York quasi-local mass has two properties one would require of a quasi-local mass functional:

(1) Nonnegativity: if (Ω^3, g) is as in Definition 3.5.1, then $\mathfrak{m}_{\mathrm{BY}}(\Omega, g) \geq 0$, with equality if and only if (Ω^3, g) is a convex bounded domain in $(\mathbf{R}^3, \delta)$. (See Theorem 3.5.2 below.)

(2) Exhaustion: if (M^3, g) is complete, asymptotically flat, with nonnegative scalar curvature and no boundary, then $\lim_{r\uparrow\infty} \mathfrak{m}_{\mathrm{BY}}(M \setminus \Phi^{-1}(\{|\mathbf{x}| > r\}), g) = \mathfrak{m}_{\mathrm{ADM}}(M^3, g)$ for any asymptotically flat chart Φ. (See [19, Theorem 1.1].)

Given Definition 3.5.1, the following question is natural:

Question. *Can one make sense of the Brown-York quasi-local mass without a positive Gauss curvature assumption?*

In [64], Shi and Tam proved the following theorem, which implies the positivity of $\mathfrak{m}_{\mathrm{BY}}(\Omega)$ when Ω has nonnegative scalar curvature. Recall that this is always the case with us, by Lemma **??**, since we are assuming our initial data sets satisfy assumptions (DEC) and (TS).

Theorem 3.5.2 (Shi-Tam [64, Theorem 1])**.** *Let (Ω^3, g) be a compact three-dimensional Riemannian manifold with nonnegative scalar curvature, and with nonempty boundary $\partial\Omega$. Suppose $\partial\Omega$ has finitely many components Σ_j, $j = 1, \ldots, k$, so that each Σ_j is a topological 2-sphere which has positive Gauss curvature and positive mean curvature $H_{g,j}$. Then*

$$\int_{\Sigma_j} H_{g,j}\, d\sigma_j \leq \int_{\Sigma_j} H_{0,j}\, d\sigma_j, \tag{3.5.2}$$

where $H_{0,j}$ is the mean curvature of the isometric embedding of Σ_j in $\mathbf{R}^3$. Moreover, equality in (3.5.2) *holds for some Σ_j if and only if $\partial\Omega$ has a unique connected component and (Ω^3, g) is isometric to a convex domain in $\mathbf{R}^3$.*

The geometric significance of inequality (3.5.2) is in that its right side is a quantity that is determined only by the induced metric on the boundary component $\Sigma_j \subseteq \partial\Omega$; it is independent of the interior of the manifold, and independent of all other boundary components. The motivation behind Section 2.3 and the study of total boundary mean curvature maximization among mean-convex fillins with nonnegative scalar curvature is that only the integral quantity

$$\int_{\partial\Omega} H_0 \, d\sigma$$

is of actual interest for the purposes of the actual definition of Brown-York mass in (3.5.2), and not the pointwise quantity H_0. Moreover, when Σ has positive Gauss curvature, Theorem 3.5.2 characterizes this integral quantity as

$$\int_{\Sigma} H_0 \, d\sigma = \sup_{(M,h)\in\mathcal{F}} \int_{\partial M} H_h \, d\sigma,$$

over an appropriate class $\mathcal{F}$ of fill-ins. More precisely, in the notation of Section 2.3, Theorem 3.5.2 asserts that when (Σ_j^2, γ_j), $j = 1, \dots, k$ are 2-spheres endowed with metrics of positive Gauss curvature, then for every $(\Omega^3, g) \in \mathcal{F}_{\{(\Sigma,\gamma_j)\}_{j=1,\dots,k}}$,

$$\int_{\Sigma_j} H_{g,j} \, d\sigma_j \leq \Lambda_{(\Sigma_j,\gamma_j)} = \int_{\Sigma_j} H_{0,j},$$

with equality for some j if and only if $k = 1$ and (Ω, g) is the unique flat fill-in $(B^3, \delta) \in \mathcal{F}_{(\Sigma_1,\gamma_1)}$ coming from Weyl's embedding theorem. Section 2.3 can be viewed as a geometric continuation of this idea.

Some results in this section have appeared in work published in [39].

3.5.2 Results

All the facts above suggest the following variational analog of this quasi-local mass that does not require positivity of the Gauss curvature on the boundary:

Definition 3.5.3 (M.-Miao [39, Theorem 1.5])**.** *Given a compact, connected (Ω^3, g) with nonnegative scalar curvature, and a mean-convex boundary Σ^2 which is a topological 2-sphere, define*

$$\mathfrak{m}(\Sigma^2, \Omega^3, g) = \frac{1}{8\pi}\left(\mathring{\Lambda}_{(\Sigma,\gamma)} - \int_{\Sigma} H_g \, d\sigma\right). \tag{3.5.3}$$

Here γ is the metric induced on Σ by g, and, $\mathring{\Lambda}_{(\Sigma,\gamma)}$ is as in Definition 3.5.4 below.

Definition 3.5.4 (Fill-ins, II, M.-Miao [39, Definition 1.2])**.** *Let (Σ^2, γ) be a closed, orientable two-dimensional Riemannian manifold. Recalling the definition of $\mathring{\mathcal{F}}_{(\Sigma,\gamma)}$ from Definition 2.3.7, we* *define*

$$\mathring{\Lambda}_{(\Sigma,\gamma)} \triangleq \sup\left\{\int_{\Sigma_O} H_g \, d\sigma : (\Omega, g) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)}\right\}.$$

It is not hard to see that $\mathcal{F}_{(\Sigma,\gamma)} \subseteq \mathring{\mathcal{F}}_{(\Sigma,\gamma)}$, and therefore $\Lambda_{(\Sigma,\gamma)} \leq \mathring{\Lambda}_{(\Sigma,\gamma)}$. We verify in Proposition 3.5.5 that, in fact, these two scalar quantities coincide. We nevertheless choose to use the enlarged class of fill-ins in the definition of $\mathfrak{m}(\Sigma, \Omega, g)$ because it is more suitable for discussion on quasi-local mass. For an element $(\Omega, g) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)}$, the portion Σ_H of the boundary, if nonempty, represents horizons of black holes detected by observers at (the outer boundary) Σ_O. For example, the region bounded by a rotationally symmetric sphere $(\mathbf{S}^2, \gamma)$ of positive mean curvature and the horizon in a Schwarzschild manifold (see Example 3.1.6) of positive mass is now a valid fill-in of $(\mathbf{S}^2, \gamma)$, while it wasn't an element of the original class of fill-ins, $\mathcal{F}_{(\Sigma,\gamma)}$.

Proposition 3.5.5 (M.-Miao [39, Proposition 5.1])**.** *For any closed, orientable* (Σ^2, γ)*,*

$$\Lambda_{(\Sigma,\gamma)} = \mathring{\Lambda}_{(\Sigma,\gamma)}.$$

Proof. It suffices to show $\mathring{\Lambda}_{(\Sigma,\gamma)} \leq \Lambda_{(\Sigma,\gamma)}$, i.e., that

$$\sup\left\{\int_{\partial\Omega} H \, d\sigma : (\Omega^3, g) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)} \text{ with } \Sigma_H \neq \emptyset\right\} \leq \Lambda_{(\Sigma,\gamma)}.$$

Pick any such $(\Omega^3, g) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)}$ with $\Sigma_H \neq \emptyset$. Let h be the metric given by the doubling lemma (Lemma 2.3.10) on the doubling D of Ω across Σ_H. Then $H_h > H_g$ on Σ_O. In combination with (2.3.1) of Theorem 2.3.5 applied to (D^3, h), on Σ_O, we have

$$\int_{\partial\Omega} H_g \, d\sigma = \int_{\Sigma_O} H_g \, d\sigma < \int_{\Sigma_O} H_h \, d\sigma \leq \Lambda_{(\Sigma,\gamma)},$$

and the result follows. □

We now prove that $\mathfrak{m}(\Sigma, \Omega, g)$ is a generalization of Brown-York quasi-local mass.

Theorem 3.5.6 (M.-Miao [39, Theorem 1.5])**.** *Let* (Σ^2, γ) *be given, with* Σ *a 2-sphere. For every* $(\Omega^3, g) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)}$*:*

(1) $0 \leq \mathfrak{m}(\Sigma, \Omega, g) < \infty$*, and* $\mathfrak{m}(\Sigma, \Omega, g) = 0$ *only if* (Ω, g) *is a flat 3-ball;*

(2) if (Σ, γ) *has positive Gauss curvature, then*

$$\mathfrak{m}(\Sigma, \Omega, g) = \frac{1}{8\pi}\int_{\Sigma} (H_0 - H_g) \, d\sigma;$$

i.e., if $\partial\Omega = \Sigma$*, then* $\mathfrak{m}(\Sigma, \Omega, g) = \mathfrak{m}_{\mathrm{BY}}(\Omega, g)$*. Here* H_0 *is as in Definition 3.5.1.*

Proof. Let us first show (1). Since $(\Omega, g) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)}$, the finiteness assertion is a direct consequence of Theorem 2.3.3 (1) and Proposition 3.5.5. Nonnegativity is a consequence of the simple fact

$$8\pi\mathfrak{m}(\Sigma, \Omega, g) = \mathring{\Lambda}_{(\Sigma,\gamma)} - \int_\Sigma H_g \, d\sigma = \left(\sup_{(M,h)\in\mathcal{F}_{(\Sigma,\gamma)}} \int_\Sigma H_h \, d\sigma \right) - \int_\Sigma H_g \, d\sigma \geq 0.$$

Let us study the equality case. We need to show that (Ω^3, g) is a flat 3-ball. We will first show that $\partial\Omega = \Sigma_O$, i.e., there are no minimal surfaces in $\partial\Omega$. Once that claim has been established, it follows that $(\Omega, g) \in \mathcal{F}_{(\Sigma,\gamma)}$ and (1) will be a consequence of Theorem 2.3.3 (2).

Claim. If $\mathfrak{m}(\Sigma, \Omega, g) = 0$, then $\partial\Omega = \Sigma_O$.

Proof of claim. We argue by contradiction. If $\partial\Omega \neq \Sigma_O$, then Σ_H would be nonempty. We may, therefore, employ the doubling lemma (Lemma 2.3.10) to double (Ω, g) across Σ_H. Write (D^3, h) for the doubled manifold, whose boundary satisfies $\partial D = S \sqcup S'$, where S is the image of Σ and S' is its mirror image.

By conclusion (1) of the doubling lemma, the metric induced on S by (D^3, h) is still γ, though the metric induced on S' may have changed. By conclusion (2) of the lemma, the resulting doubled manifold (D, h) is going to have

$$H_{S,h} > H_{\Sigma_O,g}. \tag{3.5.4}$$

Likewise, by conclusion (3) of the lemma,

$$H_{S',h} > 0 \tag{3.5.5}$$

if the lemma parameter $\eta > 0$ is chosen sufficiently small.

Let $\phi < 0$ be an arbitrary negative function on Ω, and let w be the unique solution of

$$\Delta_h w - \frac{1}{8} R_h w = \phi \text{ in } D \text{ and } w = 0 \text{ on } \partial D.$$

Let $\varepsilon > 0$ be given. For small $\tau > 0$, consider the conformally transformed Riemannian metrics $h^{(\tau)} = (1 + \tau w)^4 h$ on D. By the conformal transformation of scalar curvature formula in Lemma A.1.7 we see that $R_{h^{(\tau)}} > 0$ and, as long as $\tau > 0$ is small, the boundary mean curvatures will not vary by more than $\varepsilon > 0$. In view of (3.5.4)-(3.5.5), it will be sufficient to pick

$$\varepsilon < \min\left\{ \frac{1}{\sigma_\gamma(S)} \int_S (H_h - H_g) \, d\sigma, \, H_h, \, H'_h \right\} > 0,$$

so that

$$H_{S,h^{(\tau)}} > 0 \text{ on } S, \text{ and } \int_S H_{S,h^{(\tau)}}\, d\sigma > \int_{\Sigma_O} H_{\Sigma_O,g}\, d\sigma, \tag{3.5.6}$$

and

$$H_{S',h^{(\tau)}} > 0. \tag{3.5.7}$$

At this stage we finalize any small enough choice of $\tau > 0$ so that (3.5.6)-(3.5.7) are true, and cut the resulting $(D, h^{(\tau)})$ using the cutting lemma (Lemma 2.3.8) to isolate its boundary component S. The resulting manifold may be denoted $(M, h^{(\tau)})$. Note that $(M, h^{(\tau)}) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)}$. By (3.5.6),

$$\int_{\Sigma_O} H_{\Sigma_O,g}\, d\sigma < \int_S H_{S,h^{(\tau)}}\, d\sigma \leq \mathring{\Lambda}_{(\Sigma,\gamma)} \implies 8\pi\mathfrak{m}(\Sigma,\Omega,g) = \mathring{\Lambda}_{(\Sigma,\gamma)} - \int_{\Sigma_O} H_{\Sigma_O,g}\, d\sigma > 0,$$

in direct contradiction to $\mathfrak{m}(\Sigma,\Omega,g) = 0$. □

Next, we prove (2). Theorem 2.3.3 (2) and Proposition 3.5.5 combined tell us that

$$8\pi\mathfrak{m}(\Sigma,\Omega,g) = \mathring{\Lambda}_{(\Sigma,\gamma)} - \int_\Sigma H\, d\sigma = \Lambda_{(\Sigma,\gamma)} - \int_\Sigma H\, d\sigma = \int_\Sigma H_0\, d\sigma - \int_\Sigma H\, d\sigma.$$

This completes the proof of Theorem 3.5.6. □

Remark 3.5.7. *It seems a challenging question to check whether mean-convex domains $\Omega \subset \mathbf{R}^3$, with $\Sigma = \partial\Omega$ a 2-sphere, necessarily maximize the total mean curvature on their boundary relative to all competitors in $\mathring{\mathcal{F}}_{(\Sigma,\delta|_\Sigma)}$; i.e., is $\mathfrak{m}(\Sigma,\Omega,g) = 0$ for all mean-convex domains $\Omega \subset \mathbf{R}^3$ bounded by a 2-sphere Σ? If Σ is strictly convex in $\mathbf{R}^3$, then this is the conclusion of (2) in Theorem 3.5.6.*

We present one final way in which the $\mathfrak{m}(\Sigma,\Omega,g)$ functional that generalizes the Brown-York quasi-local mass functional. By work of Shi-Tam and Miao, it is known that the Brown-York mass is an upper bound for the so-called Hawking quasi-local mass. In the notation of this thesis, the known result is:

Theorem 3.5.8 (Miao [43, Theorem 3], Shi-Tam [65, Theorem 1.1])**.** *Suppose (Σ^2, γ) is a 2-sphere with positive Gauss curvature. If $(\Omega^3, g) \in \mathcal{F}_{(\Sigma,\gamma)}$, then*

$$\mathfrak{m}_{\mathrm{BY}}(\Sigma,\Omega,g) \geq \sqrt{\frac{\sigma_\gamma(\Sigma)}{16\pi}} \left(1 - \frac{1}{16\pi}\int_\Sigma H_g^2\, d\sigma_\gamma\right)$$

with equality if and only if (Ω^3, g) is a geodesic ball in $(\mathbf{R}^3, \delta)$.

The generalization of this result, to appear in [38], is:

Proposition 3.5.9 (M.-Miao [38, Proposition 3.1])**.** *Suppose (Σ^2, γ) is a 2-sphere that embeds* *isometrically as an outer-minimizing, mean-convex sphere in* $(\mathbf{R}^3, \delta)$*. If* $(\Omega^3, g) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)}$*, then*

$$\mathfrak{m}(\Sigma, \Omega, g) \geq \sqrt{\frac{\sigma_\gamma(\Sigma_O)}{16\pi}} \left(1 - \frac{1}{16\pi} \int_{\Sigma_O} H_g^2 \, d\sigma_\gamma\right)$$

with equality if and only if (Ω^3, g) *is a geodesic ball in* $(\mathbf{R}^3, \delta)$*.*

Proof. We have

$$8\pi\mathfrak{m}(\Sigma, \Omega, g) = \mathring{\Lambda}_{(\Sigma,\gamma)} \left(1 - \frac{1}{\mathring{\Lambda}_{(\Sigma,\gamma)}} \int_{\Sigma_O} H_g \, d\sigma_\gamma\right) \geq \frac{\mathring{\Lambda}_{(\Sigma,\gamma)}}{2} \left[1 - \left(\frac{1}{\mathring{\Lambda}_{(\Sigma,\gamma)}} \int_{\Sigma_O} H_g \, d\sigma_\gamma\right)^2\right], \quad (3.5.8)$$

where the inequality in (3.5.8) follows from the algebraic inequality $1 - x \geq \frac{1}{2}(1 - x^2)$, $x \in \mathbf{R}$. Notice that the right hand side is nonnegative, by the definition of $\mathring{\Lambda}_{(\Sigma,\gamma)}$.

From Proposition 3.5.5 and conclusion (4) of Theorem 2.3.3,

$$\mathring{\Lambda}_{(\Sigma,\gamma)} = \Lambda_{(\Sigma,\gamma)} \geq \sqrt{16\pi\sigma_\gamma(\Sigma)}. \quad (3.5.9)$$

Then, we first estimate the leading $\mathring{\Lambda}_{(\Sigma,\gamma)}$ in (3.5.8) from below by (3.5.9), and then estimate the squared term by Hölder's inequality and (3.5.9) again:

$$8\pi\mathfrak{m}(\Sigma, \Omega, g) \geq \frac{\sqrt{16\pi\sigma_\gamma(\Sigma)}}{2} \left(1 - \frac{\sigma_\gamma(\Sigma)}{\mathring{\Lambda}^2_{(\Sigma,\gamma)}} \int_{\Sigma_O} H_g^2 \, d\sigma_\gamma\right) = \sqrt{4\pi\sigma_\gamma(\Sigma)} \left(1 - \frac{1}{16\pi} \int_{\Sigma_O} H_g^2 \, d\sigma_\gamma\right).$$

The desired inequality follows after dividing through by 8π.

Let us study the equality case. This forces all inequalities to be equalities; i.e.,

(1) equality in (3.5.8) means $(\Omega^3, g) \in \mathring{\mathcal{F}}_{(\Sigma,\gamma)}$ attains $\mathring{\Lambda}_{(\Sigma,\gamma)}$, which, by (1) of Theorem 3.5.6, means (Ω^3, g) is a flat 3-ball;

(2) equality in (3.5.9) means, by conclusion (4) of Theorem 2.3.3, that (Σ^2, γ) is round.

In particular, Brown-York mass is well-defined, and $\mathfrak{m}_{\mathrm{BY}}(\Omega, g) = \mathfrak{m}(\Sigma, \Omega, g) = 0$. By the rigidity case of Theorem 3.5.2, (Ω^3, g) is isometric to a geodesic ball in $(\mathbf{R}^3, \delta)$. □

Remark 3.5.10. *By the maximum principle, we see that this applies to all* (Σ^2, γ) *that embed isometrically as starshaped mean-convex spheres in* $(\mathbf{R}^3, \delta)$*. Conjecturally, Proposition 3.5.9 holds for all 2-spheres* (Σ^2, γ)*.*

Chapter 4

Phase Transitions

4.1 Introduction

4.1.1 The equation and its variational structure

The Allen-Cahn equation is a linear elliptic partial differential equation that describes phase separation in multi-component alloy systems, and is given by:

$$\Delta u = W'(u), \tag{4.1.1}$$

where W is a double-well energy potential. (See Definition 4.1.1.) Its geometric significance is founded on De Giorgi's conjectures [15] about how the level sets of solutions u of (4.1.1) will behave like minimal hypersurfaces. The conjectures and their current status are outlined in subsection 4.1.3.

Definition 4.1.1. *A map $W : \mathbf{R} \to \mathbf{R}$ is called a double well potential provided:*

(1) W is nonnegative, and vanishes at its two global minima $t = \pm 1$:

$$W \geq 0,\ W(t) = 0 \iff t = \pm 1; \tag{H1}$$

(2) W has a unique, nondegenerate, critical point between its global minima, at $t = 0$:

$$tW'(t) < 0 \text{ for } 0 < |t| < 1, \text{ and } W''(0) \neq 0; \tag{H2}$$

(3) W is strictly convex near ± 1:

$$W''(t) \geq \kappa > 0 \text{ for } |t| > 1 - \alpha,\ \alpha \in (0,1); \tag{H3}$$

Example 4.1.2. *The standard double-well potential is $W(t) = \frac{1}{4}(1-t^2)^2$, and the corresponding equation* (4.1.1) *is $\Delta u = u^3 - u$.*

Remark 4.1.3. *There is nothing special about $t = -1,\ 0,\ 1$; one can replace these points by any other triple points $t_1 < t_2 < t_3$ in Definition 4.1.1, and everything will continue to hold just the same, up to relabeling.*

We now turn to the variational nature of (4.1.1). The equation in question arises as the Euler-Lagrange equation for the functional

$$E[u] \triangleq \int_M \left(\frac{1}{2}\|\nabla u\|^2 + W(u)\right) dv_g, \tag{4.1.2}$$

i.e., functions $u \in C^\infty_{\text{loc}}(M^n)$ that are zeroes of the **first variation** operator $\delta E[u] : C^\infty_c(M^n) \to \mathbf{R}$,

$$\delta E[u]\{\zeta\} \triangleq \int_M (\langle \nabla u, \nabla \zeta\rangle + W'(u)\zeta)\ dv_g. \tag{4.1.3}$$

4.1.2 Standard technical results on $\mathbf{R}^n$

One of the causes of interesting behavior of solutions to the Allen-Cahn equation is the particular saddle behavior of the double-well potential in its non-convex portion; in some sense the behavior of solutions to the Allen-Cahn is least interesting inside the convex portions of the functional.

We isolate a useful computation as a lemma that will be invoked often throughout this chapter.

Lemma 4.1.4 (M.). *Let $\rho_0,\ \gamma > 0$. The test function $w : \mathbf{R}^n \to \mathbf{R}$ given by*

$$w(\mathbf{x}) = \exp\left(\gamma(\|\mathbf{x}\|^2 + \rho_0^2)^{1/2}\right)$$

satisfies

$$w^{-1}\Delta w = \gamma^2 + \frac{\gamma}{(r^2+\rho_0^2)^{3/2}}\left((n-\gamma)\|\mathbf{x}\|^2 + \gamma(\|\mathbf{x}\|^2+\rho_0^2)^{1/2} + n\rho_0^2\right). \tag{4.1.4}$$

Proof. This is a straightforward computation. Writing $r = \|\mathbf{x}\|$, we have:

$$\begin{aligned} w^{-1}\Delta w &= \gamma \operatorname{div} \nabla (r^2 + \rho_0^2)^{1/2} + \gamma^2 \|\nabla (r^2+\rho_0^2)^{1/2}\|^2 \\ &= \gamma \operatorname{div} \frac{r\,\nabla r}{(r^2+\rho_0^2)^{1/2}} + \gamma^2 \left\| \frac{r\,\nabla r}{(r^2+\rho_0^2)^{1/2}} \right\|^2 \\ &= \gamma \frac{n}{(r^2+\rho_0^2)^{1/2}} - \gamma \frac{r^2}{(r^2+\rho_0^2)^{3/2}} + \gamma^2 \frac{r^2}{r^2+\rho_0^2}. \end{aligned}$$

The result follows by rearranging some terms. $\square$

Remark 4.1.5. *We will often only be interested in a crude estimate of the right hand side of* (4.1.4)*. Namely, in the case $\gamma > 0$, we are going to be interested in ensuring that*

$$\left| \frac{\gamma}{(\|\mathbf{x}\|^2+\rho_0^2)^{3/2}} \left((n-\gamma)\|\mathbf{x}\|^2 + \gamma(\|\mathbf{x}\|^2+\rho_0^2)^{1/2} + n\rho_0^2 \right) \right| \le \delta\gamma^2 \tag{4.1.5}$$

holds for a given $\delta > 0$. The first and third terms combined are

$$\left| \frac{\gamma}{(\|\mathbf{x}\|^2+\rho_0^2)^{3/2}} \left((n-\gamma)\|\mathbf{x}\|^2 + n\rho_0^2 \right) \right| \le \frac{n\gamma}{(\|\mathbf{x}\|^2+\rho_0^2)^{1/2}} \le \frac{\delta\gamma^2}{2},$$

as long as $\rho_0 \ge \frac{2n}{\gamma\delta}$. The second term is

$$\left| \frac{\gamma}{(\|\mathbf{x}\|^2+\rho_0^2)^{3/2}} \cdot \gamma(\|\mathbf{x}\|^2+\rho_0^2)^{1/2} \right| = \frac{\gamma^2}{\|\mathbf{x}\|^2+\rho_0^2} \le \frac{\delta\gamma^2}{2},$$

as long as $\rho_0 \ge \frac{2}{\delta}$. In particular, (4.1.5) *will hold true as long as $\rho \ge \frac{2}{\delta}\max\{1, \frac{n}{\gamma}\}$.*

Solutions of the Allen-Cahn equation decay exponentially quickly to ± 1 in the convex region, as the following well-known estimate shows.

Proposition 4.1.6 (Convexity estimate, cf. Gui [29, Proposition 2.2], Kowalczyk-Liu-Pacard [35, Lemma 4.2])**.** *If $u \in C^\infty_{\mathrm{loc}}(\mathbf{R}^n)$ is a solution of* (4.1.1)*, $|u| < 1$, and W satisfies* (H1)*,* (H2)*,* (H3)*, then for κ as in* (H3) *and $\theta \in (0,1)$,*

$$|u(\mathbf{x})^2 - 1| + \|\nabla u(\mathbf{x})\| + \|\nabla^2 u(\mathbf{x})\| \le c_0 \exp\left(-\theta\sqrt{\kappa}\operatorname{dist}(\mathbf{x}, \{u=0\})\right) \text{ for all } \mathbf{x} \in \mathbf{R}^2,$$

where $c_0 = c_0(n, W, \theta)$.

Proof. Without loss of generality, we may prove this for $\mathbf{x} \in \mathbf{R}^2$ such that $u(\mathbf{x}) > 0$. By (H1) and Taylor's theorem on $W'(t)$, for $t \approx 1$,

$$-W'(t) = W'(1) - W'(t) = W''(\xi)(1-t)$$

for some $\xi \in (t,1)$. By (H3), there exists $\beta = \beta(W) \in (0,1)$, such that

$$-W'(t) \geq \kappa(1-t) \text{ when } 1-\beta < t < 1. \tag{4.1.6}$$

Claim. Let $\mathbf{x}_0 \in \{u>0\}$, and $R \triangleq \operatorname{dist}(\mathbf{x}_0, \{u=0\})$. There exists $R_0 = R_0(n,W) > 0$ such that, if $R > R_0$, then $u \geq 1-\beta$ in $B_{R/2}(\mathbf{x}_0)$.

Let us first show how the $|u^2-1|$ estimate follows assuming the claim has been proven. Let $\mathbf{x}_0$ be as in the claim above. For $\gamma = \theta\sqrt{\kappa}$ and $\rho_0 > 0$ sufficiently large depending on n, θ, κ (see Remark 4.1.5), the function $w_{\mathbf{x}_0} \triangleq w(\cdot - \mathbf{x}_0) : \mathbf{R}^n \to \mathbf{R}$, with w as in Lemma 4.1.4, satisfies the differential inequality

$$(-\Delta + \kappa) w_{\mathbf{x}_0} \geq 0 \text{ in } \mathbf{R}^n. \tag{4.1.7}$$

Choose $Q \triangleq \exp\big[-\gamma(R^2/4+\rho_0^2)^{1/2}\big]$ so that $Qw_{\mathbf{x}_0} \equiv 1$ on $\partial B_{R/2}(\mathbf{x}_0)$. From (4.1.6) and (4.1.7),

$$(-\Delta+\kappa)Qw_{\mathbf{x}_0} \geq (-\Delta+\kappa)(1-u) \text{ in } B_{R/2}(\mathbf{x}_0) \text{ and } Qw_{\mathbf{x}_0} > 1-u \text{ on } \partial B_{R/2}(\mathbf{x}_0).$$

By the maximum principle (Theorem B.1.1), $Qw_{\mathbf{x}_0} > 1-u$ in $B_{R/2}(\mathbf{x}_0)$. Thus,

$$1-u(\mathbf{x}_0) < Qw_{\mathbf{x}_0}(\mathbf{x}_0) = \exp\left[\gamma(\rho_0 - (R^2/4+\rho_0^2)^{-1/2})\right],$$

and the estimate on $1-u$ follows for a suitable c_0. Similar arguments hold in case $u \leq -1+\beta$ in $B_{R/2}(\mathbf{x}_0)$, so, recalling $\gamma = \theta\sqrt{\kappa}$, we see that $|u(\mathbf{x}_0)^2-1| \leq c\exp\left(-\theta\sqrt{\kappa}\operatorname{dist}(\mathbf{x}_0,\{u=0\})\right)$.

Now we return to proving the claim.

Proof of claim. Without loss of generality, $u(\mathbf{x}_0) > 0$. By (H2), (H3), there exists $\mu > 0$ such that

$$-W'(t) \geq \mu t \text{ for } 0 \leq t \leq 1-\beta; \tag{4.1.8}$$

here, β is as in (4.1.6).

Suppose, for the sake of contradiction, that there were $\mathbf{y}_0 \in B_{R/2}(\mathbf{x}_0)$ where $u(\mathbf{y}_0) < 1-\beta$. Denote $\varphi > 0$ a Dirichlet eigenfunction of $-\Delta$ in $B_{R/2}(\mathbf{y}_0)$ such that $\varphi(\mathbf{y}_0) = 1$. Since $B_{R/2}(\mathbf{y}_0) \subset B_R(\mathbf{x}_0) \subset \{u>0\}$, there exists a maximal $Q>0$ such that $Q\varphi \leq u$ in $B_{R/2}(\mathbf{y}_0)$. Denoting the point of contact to be $\mathbf{z}_0$, we have, by virtue of the fact that φ is maximized at $\mathbf{y}_0$,

$$u(\mathbf{z}_0) = Q\varphi(\mathbf{z}_0) \leq Q\varphi(\mathbf{y}_0) \leq u(\mathbf{y}_0) < 1-\beta,$$

so, by (4.1.8),

$$-\Delta u(\mathbf{z}_0) = -W'(u(\mathbf{z}_0)) \geq \beta u(\mathbf{z}_0). \tag{4.1.9}$$

On the other hand, by virtue of being a contact point for $u \geq Q\varphi$, and by virtue of φ being an eigenfunction with eigenvalue $\lambda_1(-\Delta; B_{R/2})$,

$$-\Delta u(\mathbf{z}_0) \leq -Q\Delta\varphi(\mathbf{z}_0) \leq Q\lambda_1(-\Delta; B_{R/2})\varphi(\mathbf{z}_0) = 4R^{-2}\lambda_1(-\Delta; B_1)u(\mathbf{z}_0).$$

This contradicts (4.1.9) if R_0 is large enough that $4R_0^{-2}\lambda_1(-\Delta; B_1) < \beta$. □

The proof of Proposition 4.1.6 will be complete as soon as we establish the higher order estimates. Let $\mathbf{x}_0 \in \mathbf{R}^n$ be as in the claim above, $R = \operatorname{dist}(\mathbf{x}_0, \{u = 0\})$, and $u \geq 1 - \beta$ in $B_{R/2}(\mathbf{x}_0)$. Without loss of generality, $R \geq 2$. By applying [27, Theorem 3.9] to $\Delta(u-1) = W'(u)$,

$$\|\nabla u(\mathbf{x}_0)\| \leq C_0 \left(\sup_{B_1(\mathbf{x}_0)} |u - 1| + \sup_{B_1(\mathbf{x}_0)} |W'(u)| \right),$$

with $C_0 = C_0(n, W)$. From Taylor's theorem on $W'(t)$ and (H1),

$$W'(t) = W'(t) - W'(1) = W''(\xi)(t-1),$$

for some $\xi \in (1-t, 1)$, so

$$\|\nabla u(\mathbf{x}_0)\| \leq C_0(1 + \|W''\|_{L^\infty(\mathbf{R})}) \sup_{B_1(\mathbf{x}_0)} |u - 1|.$$

The gradient estimate follows from the supremum estimate, after possibly relaxing c_0. The Hessian estimate follows similarly; differentiate $\Delta u = W'(u)$ apply [27, Theorem 3.9] again, and relax c_0. □

Remark 4.1.7. *We used the standard fact that the first Dirichlet eigenfunction of the Laplacian in Euclidean balls is maximized at the central point. This is a simple consequence of the maximum principle (Theorem B.1.1), which shows that the first eigenfunction is rotationally symmetric, and the min-max characterization of first eigenfunctions (Proposition B.2.2).*

The assumption $|u| < 1$ in the convexity estimate is not essential. In fact, any nonconstant bounded solution turns out to satisfy this bound. This is a corollary of the following important gradient estimate due to Modica [45], whose proof we omit.

Proposition 4.1.8 (Modica gradient estimate, Modica [45, Theorems I, II])**.** *Let $u \in C^\infty_{\mathrm{loc}}(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$ be an entire solution of* (4.1.1)*, with W satisfying* (H1)*. Then*

$$\frac{1}{2}\|\nabla u\|^2 \leq W(u) \ \text{in } \mathbf{R}^n.$$

If $W(u) = 0$ at some point, then u is constant.

Corollary 4.1.9. *Suppose $u \in C^\infty_{\mathrm{loc}}(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$ is an entire solution of* (4.1.1)*, and W satisfies* (H1) *and* (H3)*. Then $|u| < 1$, $u \equiv 1$, or $u \equiv -1$.*

Proof. Given Proposition 4.1.8, it suffices to show that $|u| > 1$ cannot hold true everywhere. By continuity, and without loss of generality, it suffices to show that $u > 1$ cannot hold true everywhere. We proceed by contradiction. If $u > 1$ in $\mathbf{R}^n$, then we may repeat the part of the argument in the proof of Proposition 4.1.6 following the claim and prove that $u \equiv 1$, which is impossible. ☐

4.1.3 One-dimensional solutions, De Giorgi conjectures

The one-dimensional heteroclinic solutions $H : \mathbf{R} \to \mathbf{R}$ of (4.1.1) on $\mathbf{R}$,

$$H' = \sqrt{2W \circ H},\ H(0) = 0, \tag{4.1.10}$$

are foundational in the theory of phase transitions.

Note that there is a natural way to lift these heteroclinic solutions to higher dimensional Riemannian manifolds. Namely, for any Riemannian manifold (M^{n-1}, g), the function

$$M^n \times \mathbf{R} \ni (p,t) \mapsto H(t),$$

with $H : \mathbf{R} \to \mathbf{R}$ as in (4.1.10), is known as a "one-dimensional" heteroclinic solution of the Allen-Cahn equation on the product manifold $(M^{n-1} \times \mathbf{R}, g + dt^2)$.

One-dimensional solutions on Euclidean space stand out the most among all solutions to (4.1.1), not just because of their simplicity, but also because of their rigid variational behavior. For instance:

Lemma 4.1.10 (M.)**.** *Suppose W satisfies* (H1)*,* (H2)*, and* (H3)*. Then the only nonconstant solution of* (4.1.2) *with finite energy, up to translations and reflections, is* (4.1.10)*.*

Proof. If u has finite total energy, then $\frac{1}{2}(u')^2 + W(u) \in L^1(\mathbf{R})$, so by (H1) it would follow that $(u')^2$, $W(u) \in L^1(\mathbf{R})$. On the other hand, multiplying (4.1.1) by u', we see that

$$\frac{d}{dt}\left(\frac{1}{2}(u')^2 - W(u)\right) = 0.$$

Since no nonzero constant function is in $L^1(\mathbf{R})$, it must be that $\frac{1}{2}(u')^2 = W(u)$. Together with (H1), this readily implies that u' cannot change sign except across points with $u = \pm 1$.

We claim that $|u| \leq 1$. Indeed, without loss of generality, we may suppose—for the sake of contradiction—that there exists $x_0 \in \mathbf{R}$ with $u(x_0) > 1$, $u'(x_0) = \sqrt{2W(u(x_0))} > 0$. It is simple to see, then, that $u' > 0$ for all $x > x_0$, and thus $u > u(x_0) > 1$ for all $x > x_0$. This contradicts the finite total energy assumption by (H1) and (H3).

Next, we claim that $|u| < 1$. If this were false, then there would exist $x_0 \in \mathbf{R}$ where, without loss of generality, $u(x_0) = 1$. This would also have $u'(x_0) = 0$. By existence and uniqueness theory for ordinary differential equations, and (H1), u is constant—a contradiction.

Summarizing, we have $|u| < 1$, so u' cannot change sign and, therefore, u is strictly monotone. Without loss of generality, we may assume that $u' > 0$, so $u' = \sqrt{2W(u)}$. It is a simple consequence of (H2) that u has a root in $\mathbf{R}$. The result follows from existence and uniqueness theory for ordinary differential equations. $\square$

Certainly, Lemma 4.1.10 is no longer true—as stated—in higher dimensions. De Giorgi conjectured [15] that one-dimensional solutions are the only "monotone" solutions in low enough dimensions n of Euclidean space, similar to the role of hyperplanes in minimal hypersurface theory.

Conjecture 4.1.11 (De Giorgi conjecture, monotone version)**.** *Let $n \leq 8$. If $u : \mathbf{R}^n \to \mathbf{R}$ is a solution of* (4.1.1) *such that $|u| < 1$ and such that $\langle \nabla u, \mathbf{e}_n \rangle > 0$, then u is one-dimensional.*

This conjecture is inspired by the Bernstein theorem in minimal surface theory, which states that hyperplanes are the only minimal hypersurfaces that are graphical over a hyperplane $P^{n-1} \subset \mathbf{R}^n$, $n \leq 8$. (See [12].)

A closely related conjecture is:

Conjecture 4.1.12 (De Giorgi conjecture, minimizing version)**.** *Let $n \leq 7$. If $u : \mathbf{R}^n \to \mathbf{R}$ is a solution of* (4.1.1) *such that $|u| < 1$ and such that u is energy minimizing among all compactly supported perturbations, then u is one-dimensional.*

The first positive result was due to Ghoussoub-Gui [25], who confirmed Conjecture 4.1.11 for $n = 2$. Their proof can be adapted to confirm Conjecture 4.1.12 in the same dimension. (We refer to the work of Farina-Mari-Valdinoci [20] for this adaptation.)

Theorem 4.1.13 (cf. Ghoussoub-Gui [25, Theorem 1.1], Farina-Mari-Valdinoci [20, Theorem 1])**.** *Conjectures 4.1.11 and 4.1.12 are true when $n = 2$.*

Next, Ambrosio-Cabré [1] confirmed Conjecture 4.1.11 for $n = 3$. Their proof, too, can be adapted to confirm Conjecture 4.1.12 in the same dimension. (Again, we refer to the work of Farina-Mari-Valdinoci [20] for this adaptation.)

Theorem 4.1.14 (cf. Ambrosio-Cabré [1, Theorem 1.2], Farina-Mari-Valdinoci [20, Theorem 1]). *Conjectures 4.1.11, 4.1.11 are true when* $n = 3$.

Finally, Savin [58] confirmed the remaining cases of Conjecture 4.1.12 as well as a weaker form of the remaining cases of Conjecture 4.1.11.

Theorem 4.1.15 (cf. Savin [58, Theorems 2.3, 2.4]). *Conjecture 4.1.12 is true when* $4 \leq n \leq 7$. *Conjecture 4.1.11 is true for* $4 \leq n \leq 8$, *under the additional asymptotic assumption*

$$\lim_{t\to\infty} u(\mathbf{x}', t) = \lim_{t\to\infty} (-u(\mathbf{x}', -t)) = 1 \textit{ for all } \mathbf{x}' \in \mathbf{R}^{n-1}. \tag{4.1.11}$$

Both conjectures are known to fail for the standard double-well potential $W = \frac{1}{4}(1-u^2)^2$ in dimensions higher than those mentioned. Specifically:

(1) Del Pino-Kowalczyk-Wei [17] constructed monotone solutions in $\mathbf{R}^n$, $n \geq 9$, which are not one dimensional.

(2) Liu-Wang-Wei [36] constructed energy-minimizing solutions in $\mathbf{R}^n$, $n \geq 8$, which are not one-dimensional.

(3) Liu-Wang-Wei's construction already yields stable counterexamples to the "one-dimensional" conjecture, but we also mention that Pacard-Wei [49] constructed stable solutions in $\mathbf{R}^n$, $n \geq 8$, which are not one-dimensional.

4.2 Results on the space $\mathcal{M}_{2k}$ of $2k$-ended solutions in $\mathbf{R}^2$

4.2.1 Introduction to $\mathcal{M}_{2k}$

We now move one level of complexity up from the standard one-dimensional solutions of (4.1.1), and study nontrivial solutions on $\mathbf{R}^2$. Del Pino-Kowalczyk-Pacard defined in [16] a space of solutions of (4.1.1) on $(\mathbf{R}^2, \delta)$, $\mathcal{M}_{2k}$, that looks from infinity like a collection of $2k$ copies of the one-dimensional heteroclinic solution. We recall the construction of this space here (after [16]) for the sake of completeness.

Remark 4.2.1. *We assume hypotheses* (H1), (H2), (H3) *throughout this section.*

Fix $k \in \{1, 2, \ldots\}$. We denote by Λ^{2k} (denoted $\Lambda^{2k}_{\text{ord}}$ in [16]) the space of ordered $2k$-tuples $\lambda = (\lambda_1, \ldots, \lambda_{2k})$ of oriented affine lines on $\mathbf{R}^2$, parametrized as

$$\lambda_j = (r_j, \mathbf{f}_j) \in \mathbf{R} \times \mathbf{S}^1, \ j = 1, \ldots, 2k,$$

where $\mathbf{f}_j = (\cos\theta_j, \sin\theta_j)$, $\theta_1 < \ldots < \theta_{2k} < 2\pi + \theta_1$. For $\lambda \in \Lambda^{2k}$, we denote

$$\theta_\lambda \triangleq \frac{1}{2}\min\{\theta_2 - \theta_1, \ldots, \theta_{2k} - \theta_{2k-1}, 2\pi + \theta_1 - \theta_{2k}\}. \tag{4.2.1}$$

Fix $\lambda \in \Lambda^{2k}$. For large $R > 0$ and all $j = 1, \ldots, 2k$, there exists $s_j \in \mathbf{R}$ such that $r_j\mathbf{J}\mathbf{f}_j + s_j\mathbf{f}_j \in \partial B_R(\mathbf{0})$, the half-lines $\lambda_j^+ \triangleq r_j\mathbf{J}\mathbf{f}_j + s_j\mathbf{f}_j + \mathbf{R}_+\mathbf{f}_j$ are disjoint and contained in $\mathbf{R}^2 \setminus B_R(\mathbf{0})$, and the minimum distance of any two distinct λ_i^+, λ_j^+ is ≥ 4. (Here, $\mathbf{J} \in \mathrm{End}(\mathbf{R}^2)$ is the counterclockwise rotation map by $\frac{\pi}{2}$.) The affine half-lines $\lambda_1^+, \ldots, \lambda_{2k}^+$ and the circle $\partial B_R(\mathbf{0})$ induce a decomposition of $\mathbf{R}^2$ into $2k+1$ open sets,

$$\begin{aligned}&\Omega_0 \triangleq B_{R+1}(\mathbf{0}), \text{ and}\\ &\Omega_j \triangleq \bigcap_{i\neq j}\{\mathbf{x} \in \mathbf{R}^2 \setminus B_{R-1}(\mathbf{0}) : \mathrm{dist}(\mathbf{x}, \lambda_j^+) < \mathrm{dist}(\mathbf{x}, \lambda_i^+) + 2\},\ j = 1, \ldots, 2k\end{aligned} \tag{4.2.2}$$

Note that these open sets are not disjoint. Then, we define $\chi_{\Omega_0}, \ldots, \chi_{\Omega_{2k}}$ to be a smooth partition of unity of $\mathbf{R}^2$ subordinate to $\Omega_0, \ldots, \Omega_{2k}$, and such that

$$\begin{aligned}&\chi_{\Omega_0} \equiv 1 \text{ on } \Omega_0' \triangleq B_{R-1}(\mathbf{0}),\ \text{and}\\ &\chi_{\Omega_j} \equiv 1 \text{ on } \Omega_j' \triangleq \cap_{i\neq j}\{\mathbf{x} \in \mathbf{R}^2 \setminus B_{R+1}(\mathbf{0}) : \mathrm{dist}(\mathbf{x}, \lambda_j^+) < \mathrm{dist}(\mathbf{x}, \lambda_i^+) - 2\},\ j = 1, \ldots, 2k.\end{aligned} \tag{4.2.3}$$

Note that these new open sets are disjoint. Without loss of generality, $|\chi_{\Omega_j}| + \|\nabla\chi_{\Omega_j}\| + \|\nabla^2\chi_{\Omega_j}\| \leq c_1$ for all $j = 0, \ldots, 2k$, with $c_1 = c_1(\theta_\lambda)$. Finally, we define

$$u_\lambda \triangleq \sum_{j=1}^{2k}(-1)^{j+1}\chi_{\Omega_j}H(\mathrm{dist}^s(\cdot, \lambda_j)), \tag{4.2.4}$$

where $\mathrm{dist}^s(\cdot, \lambda_j)$ denotes the signed distance to λ_j, taking $\mathbf{J}\mathbf{f}_j$ to be the positive direction. Here, H is the heteroclinic solution (4.1.10).

Definition 4.2.2 (Del Pino-Kowalczyk-Pacard [16, Definition 2.2])**.** *For $k \geq 1$, we denote*

$$\mathcal{S}_{2k} \triangleq \bigcup_{\lambda\in\Lambda^{2k}}\{u \in C^\infty(\mathbf{R}^2) : u - u_\lambda \in W^{2,2}(\mathbf{R}^2)\}. \tag{4.2.5}$$

We endow $\mathcal{S}_{2k}$ with the weak topology of the operator $\mathcal{J} : \mathcal{S}_{2k} \to W^{2,2}(\mathbf{R}^2) \times \Lambda^{2k}$,

$$\mathcal{J}(u) \triangleq (u - u_\lambda, \lambda). \tag{4.2.6}$$

Finally, we define the space of "2k-ended solutions" to be

$$\mathcal{M}_{2k} \triangleq \{u \in \mathcal{S}_{2k} \ \textit{satisfying}\ (4.1.1)\}. \tag{4.2.7}$$

Remark 4.2.3. *The space $\mathcal{M}_{2k}$ is independent of our particular choices of $\Omega_0, \ldots, \Omega_{2k}$, $\chi_{\Omega_0}, \ldots, \chi_{\Omega_{2k}}$, and the operator $\mathcal{J}$ is well-defined, because*

$$\begin{aligned} &\lambda \neq \lambda' \in \Lambda^{2k} \ \textit{up to rotation} \\ &\implies \{u \in (u_\lambda + W^{2,2}(\mathbf{R}^2)) \ \textit{satisfying}\ (4.1.1)\} \cap \{u \in (u_{\lambda'} + W^{2,2}(\mathbf{R}^2)) \ \textit{satisfying}\ (4.1.1)\} = \emptyset. \end{aligned}$$

This follows from Proposition 4.1.6.

Remark 4.2.4. *It was shown in [16, Theorem 2.2] that $\mathcal{M}_{2k}$ is a $2k$-dimensional Banach manifold in neighborhoods of points $u \in \mathcal{M}_{2k}$ satisfying a certain "nondegeneracy" condition, namely, the nonexistence of exponentially decaying Jacobi fields (see Definition 4.2.13).*

Example 4.2.5 (One-dimensional solutions)**.** *All one-dimensional solutions on $\mathbf{R}^2$,*

$$\mathbf{R}^2 \ni \mathbf{x} \mapsto H(\langle \mathbf{x}, \mathbf{e}\rangle - \beta), \ \textit{with}\ (\mathbf{e}, \beta) \in \mathbf{S}^1 \times \mathbf{R},$$

are elements of $\mathcal{M}_2$. This family accounts for all dimensions of $\mathcal{M}_2$ predicted by Remark 4.2.4. See also Proposition 4.2.16.

Example 4.2.6 (Saddle solution)**.** *A solution $u \in C^\infty_{\mathrm{loc}}(\mathbf{R}^2) \cap L^\infty(\mathbf{R}^2)$ of the Allen-Cahn equation* (4.1.2) *with $u = 0$ on $\{xy = 0\}$ and $u(x, y) = \mathrm{sign}(xy)$ when $xy \neq 0$ is called a "saddle solution." Such a solution is an element of $\mathcal{M}_4$, and is known to be unique [11] and nondegenerate [34], [35, Theorem 2.4].*

The following result, due to Del Pino-Kowalczyk-Pacard, significantly improves the a priori $W^{2,2}$ decay of $u - u_\lambda$ to an exponential decay:

Theorem 4.2.7 (Refined asymptotics, Del Pino-Kowalczyk-Pacard [16, Theorem 2.1])**.** *Let $u_0 \in \mathcal{M}_{2k}$. There exists a neighborhood $U \subset \mathcal{M}_{2k}$ and a $\delta = \delta(u_0) > 0$ such that*

$$\mathcal{J}(U) \subseteq e^{-\delta\|\mathbf{x}\|} W^{2,2}(\mathbf{R}^2) \times \Lambda^{2k}, \tag{4.2.8}$$

and, moreover, such that the restricted map

$$\mathcal{J}|_U : U \to e^{-\delta\|\mathbf{x}\|} W^{2,2}(\mathbf{R}^2) \times \Lambda^{2k} \tag{4.2.9}$$

is continuous with respect to the corresponding topologies; here, $\mathcal{J}$ is the map defined in (4.2.6).

Remark 4.2.8. *Del Pino-Kowalczyk-Pacard assume in [16] that their double-well potential is even, i.e., that $W(x) = W(-x)$ for all $x \in \mathbf{R}$. This assumption simplifies their analysis as well as their statements. Nonetheless, this assumption is not necessary if one just requires Theorem 4.2.7 above.*

4.2.2 Effects of topology at infinity on the Morse index

One can make precise the notions of **stability** and **Morse index** in the Allen-Cahn setting by turning to the **second variation** operator, $\delta^2 E[u] : W_0^{1,2}(M) \otimes W_0^{1,2}(M) \to \mathbf{R}$,

$$\delta^2 E[u]\{\zeta, \psi\} \triangleq \int_M \left(\langle \nabla\zeta, \nabla\psi\rangle + W''(u)\zeta\psi\right) \, dv_g. \tag{4.2.10}$$

Definition 4.2.9 (Stability)**.** *A critical point u of E is called stable inside an open subset $U \subseteq M$ if $\delta^2 E[u]\{\zeta, \zeta\} \geq 0$ for all $\zeta \in W_0^{1,2}(U)$.*

Definition 4.2.10 (Morse index)**.** *A critical point u of E is said to have Morse index k inside an open subset $U \subseteq M$, denoted* $\operatorname{ind}(u; U) = k$, *provided*

$$\max\{\dim V : V \subset W_0^{1,2}(U) \text{ is a vector space such that } \delta^2 E[u]\{\zeta, \zeta\} < 0 \text{ for all } \zeta \in V \setminus \{0\}\}.$$

If the choice of the open subset U is clear or $U = M$ then we will simply write $\operatorname{ind}(u)$*. Stable critical points, by definition, have* $\operatorname{ind}(u) = 0$*.*

The Morse index of a critical point measures the number of linearly independent unstable directions for energy. From a physical perspective, unstable critical points are a lot less likely to be observed than stable ones.

Associated to $\delta^2 E[u]$ is the linear elliptic operator $-\Delta_g + W''(u)$, called the **Jacobi operator**. These notions of stability and Morse index are developed in Section 4.3.1.

In this section we prove the following theorem, which relates the Morse index of $2k$-ended solutions to (4.1.1) to their structure at infinity.

Theorem 4.2.11 (M.)**.** *Let $u \in \mathcal{M}_{2k}$, $k \geq 2$. Then* $\operatorname{ind}(u) \geq k - 1$*.*

Remark 4.2.12. *Notice that we assume that $k \geq 2$, seeing as to how Proposition 4.2.16 establishes a stronger result when $k = 1$.*

To prove this Theorem, we will need to obtain a precise pointwise understanding of kernel elements of the Jacobi operator, seeing as to how they will play a significant role in the relevant variational theory:

Definition 4.2.13 (Jacobi fields)**.** *If u is a critical point of E in U, then the space of its Jacobi fields consists of all functions v that satisfy $-\Delta v + W''(u)v = 0$ in U in the classical sense.*

Denote R, $\lambda \in \Lambda^{2k}$, and u_λ the objects associated with u by its construction as an element of $\mathcal{M}_{2k}$ in Section 4.2.1. Also, denote $\lambda = (\lambda_1, \ldots, \lambda_{2k})$, with $\lambda_i = (\tau_i, \mathbf{f}_i) \in \mathbf{R} \times \mathbf{S}^1$. Recall from [16, (2.16)] that:

$$\sum_{i=1}^{2k} \mathbf{f}_i = \mathbf{0}, \tag{4.2.11}$$

and that, after possibly enlarging $R > 0$, $\{u = 0\} \setminus B_R(\mathbf{0})$ decomposes into $2k$ disjoint curves Γ_i, $i = 1, \ldots, 2k$, and, for some $\delta < \theta_\lambda(u)$, $\Gamma_i \subset S(\mathbf{f}_i, \delta/2, R)$ with $S(\mathbf{f}_i, \delta, R)$ all pairwise disjoint. Here,

$$S(\mathbf{e}, \theta, R) \triangleq \{r\mathbf{f} : r \geq R, \mathrm{dist}_{\mathbf{S}^1}(\mathbf{f}, \mathbf{e}) < \theta\}. \tag{4.2.12}$$

Finally, using Theorem 4.2.7 we see that, perhaps after shrinking $\delta > 0$ and enlarging $R > 0$, and perhaps after an ambient rigid motion,

$$\frac{\nabla u}{\|\nabla u\|} \approx (-1)^{i+1} \mathbf{J}\mathbf{f}_i \text{ in } S(\mathbf{f}_i, \delta/2, R). \tag{4.2.13}$$

Lay out $\mathbf{f}_1, \ldots, \mathbf{f}_{2k} \in \mathbf{S}^1$, and color them red (negative) or blue (positive) depending on the sign of $\langle (-1)^{i+1} \mathbf{J}\mathbf{f}_i, \mathbf{e} \rangle$. Here, $\mathbf{e} \in \mathbf{S}^1$ is a fixed direction, chosen generically, so that $\langle \mathbf{J}\mathbf{f}_i, \mathbf{e} \rangle \neq 0$ for all $i = 1, \ldots, 2k$. We will temporarily need the following generalization of $\mathbf{J}$:

$$\mathbf{J}_\theta \in \mathrm{End}(\mathbf{R}^2) \text{ acting by } \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

There exist unique $\varphi_1, \ldots, \varphi_{2k} \in (0, 2\pi)$ such that $\mathbf{f}_{i+1} = \mathbf{J}_{\varphi_i}(\mathbf{f}_i)$ for all $i = 1, \ldots, 2k$. It's easy to see that

$$\varphi_i \in (0, \pi) \text{ for all } i = 1, \ldots, 2k \tag{4.2.14}$$

by combining (4.2.11) with $k \geq 2$ (recall Remark 4.2.12).

Claim. If $\mathbf{f}_{2\ell-1}$, $\mathbf{f}_{2\ell}$ have the same color, blue, then

$$\mathbf{J}_{-\varphi_{2\ell-1}}\mathbf{e}, \mathbf{f}_{2\ell-1}, \mathbf{e} \text{ lie counterclockwise on } \mathbf{S}^1 \text{ in the order listed;}$$

else, if their common color is red, then

$$\mathbf{J}_{-\varphi_{2\ell-1}}(-\mathbf{e}), \mathbf{f}_{2\ell-1}, -\mathbf{e} \text{ lie counterclockwise on } \mathbf{S}^1 \text{ in the order listed.}$$

Proof. Without loss of generality, let us assume $\ell = 1$. Recall that the respective colors are determined by the signs of $\langle \mathbf{J}_{\frac{\pi}{2}}\mathbf{f}_1, \mathbf{e}\rangle$ and $\langle -\mathbf{J}_{\frac{\pi}{2}}\mathbf{f}_2, \mathbf{e}\rangle = \langle \mathbf{J}_{-\frac{\pi}{2}+\varphi_1}\mathbf{f}_1, \mathbf{e}\rangle$.

Denote $\mathcal{P} \triangleq \{\mathbf{f} \in \mathbf{S}^1 : \langle \mathbf{f}, \mathbf{e}\rangle > 0\}$. If both colors are blue, then

$$\mathbf{J}_{\frac{\pi}{2}}\mathbf{f}_1 \in \mathcal{P} \iff \mathbf{f}_1 \in \mathbf{J}_{-\frac{\pi}{2}}(\mathcal{P})$$

and

$$\mathbf{J}_{-\frac{\pi}{2}+\varphi_1}\mathbf{f}_1 \in \mathcal{P} \iff \mathbf{f}_1 \in \mathbf{J}_{\frac{\pi}{2}-\varphi_1}(\mathcal{P}),$$

i.e., $\mathbf{f}_1 \in \mathbf{J}_{-\frac{\pi}{2}}(\mathcal{P}) \cap \mathbf{J}_{\frac{\pi}{2}-\varphi_1}(\mathcal{P})$. Using (4.2.14), we see that the three vertices $\mathbf{J}_{-\varphi_1}\mathbf{e}$, $\mathbf{f}_1$, and $\mathbf{e}$, must lie counterclockwise in this order on $\mathbf{S}^1$.

If both colors are red, then, by a similar argument, $\mathbf{f}_1 \in \mathbf{J}_{-\frac{\pi}{2}}(-\mathcal{P}) \cap \mathbf{J}_{\frac{\pi}{2}-\varphi_1}(-\mathcal{P})$, and we see that the three vertices $\mathbf{J}_{-\varphi_1}(-\mathbf{e})$, $\mathbf{f}_1$, and $-\mathbf{e}$, must lie counterclockwise in this order on $\mathbf{S}^1$. $\square$

In a completely analogous manner, one also checks that:

Claim. If $\mathbf{f}_{2\ell}$, $\mathbf{f}_{2\ell+1}$ have the same color, blue, then

$$\mathbf{J}_{-\varphi_{2\ell}}(-\mathbf{e}), \mathbf{f}_{2\ell}, -\mathbf{e} \text{ lie counterclockwise on } \mathbf{S}^1 \text{ in the order listed;}$$

else, if their common color is red, then

$$\mathbf{J}_{-\varphi_{2\ell}}\mathbf{e}, \mathbf{f}_{2\ell}, \mathbf{e} \text{ lie counterclockwise on } \mathbf{S}^1 \text{ in the order listed.}$$

We now make the following key observation:

Claim. There exist at least $2k - 2$ groups of consecutive same-colored vertices.

Proof of claim. Within the space of valid colorings,

$$\{\text{existence of blue } \mathbf{f}_{2\ell-1}, \mathbf{f}_{2\ell}\} \cap \{\text{existence of red } \mathbf{f}_{2m}, \mathbf{f}_{2m+1}\} = \emptyset. \tag{4.2.15}$$

This follows by combining the previous two claims. Likewise

$$\{\text{existence of red } \mathbf{f}_{2\ell-1}, \mathbf{f}_{2\ell}\} \cap \{\text{existence of blue } \mathbf{f}_{2m}, \mathbf{f}_{2m+1}\} = \emptyset. \tag{4.2.16}$$

There are now the following cases to consider:

(1) There exist three consecutive same-colored vertices. Then, by combining the previous two claims and engaging in elementary angle-chasing, it follows that there do not exist any more

consecutive same-colored vertices. In this case, it follows that there are precisely $2k-2$ groups of consecutive same-colored vertices.

(2) There are no three consecutive same-colored vertices. Then, together with (4.2.15), (4.2.16), it follows that there are at least $2k-2$ groups of consecutive same-colored vertices.

This concludes the proof of the claim. $\square$

Given this claim, differentiate (4.1.1) in the direction of $\mathbf{e} \in \mathbf{S}^1$. We see that $v \triangleq \langle \nabla u, \mathbf{e} \rangle$ satisfies

$$\Delta v = W''(u) v \text{ in } \mathbf{R}^2. \tag{4.2.17}$$

Define

$$\mathcal{N} \triangleq \{v = 0\} \text{ (the "nodal set"), and}$$
$$\mathcal{S} \triangleq \mathcal{N} \cap \{\nabla v = \mathbf{0}\} \text{ (the "singular set").}$$

By the implicit function theorem, $\mathcal{N} \setminus \mathcal{S}$ consists of smooth, injectively immersed curves in $\mathbf{R}^2$. By [**?**], $\mathcal{S}$ consists of at most countably many points and, for each $p \in \mathcal{S}$, there exists $r = r(p)$ such that, up to a diffeomorphism of $B_r(p)$,

$$\mathcal{N} \cap B_r(p) \approx \text{the zero set of a homogeneous even-degree harmonic polynomial.} \tag{4.2.18}$$

Denote $\Omega_1, \ldots, \Omega_q \subset \mathbf{R}^2 \setminus \mathcal{N}$ the nodal domains (i.e., connected components of $\{v \neq 0\}$), labeling so that $\Omega_1, \ldots, \Omega_p$ are the unbounded ones, and $\Omega_{p+1}, \ldots, \Omega_q$ are the bounded ones. By virtue of our precise understanding of $\mathcal{N}$, $\mathcal{S}$, as discussed above, we know that they are all open, connected, Lipschitz domains.

Remark 4.2.14. *The notation used here implicitly asserts that there are finitely many nodal domains. This follows, a posteriori, by the proof of the following claim and [35, Theorem 2.8].*

Claim. $\operatorname{ind}(u) \geq q - 1$.

Proof of claim. First, it's standard that for every bounded nodal domain $\Omega_{p+1}, \ldots, \Omega_q$ we have

$$\operatorname{nul}(u; \Omega_i) \geq 1 \text{ for all } i = p+1, \ldots, q. \tag{4.2.19}$$

Now we move on to unbounded nodal domains. It is not hard to see that we have at least two such. Suppose that Ω_1 is an unbounded nodal domain, and suppose Ω_2 is its counterclockwise neighboring unbounded nodal domain. By (4.2.18), v attains opposite signs on Ω_1, Ω_2. Thus, v is a bounded, sign-changing Jacobi field in $\Omega_{12} \triangleq \operatorname{int} \overline{\Omega}_1 \cup \overline{\Omega}_2$, which is itself an open, connected, unbounded Lischitz domain. By Lemma 4.3.8, $\operatorname{ind}(u; \Omega_{12}) \geq 1$. Since Ω_{12} is unbounded, we have

$$\operatorname{ind}(u; \widetilde{\Omega}_2) \geq 1 \text{ for some bounded } \widetilde{\Omega}_2 \subsetneq \Omega_{12}$$

which is itself open, connected, and Lipschitz. Denote $\widetilde{\Omega}_1 = \emptyset$.

Proceeding similarly (and labeling accordingly) in the counterclockwise direction, we can construct disjoint, bounded, open, connected, Lipschitz $\widetilde{\Omega}_3, \ldots, \widetilde{\Omega}_p$,

$$\operatorname{ind}(u; \widetilde{\Omega}_i) \geq 1 \text{ for all } i = 2, \ldots, p, \tag{4.2.20}$$

where $\widetilde{\Omega}_i \subset \operatorname{int}(\overline{\Omega}_{i-1} \cup \overline{\Omega}_i)$. More precisely, at each stage i we have to sacrifice a bounded portion of $\overline{\Omega}_1 \cup \cdots \cup \overline{\Omega}_{i-1} \setminus (\widetilde{\Omega}_1 \cup \cdots \cup \widetilde{\Omega}_{i-1})$ to give rise to a negative eigenvalue on a slight enlargement of Ω_i, which is bounded and disjoint from $\widetilde{\Omega}_1 \cup \cdots \cup \widetilde{\Omega}_{i-1}$.

The claim follows by combining (4.2.19), (4.2.20), and Theorem 4.3.7. □

We now estimate $q - 1$ from below. It will be convenient to assume that $\mathcal{S}$ and the set of connected components of $\mathcal{N} \setminus \mathcal{S}$ are both finite sets—refer to Remark 4.2.15 for the minor necessary adjustments to deal with the general case. From Euler's formula for planar graphs, we know that

$$q = 1 + |\{\text{connected components of } \mathcal{N} \setminus \mathcal{S}\}| - |\mathcal{S}|, \tag{4.2.21}$$

where $|\cdot|$ denotes the cardinality of a set. By (4.2.18), every connected component Γ of $\mathcal{N} \setminus \mathcal{S}$ is a smooth curve with

$$|\partial\Gamma| = 0, 1, \text{ or } 2, \tag{4.2.22}$$

depending on whether Γ is infinite in both directions, one direction, or is finite. Counting the set of pairs (v, e) of vertices and edges in $\mathcal{N}$ in two ways, we see that

Claim. $q \geq k$.

Proof. The fact that there exist at least $2k - 2$ groups of consecutive same-colored vertices implies that there exists $R > 0$ sufficiently large so that $\mathcal{S} \subset B_R(\mathbf{0})$ and $\mathcal{N} \setminus B_R(\mathbf{0})$ has at least $2k - 2$ components. By a straightforward counting argument combined with (4.2.22), this implies

$$2k - 2 \leq \sum_{\ell=0}^{2} (2 - \ell) \cdot |\{\text{connected components } \Gamma \subset \mathcal{N} \setminus \mathcal{S} : |\partial\Gamma| = \ell\}|. \tag{4.2.23}$$

On the other hand, by counting the elements of the set

$$\mathcal{A} \triangleq \{(p, \Gamma) : p \in \mathcal{S},\ \Gamma = \text{connected component of } \mathcal{N} \setminus \mathcal{S} \text{ incident to } p\}$$

in one way, we find that

$$|\mathcal{A}| = \sum_{\ell=0}^{2} \ell \cdot |\{\text{connected components } \Gamma \subset \mathcal{N} \setminus \mathcal{S} : |\partial\Gamma| = \ell\}|. \tag{4.2.24}$$

Adding (4.2.23), (4.2.24), and rearranging, we get

$$\begin{aligned} &2 \cdot |\{\text{connected components of } \mathcal{N} \setminus \mathcal{S}\}| \geq |\mathcal{A}| + 2k - 2 \\ &\qquad \iff |\{\text{connected components of } \mathcal{N} \setminus \mathcal{S}\}| \geq \frac{1}{2}|\mathcal{A}| + k - 1. \end{aligned} \tag{4.2.25}$$

Plugging (4.2.25) into (4.2.21) yields the estimate

$$q \geq k + \frac{1}{2}|\mathcal{A}| - |\mathcal{S}|. \tag{4.2.26}$$

On the other hand, because of (4.2.18), each $p \in \mathcal{S}$ contributes at least two elements to $\mathcal{A}$; i.e., $|\mathcal{A}| \geq 2 \cdot |\mathcal{S}|$. The claim follows. $\square$

Remark 4.2.15. *The proof above assumed that*

$$|\mathcal{S}| + |\{\textit{connected components of } \mathcal{N} \setminus \mathcal{S}\}| < \infty,$$

so let us discuss the necessary adjustments for it to go through in the general case. By the finiteness of q (see Remark 4.2.14), we know that there exists a large enough radius R so that $\Omega_i \cap B_R(\mathbf{0})$ is connected and nonempty for every $i = 1, \ldots, q$. By the local finiteness of $\mathcal{S}$, we may further arrange for $\partial B_R(\mathbf{0}) \cap \mathcal{S} = \emptyset$ and for all intersections $\partial\Omega_i \cap \partial B_R(\mathbf{0})$, $i = 1, \ldots, 2k$, to be transverse. The finite planar graph arrangement contained within $B_R(\mathbf{0})$ has the same number of faces as the original infinite planar graph arrangement. We may, therefore, repeat the previous proof, starting at Remark 4.2.14, discarding all elements of $\mathcal{S}$ and components of $\mathcal{N} \setminus \mathcal{S}$ that lie fully outside of $B_R(\mathbf{0})$, and identifying $\partial B_R(\mathbf{0})$ with infinity.

Combining everything above, we obtain the thesis of Theorem 4.2.11.

4.2.3 Finite Morse index solutions of linear energy growth

There is a finer characterization of $\mathcal{M}_2$ in terms of Morse index than the one in Theorem 4.2.11:

Proposition 4.2.16 (M.). *The following are all equivalent:*

(1) $u \in \mathcal{M}_2$*;*

(2) $u(\mathbf{x}) \equiv H(\langle \mathbf{x}, \mathbf{e}\rangle - \beta)$ *where* H *is as in* (4.1.10)*, and* $(\mathbf{e}, \beta) \in \mathbf{S}^1 \times \mathbf{R}$*;*

(3) $u \in C^\infty_{\text{loc}}(\mathbf{R}^2) \cap L^\infty(\mathbf{R}^2)$ *satisfies* (4.1.1) *and* $\langle \nabla u, \mathbf{e}\rangle > 0$ *for some fixed* $\mathbf{e} \in \mathbf{S}^1$*;*

(4) $u \in C^\infty_{\text{loc}}(\mathbf{R}^2) \cap L^\infty(\mathbf{R}^2)$ *is a nonconstant minimizer of the energy in* (4.1.2) *among compact perturbations;*

(5) $u \in C^\infty_{\text{loc}}(\mathbf{R}^2) \cap L^\infty(\mathbf{R}^2)$ *is a nonconstant stable critical point of the energy in* (4.1.2)*; i.e.,* $\text{ind}(u) = 0$.

Remark 4.2.17. *We assume hypotheses* (H1)*,* (H2)*,* (H3) *throughout this section.*

Recall, from Remark 4.2.4, that near nondegenerate points, the dimension of $\mathcal{M}_2$ should be 2. According to this theorem, the moduli space $\mathcal{M}_2$ is diffeomorphic to the two-dimensional cylinder $\mathbf{S}^1 \times \mathbf{R}$ that consists of all distinct one-dimensional solutions in $\mathbf{R}^2$.

Proof of Proposition 4.2.16. Let's first show (2) $\iff$ (3) $\iff$ (4) $\iff$ (5). The first three of these equivalences follow from Theorems 4.1.13, 4.1.14. Next, (4) $\implies$ (5) is trivial by the definition of stability. Finally, (5) $\implies$ (3) follows because $\langle \nabla u, \mathbf{e}\rangle$ is an L^∞ Jacobi field, so, by Theorem 4.3.7, it is either identically zero or has constant sign. Since u is not constant, there will exist at least one $\mathbf{e} \in \mathbf{S}^1$ for which $\langle \nabla u, \mathbf{e}\rangle > 0$. For an alternative proof of (5) $\implies$ (2), see the work of Farina-Mari-Valdinoci [20].

It remains to prove the equivalence of (1) with (2)-(5). Since (2) $\implies$ (1) is clear, all we need to show is (1) $\implies$ (2), i.e., that all elements of $\mathcal{M}_2$ are one-dimensional. Without loss of generality, after precomposing with a rigid motion, we may assume that $u(\mathbf{0}) = 0$. Let ε_A, τ_A, $\alpha_A \in (0, 1)$, R_A, $K_A > 0$ be as in [69, Theorem 9.1]. By construction, all $u \in \mathcal{M}_2$ have

$$\lim_{R\uparrow\infty} h_0^{-1}\omega_1^{-1}R^{-1}(E \llcorner B_R(\mathbf{0}))[u] = 1. \tag{4.2.27}$$

Therefore, there exists $R_0 \geq \varepsilon_A^{-1}R_A$ such that, for all $R \geq R_0$, the rescaled function

$$u_R(\mathbf{x}) \triangleq u\left(\frac{R\mathbf{x}}{R_A}\right),\ \mathbf{x} \in B_R(\mathbf{0}),$$

satisfies $h_0^{-1}\omega_1^{-1}R^{-1}(E \llcorner B_R(\mathbf{0}))[u] \leq 1 + \tau_A$, $\Delta u_R(\mathbf{x}) = \left(\frac{R}{R_A}\right)^2 W'(u)$ in $B_R(\mathbf{0})$, and $\varepsilon \triangleq \frac{R_A}{R} \leq \varepsilon_A$. Invoking [69, Theorem 9.1], and passing to a subsequence $R_j \uparrow \infty$, there exists a fixed line $\ell \subset \mathbf{R}^2$, such that, for all $t \in [-1/2, 1/2]$, $j = 1, 2, \ldots$,

$$\{u_{R_j} = t\} \cap \{(x,y) : |x| \leq 1\} = \{(x, h_j^t(x)) : |x| \leq 1\},$$

where $h_j^t : [-1,1] \to \mathbf{R}$ is a $C^{1,\alpha_A}([-1,1])$ function with $|h_j^t|_{C^{1,\alpha_A}([-1,1])} \leq K_A$. Undoing the scaling, this implies that, for all $t \in [-1/2, 1/2]$, $j = 1, 2, \ldots$,

$$\{u = t\} \cap \left\{(x,y) : |x| \leq \frac{R_j}{R_A}\right\} = \left\{(x, h^t(x)) : |x| \leq \frac{R_j}{R_A}\right\},$$

where h^t is a C^{1,α_A} function, whose derivative satisfies the scale-invariant estimate

$$\left(\frac{R}{R_A}\right)^{\alpha_A} [(h^t)']_{C^{\alpha_A}([-R_j/R_A, R_j/R_A])} \leq K_A.$$

Letting $j \uparrow \infty$, we conclude that h^t is constant for each $t \in [-1/2, 1/2]$. Thus, u is one-dimensional by unique continuation (see [24, Theorem 4.2]). $\square$

We conclude this section by remarking that $2k$-ended solutions exhaust all finite Morse index solutions with linear energy growth:

Proposition 4.2.18 (M.). *The following equality of sets holds true*

$$\bigsqcup_{k=1}^{\infty} \mathcal{M}_{2k} = \Big\{ u \in C^\infty_{\mathrm{loc}}(\mathbf{R}^2) \cap L^\infty(\mathbf{R}^2) \textit{ satisfying } (4.1.1), \textit{ and such that}$$

$$\mathrm{ind}(u) < \infty \textit{ and } 0 < \limsup_{R\uparrow\infty} R^{-1}(E \llcorner B_1)[u] < \infty \Big\}. \quad (4.2.28)$$

Remark 4.2.19. *Wang-Wei have recently announced [71] that the* $\lim_{R\uparrow\infty} R^{-1}(E \llcorner B_1)[u] < \infty$ *assumption above is unnecessary; they show it is automatically true whenever* $\mathrm{ind}(u) < \infty$.

Remark 4.2.20. *We assume hypotheses* (H1), (H2), (H3) *throughout this section.*

Proof of Proposition 4.2.18. The "$\subseteq$" direction of (4.2.28) is precisely [35, Theorem 2.8]. The "$\supseteq$" direction essentially follows from the work of Wang in [70], as we explain now. By virtue of [70, Theorems 1.2, 1.3], we know that there exist $k = k(u) \geq 1$, $R = R(u) > 1$, disjoint, embedded curves $\Gamma_1, \ldots, \Gamma_{2k}$, and angles $\underline{\varphi}_1 < \overline{\varphi}_1 = \underline{\varphi}_2 < \ldots < \overline{\varphi}_{2k-1} = \underline{\varphi}_{2k} < \overline{\varphi}_{2k} = 2\pi + \underline{\varphi}_1$ such that

$$\{u = 0\} \setminus B_R(\mathbf{0}) = \sqcup_{i=1}^{2k} \Gamma_i, \quad (4.2.29)$$

where

$$\Gamma_i \subset S_i \triangleq \left\{ r\mathbf{f}(\theta) : r \geq R \text{ and } \underline{\varphi}_i < \theta < \overline{\varphi}_i \right\} \text{ for } i = 1, \ldots, 2k;$$

here, $\mathbf{f}(\theta) = (\cos\theta, \sin\theta) \in \mathbf{R}^2$. Following the argument in [70, Theorem 1.3], we can write each Γ_i as a smooth graph over a ray $\rho_i \triangleq \{r\mathbf{f}(\theta_i) : r \geq R\}$ with $\underline{\varphi}_i < \theta_i < \overline{\varphi}_i$, after possibly enlarging R. That is,

$$\Gamma_i = \{r\mathbf{f}(\theta_i) + h_i(r)\mathbf{J}\mathbf{f}(\theta_i) : r \geq R\}$$

with $h_i : C^\infty([R, \infty))$, for all $i = 1, \ldots, 2k$. We also have

$$\sum_{i=1}^{2k} \mathbf{f}(\theta_i) = \mathbf{0} \in \mathbf{R}^2 \tag{4.2.30}$$

by [70, Theorem 1.1 (v)]. Next, [70, Theorem 3.3] implies that

$$|h_i(r) - \tau_i| \leq C_0 e^{-C_0^{-1} r} \tag{4.2.31}$$

for some $\tau_i = \tau_i(u) \in \mathbf{R}$ and $C_0 = C_0(u) > 0$, and, up to a possible change of sign, that

$$\left| u(r\mathbf{f}(\theta)) - (-1)^{i+1} H(r\sin(\theta - \theta_i) - \tau_i) \right| \leq C_1 e^{-C_1^{-1} r} \tag{4.2.32}$$

for some $C_1 = C_1(u) > 0$ and for all $r \geq R$, $\theta \in (\underline{\varphi}_i, \overline{\varphi}_i)$, $i = 1, \ldots, 2k$. ("Up to a possible change of sign" means that $(-1)^{i+1}$ may have to be replaced by a $(-1)^i$.)

From elliptic regularity [27, Theorem 3.9], (4.2.32) readily implies

$$\left\| \nabla u(r\mathbf{f}(\theta)) - (-1)^i H'(r\sin(\theta - \theta_i) - \tau_i)\mathbf{J}\mathbf{f}(\theta_i) \right\| \leq C_2 e^{-C_2^{-1} r} \tag{4.2.33}$$

and

$$\left\| \nabla^2 u(r\mathbf{f}(\theta)) - (-1)^i H''(r\sin(\theta - \theta_i) - \tau_i)\mathbf{J}\mathbf{f}(\theta_i) \otimes \mathbf{J}\mathbf{f}(\theta_i) \right\| \leq C_2 e^{-C_2^{-1} r}, \tag{4.2.34}$$

for $C_2 = C_2(u, W)$, and all $r \geq R$, $\theta \in (\underline{\varphi}_i, \overline{\varphi}_i)$, $i = 1, \ldots, 2k$. On the other hand, from Proposition 4.1.6, for any $0 < \varepsilon < \frac{1}{2} \min_{i=1,\ldots,2k}\{\overline{\varphi}_i - \theta_i, \theta_i - \underline{\varphi}_i\}$, (4.2.29), and (4.2.31),

$$\sup\left\{ \sum_{\ell=0}^{2} \left\| \nabla^\ell u(r\mathbf{f}(\theta)) \right\| : r \geq R, |\theta - \theta_i| \geq \varepsilon \text{ for all } i = 1, \ldots, 2k \right\} \leq C_3 e^{-C_3^{-1} r}, \tag{4.2.35}$$

where $C_3 = C_3(W, \varepsilon, u)$. It follows from (4.2.32), (4.2.33), (4.2.34), and (4.2.35), that $u - u_\lambda \in W^{2,2}$ if u_λ is an approximate $2k$-ended solution with $\lambda = (\lambda_1, \ldots, \lambda_{2k}) \in \Lambda^{2k}$ given by $\lambda_i \triangleq (\tau_i, \mathbf{f}(\theta_i))$, $i = 1, \ldots, 2k$. The result follows. $\square$

4.3 Technical results

4.3.1 Morse index on spaces with quadratic area growth

In this section we will study general Schrödinger operators

$$L \triangleq -\Delta_g + V, \text{ where } V \in C^\infty_{\mathrm{loc}}(\Sigma) \cap L^\infty(\Sigma) \tag{4.3.1}$$

on complete, noncompact Riemannian manifolds without boundary, and with quadratic volume growth; the latter condition means that their volume measure υ_g satisfies

$$\upsilon_g(B_R(p)) \leq c_\upsilon R^2, \text{ for all } p \in \Sigma^n, R \geq 1. \tag{4.3.2}$$

Example 4.3.1. *The most obvious example of such a manifold is* $(\mathbf{R}^2, \delta)$*. More generally, one may consider isometric products of closed manifolds with* $\mathbf{R}$ *or* $\mathbf{R}^2$*.*

Our goal is to generalize results from Appendix B.2, on eigenvalues of Schrödinger operators on compact domains to a noncompact setting that arises in the study of phase transitions and prove a noncompact Courant nodal domain theorem.

The following lemma is classical:

Lemma 4.3.2 (Logarithmic cutoff functions)**.** *Let* (Σ^n, g) *be complete, noncompact, without boundary, and with quadratic volume growth* (4.3.2)*, and fix a point* $p \in \Sigma$*. For every* $R > 0$*, there exists* $\xi_R \in W^{1,\infty}(\Sigma) \cap C^0_c(\Sigma)$ *such that*

(1) $0 \leq \xi_R \leq 1$, $\xi_R \equiv 1$ *on* $B_R(p)$, $\xi_R \equiv 0$ *outside* $B_{R^2}(p)$,

(2) $\lim_{R\uparrow\infty} \|\nabla \xi_R\|_{L^2(\Sigma)} = 0$, *and*

(3) $\lim_{R\uparrow\infty} \|\nabla \xi_R\|_{L^\infty(\Sigma)} = 0$.

Proof. We have already prescribed the behavior of ξ_R of $B_R(p)$ and $\Sigma \setminus B_{R^2}(p)$, so it remains to define it on $B_{R^2}(p) \setminus B_R(p)$. We do so as

$$\xi_R(x) \triangleq 2 - \frac{\log r}{\log R},$$

where we write $r = r(x)$ for $\mathrm{dist}(x, p)$. Such a function satisfies $\xi_R \in W^{1,\infty}(\Sigma) \cap C^0_c(\Sigma)$, and

$$\nabla \xi_R = -\frac{\nabla r}{r \log R} \text{ a.e. on } \Sigma,$$

and, therefore, that $\lim_{R\uparrow\infty} \|\nabla\xi_R\|_\infty = 0$. Letting σ_g denote $(n-1)$-dimensional hypersurface measure, from the coarea formula and integration by parts, one sees that

$$\begin{aligned}\int_\Sigma \|\nabla\xi_R\|^2\,d\upsilon_g &= \frac{1}{\log^2 R}\int_{B_{R^2}(p)\setminus B_R(p)} \frac{d\upsilon_g}{r^2\log^2 R} = \frac{1}{\log^2 R}\int_R^{R^2} \frac{1}{r^2}\sigma_g(\partial B_r(p))\,dr \\ &= \frac{1}{\log^2 R}\left[\frac{1}{r^2}\upsilon_g(B_r(p))\right]_{r=R}^{R^2} + \frac{2}{\log^2 R}\int_R^{R^2} \frac{\upsilon_g(B_r(p))}{r^3}\,dr.\end{aligned}$$

This goes to zero as $R\uparrow\infty$ because of (4.3.2). $\square$

We recall that the quadratic form associated to L is $\mathcal{Q} : W^{1,2}(\Sigma)\otimes W^{1,2}(\Sigma)\to\mathbf{R}$, with

$$\mathcal{Q}(\zeta,\psi) \triangleq \int_\Sigma \left[\langle\nabla\zeta,\nabla\psi\rangle + V\zeta\psi\right]\,d\upsilon_g,\ \zeta,\psi\in W^{1,2}(\Sigma), \tag{4.3.3}$$

and the corresponding Rayleigh quotient is $\mathcal{Q} : W^{1,2}(\Sigma)\setminus\{0\}\to\mathbf{R}$, with

$$\mathcal{R}[\zeta] \triangleq \frac{\mathcal{Q}(\zeta,\zeta)}{\|\zeta\|^2_{L^2(\Sigma)}},\ \zeta\in W^{1,2}(\Sigma)\setminus\{0\}. \tag{4.3.4}$$

Definition 4.3.3 (Morse index, nullity)**.** *Let (Σ^n, g) be complete, noncompact, without boundary, with quadratic volume growth* (4.3.2)*, let $L \triangleq -\Delta_g + V$, with $V\in C^\infty_{\mathrm{loc}}(\Sigma)\cap L^\infty(\Sigma)$, and suppose $\Omega\subseteq\Sigma$ is an open, connected, Lipschitz domain. We define the Morse index of L on Ω as*

$$\operatorname{ind}(L;\Omega) \triangleq \sup\Big\{\dim V : V\subset W_0^{1,2}(\Omega)\ \textit{a subspace such that}\ \mathcal{Q}(\zeta,\zeta)<0\ \textit{for all}\ \zeta\in W_0^{1,2}(\Omega)\setminus\{0\}\Big\} \tag{4.3.5}$$

and the nullity of L on Ω as

$$\operatorname{nul}(L;\Omega) \triangleq \dim\{u\in W_0^{1,2}(\Omega) : Lu = 0\ \textit{weakly in}\ \Omega\}. \tag{4.3.6}$$

Remark 4.3.4. *We are implicitly (and will often do so) using the fact that $W_0^{1,2}(\Omega)\subseteq W_0^{1,2}(\Sigma)$.*

The Morse index counts the dimensionality of the space instabilities for a particular critical point. Heuristically, this corresponds to the number of negative eigenvalues, counted with multiplicity.

We prove two important results for negative eigenvalues. First, they cannot be “too” negative. Second, when the Morse index is finite, the space of instabilities can be “exhausted” by finitely many eigenfunctions, with negative eigenvalue, in a suitable Sobolev space.

Lemma 4.3.5 (M.)**.** *Let (Σ^n, g) be complete, noncompact, without boundary, and $\Omega\subseteq\Sigma$ be an* *open, connected, Lipschitz domain. For every $f \in W_0^{1,2}(\Omega)$, with $Lf = \lambda f$ weakly in Ω, we have*

$$-\lambda \leq \max\{0, -\inf V\}.$$

Proof. Assume, without loss of generality, that $\|f\|_{L^2(\Omega)} = 1$. Fix $p \in \Omega$, and consider cutoff functions $\chi_R \in W^{1,\infty}(\Sigma)$ so that

$$\chi_R = 1 \text{ on } B_R(p),\ \chi_R = 0 \text{ outside } B_{2R}(p),\ \|\nabla \chi_R\| \leq c_0 R^{-1}.$$

Notice that $\chi_R^2 f \in W_0^{1,2}(\Omega)$, so we may use it as a test function on $Lf = \lambda f$. We get:

$$\begin{aligned}
&\int_\Omega \left[\langle \nabla(\chi_R^2 f), \nabla f\rangle + V\chi_R^2 f^2\right] \, dv_g = \lambda \int_\Omega \chi_R^2 f^2 \, dv_g \\
&\iff \int_\Omega \left[2\chi_R f \langle \nabla \chi_R, \nabla f\rangle + \chi_R^2 \|\nabla f\|^2 + V\chi_R^2 f^2\right] \, dv_g = \lambda \int_\Omega \chi_R^2 f^2 \, dv_g;
\end{aligned}$$

using Cauchy-Schwarz on $2\chi_R f\langle \nabla\chi_R, \nabla f\rangle$,

$$-\lambda \int_\Omega \chi_R^2 f^2 \, dv_g \leq \int_\Omega \left[f^2 \|\nabla \chi_R\|^2 + \max\{0, -\inf V\}\chi_R^2 f^2\right] \, dv_g.$$

The result follows by letting $R \uparrow \infty$. □

The following result should be well-known to experts of elliptic partial differential equations:

Proposition 4.3.6 (Negative eigenfunction representation, cf. Fischer-Colbrie [22, Proposition 2])**.** *Let (Σ^n, g) be complete, noncompact, without boundary, with quadratic volume growth* (4.3.2)*, L as in* (4.3.1)*, and $\Omega \subseteq \Sigma$ an open, connected, Lipschitz domain, such that $k = \operatorname{ind}(L; \Omega) < \infty$. There are $\lambda_1 \leq \ldots \leq \lambda_k < 0$ and L^2-orthonormal functions $\varphi_1, \ldots, \varphi_k \in W_0^{1,2}(\Omega) \cap C^\infty_{\text{loc}}(\Omega)$, such that $L\varphi_i = \lambda_i \varphi_i$ weakly in Ω. Moreover,*

$$\mathcal{Q}(\zeta, \zeta) \geq 0 \textit{ for all } \zeta \in W_0^{1,2}(\Omega) \cap \{\varphi_1, \ldots, \varphi_k\}^\perp, \tag{4.3.7}$$

with equality if and only if $\zeta \in \operatorname{nul}(L; \Omega)$; here, $\perp$ is taken with respect to $L^2(\Omega)$.

Proof. By definition, if $\operatorname{ind}(L; \Omega) = k < \infty$, then there exists $R_0 > 1$ large enough such that $\operatorname{ind}(L; \Omega \cap B_R) = k$ for all $R \geq R_0$. Seeing as to how $\Omega \cap B_R$ is precompact for all $R \geq R_0$, and Lipschitz for a.e. $R \geq R_0$, the theory of Appendix B.2 applies and shows that, for a.e. $R \geq R_0$, there exists a sequence of eigenvalues $\lambda_{1,R} < \lambda_{2,R} \leq \ldots \leq \lambda_{k,R} < 0$ and corresponding eigenfunctions $\varphi_{1,R}, \varphi_{2,R}, \ldots, \varphi_{k,R} \in W_0^{1,2}(\Omega \cap B_R)$ that are L^2-orthonormal. By the monotonicity of Dirichlet eigenvalues (Corollary B.2.3), we know that $R \mapsto \lambda_{i,R}$ is decreasing in R, for each $i = 1, \ldots, k$. In particular, there exists $\mu > 0$ such that

$$\lambda_{i,R} \leq -\mu \text{ for all } i = 1, \dots, k \text{ and a.e. } R \geq R_0. \tag{4.3.8}$$

Moreover, from Lemma 4.3.5, there exists $M = M(V) \geq 0$ such that

$$\lambda_{i,R} \geq -M \text{ for all } i = 1, \dots, k \text{ and a.e. } R \geq R_0. \tag{4.3.9}$$

Claim. There exist $\lambda_1, \dots, \lambda_k$ such that $-M \leq \lambda_1 \leq \dots \leq \lambda_k \leq -\mu$, a corresponding orthonormal sequence $\{\varphi_i\}_{i=1,\dots,k} \subset W_0^{1,2}(\Omega)$, and a subsequence $R_\ell \uparrow \infty$ such that

$$\lim_{\ell\to\infty} \lambda_{i,R_\ell} = \lambda_i \text{ and } \lim_{\ell\to\infty} \varphi_{i,R_\ell} = \varphi_i$$

strongly in $L^2(\Omega)$ and weakly in $W^{1,2}(\Omega)$, with $L\varphi_i = \lambda_i\varphi_i$ weakly in Ω.

Proof of claim. First, note that $L\varphi_i = \lambda_i\varphi_i$ holding true weakly is a direct consequence of all the previous claims, so it suffices to prove those.

In what follows, suppose that $R_0 < S^{1/2} < R^{1/4}$. We will make use of the logarithmic cutoff functions from Lemma 4.3.2. Testing $L\varphi_{i,R} = \lambda_{i,R}\varphi$ in $\Omega \cap B_R$ with $\xi_S^2\varphi_{i,R}$, we find

$$\int_{\Omega\cap B_R} \left[\xi_S^2(\|\nabla\varphi_{i,R}\|^2 + V\varphi_{i,R}^2 - \lambda_{i,R}\varphi_{i,R}^2) - 2\xi_S\varphi_{i,R}\langle\nabla\xi_S, \nabla\varphi_{i,R}\rangle\right] dv_g = 0.$$

Using Cauchy-Schwarz on the last term, $\lambda_{i,R} \leq 0$, $V \in L^\infty$, and conclusions (1) and (3) of Lemma 4.3.2, we get the estimate

$$\int_{\Omega\cap B_S} \|\nabla\varphi_{i,R}\|^2 \, dv_g \leq \int_{\Omega\cap B_R} \xi_S^2\|\nabla\varphi_{i,R}\|^2 \leq c_0 \int_{\Omega\cap B_R} \|\varphi_{i,R}\|^2 = c_0, \tag{4.3.10}$$

for a fixed $c_0 > 0$ that applies for all $R_0 < S^{1/2} < R^{1/4}$, $i = 1, \dots, k$.

Next, recall that $(1-\xi_{S^{1/2}})\varphi_{i,R} \in W_0^{1,2}(\Omega\cap B_R\backslash\overline{B}_S)$ and that, by our choice of R_0, $\operatorname{ind}(L; \Omega\backslash\overline{B}_S) = 0$. From the variational characterization of eigenvalues on compact domains (Proposition B.2.2), and integration by parts, it follows that

$$\begin{aligned} 0 &\leq \int_{\Omega\cap B_R} \left[\|\nabla((1-\xi_{S^{1/2}})\varphi_{i,R})\|^2 + V(1-\xi_{S^{1/2}})^2\varphi_{i,R}^2\right] dv_g \\ &= \int_{\Omega\cap B_R} (1-\xi_{S^{1/2}})^2\varphi_{i,R}L\varphi_{i,R} \, dv_g + \int_{\Omega\cap B_R} \|\nabla(1-\xi_{S^{1/2}})\|^2\varphi_{i,R}^2 \, dv_g. \end{aligned}$$

Recalling that $L\varphi_{i,R} = \lambda_{i,R}\varphi_{i,R}$ classically in $\Omega \cap B_R$, $\lambda_{i,R} \leq -\mu < 0$, and conclusion (3) of Lemma

4.3.2, we obtain, after rearranging,

$$\int_{\Omega\cap B_R\setminus\overline{B}_S}\varphi_{i,R}^2\,dv_g \le \int_{\Omega\cap B_R}(1-\xi_{S^{1/2}})^2\varphi_{i,R}^2\,dv_g \le \mu^{-1}\int_{\Omega\cap B_R}\|\nabla(1-\xi_{S^{1/2}})\|^2\varphi_{i,R}^2\,dv_g$$
$$\le \mu^{-1}\|\nabla\xi_{S^{1/2}}\|^2_{L^\infty(\Sigma)}\int_{\Omega\cap B_R}\varphi_{i,R}^2\,dv_g = \mu^{-1}\|\nabla\xi_{S^{1/2}}\|^2_{L^\infty(\Sigma)} = c_1(\mu,S), \quad (4.3.11)$$

where $\lim_{S\uparrow\infty} c_1(\mu,S) = 0$ for all $\mu > 0$, by conclusion (3) of Lemma 4.3.2. We can now extract a convergent subsequence by recalling that the embedding $W_0^{1,2}(\Omega\cap B_R) \hookrightarrow L^2(\Omega\cap B_R)$ is compact, (4.3.10), and the uniformly decaying exterior L^2 bound (4.3.11). $\square$

Suppose, now, that $\zeta \in C_c^\infty(\Omega)$. Fix n_0 sufficiently large so that $\operatorname{spt}\zeta \subset\subset B_{R_\ell}$ for all $\ell \ge n_0$. Let

$$\zeta_\perp \triangleq \zeta - \sum_{i=1}^k \langle\zeta,\varphi_i\rangle_{L^2(\Omega)}\varphi_i \in W_0^{1,2}(\Omega)\cap\{\varphi_1,\dots,\varphi_k\rangle^\perp,$$
$$\text{and for } \ell \ge n_0,\ \zeta_{\ell,\perp} \triangleq \zeta - \sum_{i=1}^k \langle\zeta,\varphi_{i,R_\ell}\rangle_{L^2(\Omega\cap B_{R_\ell})}\varphi_{i,R_\ell} \in W_0^{1,2}(\Omega\cap B_{R_\ell})\cap\{\varphi_{1,R_\ell},\dots,\varphi_{k,R_\ell}\}^\perp,$$

where $\perp$ is taken with respect to $L^2(\Omega)$ and $L^2(\Omega\cap B_{R_\ell})$, respectively. Then, by the strong L^2 convergence, and the fact that $\mathcal{Q}(\zeta_{\ell,\perp},\zeta_{\ell,\perp}) \ge 0$ from the variational characterization of eigenvalues in the compact setting (Proposition B.2.2) and $\operatorname{ind}(L;\Omega\cap B_{R_\ell}) = k$, we have

$$\begin{aligned}\mathcal{Q}(\zeta_\perp,\zeta_\perp) &= \mathcal{Q}(\zeta,\zeta) - \sum_{i=1}^k \lambda_i\langle\zeta,\varphi_i\rangle^2_{L^2(\Omega)}\\ &= \lim_{\ell\to\infty}\left(\mathcal{Q}(\zeta,\zeta) - \sum_{i=1}^k \lambda_{i,R_\ell}\langle\zeta,\varphi_{i,R_\ell}\rangle^2_{L^2(\Omega\cap B_{R_\ell})}\right)\\ &= \lim_{\ell\to\infty}\mathcal{Q}(\zeta_{\ell,\perp},\zeta_{\ell,\perp}) \ge 0.\end{aligned}$$

Inequality (4.3.7) follows from this, since $C_c^\infty(\Omega)$ is dense in $W_0^{1,2}(\Omega)$.

For the rigidity case, we proceed as follows. A function $\zeta \in W_0^{1,2}(\Omega)\cap\{\varphi_1,\dots,\varphi_k\}^\perp$ attaining equality in (4.3.7) will be a global minimizer for

$$W_0^{1,2}(\Omega) \ni \psi \mapsto \mathcal{Q}(\psi_\perp,\psi_\perp) = \mathcal{Q}(\psi,\psi) - \sum_{i=1}^k \lambda_i\langle\psi,\varphi_i\rangle^2_{L^2(\Omega)},$$

so it will also be a critical point, i.e.,

$$\mathcal{Q}(\zeta,\psi) - \sum_{i=1}^k \lambda_i\langle\zeta,\varphi_i\rangle_{L^2(\Omega)}\langle\psi,\varphi_i\rangle_{L^2(\Omega)} = 0 \text{ for all } \psi \in W_0^{1,2}(\Omega);$$

however, $\zeta \in W_0^{1,2}(\Omega) \cap \{\varphi_1, \ldots, \varphi_k\}^\perp$, so $Q(\zeta, \psi) = 0$ for all ψ, and the claim follows. $\square$

From this we get:

Theorem 4.3.7 (Noncompact Courant nodal domain theorem, M., cf. Montiel-Ros [46, Lemma 12])**.** *Let (Σ^n, g) be complete, noncompact, without boundary, and with quadratic volume growth* (4.3.2)*. Suppose the open, connected, Lipschitz domain Ω can be partitioned into open, connected, disjoint, Lipschitz domains $\Omega_1, \ldots, \Omega_m$. For any L as in* (4.3.1)*,*

$$\operatorname{ind}(L;\Omega) \geq \sum_{i=1}^{m-1} (\operatorname{ind}(L;\Omega_i) + \operatorname{nul}(L;\Omega_i)) + \operatorname{ind}(L;\Omega_m). \tag{4.3.12}$$

Proof. Without loss of generality, we may assume $\operatorname{ind}(L;\Omega) < \infty$. Invoking Proposition 4.3.6 $m+1$ times, we obtain:

(1) $N \subset W_0^{1,2}(\Omega)$ such that $\dim N = \operatorname{ind}(L;\Omega)$,

(2) $N_i \subset W_0^{1,2}(\Omega_i)$ such that $\dim N_i = \operatorname{ind}(L;\Omega)$, $i = 1, \ldots, m$.

Likewise, let $K_i = \operatorname{nul}(L;\Omega_i)$, $i = 1, \ldots, m-1$.

Suppose, for the sake of contradiction, that (4.3.12) were false. By an elementary dimension counting argument, there would exist a choice of f_i, $i = 1, \ldots, m$, where $f_i \in N_i \oplus K_i$, $i = 1, \ldots, m-1$, and $f_m \in N_m$, such that $f = f_1 + \ldots + f_m \in N^\perp \setminus \{0\}$, where $\perp$ is taken with respect to $L^2(\Omega)$, and we have extended each f_i by zero on $\Omega \setminus \Omega_i$. By Proposition 4.3.6,

$$\mathcal{Q}(f,f) \geq 0, \tag{4.3.13}$$

with equality if and only if $f \in \operatorname{nul}(L;\Omega)$. On the other hand, since $\operatorname{spt} f_i \subseteq \overline{\Omega}_i$, and the domains Ω_i are Lipschitz and partition Ω,

$$\mathcal{Q}(f,f) = \sum_{i=1}^{m} \mathcal{Q}(f_i, f_i) \leq 0, \tag{4.3.14}$$

with equality if and only if $f_i \in K_i$, $i = 1, \ldots, m-1$, and $f_m = 0$ a.e. From (4.3.13) and (4.3.14) we see that equality must hold. In particular, $f \in \operatorname{nul}(L;\Omega)$ and $f_m = 0$ a.e. on N_m. In particular, since $f = f_1 + \ldots + f_m$ and $f_1, \ldots, f_{m-1}$ have no essential support in Ω_m, it must be that $f = 0$ a.e. in Ω_m. Since $Lf = 0$ weakly on Ω, unique continuation (see [24, Theorem 4.2]) forces $f = 0$ a.e. in Ω, a contradiction to our having chosen $f \in N^\perp \setminus \{0\}$. $\square$

Finally, the following proposition will be very useful in the noncompact setting. It is motivated by Ghoussoub-Gui's original proof of Conjecture 4.1.11 in $\mathbf{R}^2$ [25, Theorem 1.1].

Lemma 4.3.8 (M.)**.** *Let (Σ^n, g) be complete, noncompact, without boundary, with quadratic area growth* (4.3.2)*, and let L be as in* (4.3.1)*. Suppose $u \not\equiv 0$ is a bounded Jacobi field. If $\Omega \subseteq \Sigma$ is an open, connected, Lipschitz domain, with $u|_{\partial\Omega} \equiv 0$, then on every open, connected, Lipschitz $\Omega' \supsetneq \Omega$,* $\operatorname{ind}(L; \Omega') \geq 1$.

Proof. We define the auxiliary function $u' = u\mathbf{1}_\Omega$. We will argue by contradiction by means of the following claim:

Claim. If $\operatorname{ind}(L; \Omega') = 0$, then $Lu' = 0$ weakly on Ω'.

Proof of claim. Consider cutoff functions ξ_R as in Lemma 4.3.2. By elliptic regularity (e.g., [27, Corollary 8.11]), $u \in \ker(L)$ implies that $Lu = 0$ holds classically in all of Σ. Multiplying by $\xi_R u$ and integrating by parts over the nodal domain Ω (recall $u|_{\partial\Omega} \equiv 0$), we get

$$\begin{aligned}
0 &= \int_\Omega (-\Delta_g u + Vu)\xi_R^2 u \, dv_g \\
&= \int_\Omega \left[\xi_R^2 \|\nabla u\|^2 + V\xi_R^2 u^2 + 2\xi_R u \langle \nabla \xi_R, \nabla u\rangle\right] \, dv_g \\
&= \int_\Omega \left[\|\xi_R \nabla u + u \nabla \xi_R\|^2 + V\xi_R^2 u^2 - \|\nabla \xi_R\|^2 u^2\right] \, dv_g \\
&= \int_\Omega \left[\|\nabla(\xi_R u)\|^2 + V\xi_R^2 u^2 - \|\nabla \xi_R\|^2 u^2\right] \, dv_g.
\end{aligned}$$

From the definition of $u' = u\mathbf{1}_\Omega \in W_0^{1,2}(\Omega')$, we find—by rearranging—that

$$\int_{\Omega'} \left[\|\nabla(\xi_R u')\|^2 + V\xi_R^2 (u')^2\right] \, dv_g = \int_{\Omega'} \|\nabla \xi_R\|^2 (u')^2 \, dv_g.$$

Thus, the Rayleigh quotient $\mathcal{R}$ (4.3.4) satisfies

$$\mathcal{R}[\xi_R u'] = \|\xi_R u'\|_{L^2(\Omega')}^{-2} \int_{\Omega'} \|\nabla \xi_R\|^2 (u')^2 \, dv_g. \tag{4.3.15}$$

From Lemma 4.3.2 and $u' \in L^\infty$, we find that

$$\lim_{R\uparrow\infty} \int_{\Omega'} \|\nabla \xi_R\|^2 (u')^2 \, dv_g = 0, \tag{4.3.16}$$

and, therefore, since $u' \not\equiv 0$ a.e.,

$$\lim_{R\uparrow\infty} \mathcal{R}[\xi_R u'] = 0. \tag{4.3.17}$$

Since we're assuming $\operatorname{ind}(L; \Omega') = 0$, Proposition 4.3.6 and (4.3.17) together show that $R \mapsto \xi_R u'$ is a minimizing sequence for the Banach space operator $\mathcal{R} : W_0^{1,2}(\Omega') \to \mathbf{R} \cup \{\infty\}$, which is bounded from below by zero (by assumption), and which is lower semicontinuous everywhere and Gâteaux-differentiable wherever it is finite.

We apply Corollary B.4.2 at each $\xi_R u'$, with parameters

$$\varepsilon_R \triangleq \mathcal{R}[\xi_R u'],\ \delta_R \triangleq \left(\int_{\Omega'} \|\nabla \xi_R\|^2 (u')^2\, dv_g\right)^{1/2}. \tag{4.3.18}$$

We get $\psi_R \in W_0^{1,2}(\Omega')$ with

$$\mathcal{R}[\psi_R] \le \mathcal{R}[\xi_R u'] = \varepsilon_R,\ \|\xi_R u' - \psi_R\|_{W^{1,2}(\Omega')} \le \delta_R,\ \text{and } \|\delta\mathcal{R}[\psi_R]\|_{W_0^{1,2}(\Omega')^*} \le \frac{\varepsilon_R}{\delta_R}. \tag{4.3.19}$$

Let us first make some observations. First, by our choice of parameters and (4.3.15), we have

$$\varepsilon_R \|\xi_R u'\|^2_{L^2(\Omega')} = \int_{\Omega'} \|\nabla \xi_R\|^2 (u')^2\, dv_g = \delta_R^2 \implies \|\xi_R u'\|_{L^2(\Omega')} = \varepsilon_R^{-1/2}\delta_R. \tag{4.3.20}$$

Second, by the triangle inequality, (4.3.19), and (4.3.20), we have

$$\|\psi_R\|_{L^2(\Omega')} \le \|\xi_R u'\|_{L^2(\Omega')} + \delta_R \le (1+\varepsilon_R^{-1/2})\delta_R. \tag{4.3.21}$$

Third, by (4.3.16) and $u' \not\equiv 0$ a.e., the rightmost term in (4.3.20) satisfies

$$\lim_R \frac{\varepsilon_R}{\delta_R} = \lim_R \|\xi_R u'\|^{-2}_{L^2(\Omega')} \left(\int_{\Omega'} \|\nabla \xi_R\|^2 (u')^2\, dv\right)^{1/2} = 0. \tag{4.3.22}$$

Recall that the Gâteaux differential satisfies

$$\begin{aligned}\delta\mathcal{R}[\psi_R]\{\zeta\} &= 2\left[\frac{\int_{\Omega'} \left[\langle \nabla\psi_R, \nabla\zeta\rangle + V\psi_R\zeta\right]\, dv_g}{\int_{\Omega'} \psi_R^2\, dv_g} - \frac{\int_{\Omega'} \left[\|\nabla\psi_R\|^2 + V\psi_R^2\right]\, dv_g}{\int_{\Omega'} \psi_R^2\, dv_g} \frac{\int_{\Omega'} \psi_R\zeta\, dv_g}{\int_{\Omega'} \psi_R^2\, dv_g}\right] \\ &= 2\|\psi_R\|^{-2}_{L^2(\Omega')}\mathcal{Q}(\psi_R,\zeta) - 2\|\psi_R\|^{-2}_{L^2(\Omega')}\mathcal{R}[\psi_R]\int_{\Omega'} \psi_R\zeta\, dv_g\end{aligned}$$

and, therefore,

$$\mathcal{Q}(\psi_R,\zeta) = \mathcal{R}[\psi_R]\int_{\Omega'} \psi_R\zeta\, dv_g + \frac{1}{2}\|\psi_R\|^2_{L^2(\Omega')}\delta\mathcal{R}[\psi_R]\{\zeta\}. \tag{4.3.23}$$

We plan to let $R \uparrow \infty$ while holding ζ fixed. For convenience, assume for now that $\zeta \in C_c^\infty(\Omega')$. We have, from the triangle inequality, (4.3.19), and (4.3.21), that

$$\|\psi_R\|^2_{L^2(\Omega)}\|\delta\mathcal{R}[\psi_R]\|_{W_0^{1,2}(\Omega')^*} \le (1+\varepsilon_R^{-1/2})^2\delta_R^2\frac{\varepsilon_R}{\delta_R} = (1+\varepsilon^{1/2})^2\delta_R \to 0 \text{ as } R \uparrow \infty.$$

Moreover, Cauchy-Schwarz, (4.3.19), and (4.3.21),

$$\mathcal{R}[\psi_R]\int_{\Omega'}\psi_R\zeta\,dv_g \leq \frac{\varepsilon_R}{\delta_R}(1+\varepsilon_R^{-1/2})\delta_R\|\zeta\|_{L^2(\Omega')} = (\varepsilon_R+\varepsilon_R^{1/2})\|\zeta\|_{L^2(\Omega')} \to 0 \text{ as } R\uparrow\infty.$$

Plugging these last two inequalities into (4.3.23), we get

$$\lim_{R\uparrow\infty}\mathcal{Q}(\psi_R,\zeta)=0. \tag{4.3.24}$$

On the other hand, since $\operatorname{spt}\zeta\subset\subset\Sigma$, from (4.3.19) and the definition of ξ_R, we see that

$$\lim_{R\uparrow\infty}\mathcal{Q}(\psi_R,\zeta)=\lim_{R\uparrow\infty}\mathcal{Q}(\xi_R u',\zeta)=\mathcal{Q}(u',\zeta).$$

From this, together with (4.3.24), it follows that $\mathcal{Q}(u',\zeta)=0$ for all $\zeta\in C_c^\infty(\Omega')$. It is well-known that $C_c^\infty(\Omega')$ is dense in $W_0^{1,2}(\Omega')$, so $\mathcal{Q}(u',\zeta)=0$ for all $\zeta\in W_0^{1,2}(\Omega')$, and the claim follows. □

Let's now see how the result follows from this claim. Indeed, if $Lu'=0$ weakly in Ω', and $u'=0$ a.e. in $\Omega'\setminus\Omega$, then by unique continuation (see [24, Theorem 4.2]) we would have $u'=0$ in Ω', so $u=0$ a.e. in Ω, so by elliptic unique continuation (again) we would have $u=0$ in Σ—a contradiction. □

Appendix A

Riemannian Geometry

A.1 Curvature

In this appendix we survey some well-known facts about curvature that are needed in this thesis. We refer the reader to [55] for a thorough treatment in the Riemannian setting and to [48] for the Lorentzian setting.

Definition A.1.1 (Intrinsic curvature)**.** *Let (M^n, g) be a Riemannian manifold. We define:*

(1) the Riemann curvature as the covariant 4-tensor

$$\mathrm{Rm}_M(\mathbf{X},\mathbf{Y},\mathbf{Z},\mathbf{W}) \triangleq \langle \nabla_{\mathbf{X}}\nabla_{\mathbf{Y}}\mathbf{Z} - \nabla_{\mathbf{Y}}\nabla_{\mathbf{X}}\mathbf{Z} - \nabla_{[\mathbf{X},\mathbf{Y}]}\mathbf{Z}, \mathbf{W}\rangle,\ \mathbf{X},\mathbf{Y},\mathbf{Z},\mathbf{W} \in \Gamma(TM);$$

(2) the Ricci curvature as the covariant 2-tensor

$$\mathrm{Ric}_M(\mathbf{X},\mathbf{Y}) \triangleq \mathrm{Tr}_M \mathrm{Rm}(\mathbf{X},\cdot,\cdot,\mathbf{Y}),\ \mathbf{X},\mathbf{Y} \in \Gamma(TM);$$

(3) its scalar curvature as the function

$$R_M \triangleq \mathrm{Tr}_M \mathrm{Ric}_M\,.$$

When $n = 2$, we also define the Gauss curvature as the function

$$K_M \triangleq \frac{1}{2} R_M.$$

The following properties are all well-known:

Proposition A.1.2 (Curvature properties, cf. Petersen [55, Section 3.2]). *Let (M^n, g) be a Riemannian manifold. Its curvature tensors have the following properties:*

(1) $\mathrm{Rm}_M(\mathbf{X},\mathbf{Y},\mathbf{Z},\mathbf{W}) = -\mathrm{Rm}_M(\mathbf{Y},\mathbf{X},\mathbf{Z},\mathbf{W}) = -\mathrm{Rm}_M(\mathbf{X},\mathbf{Y},\mathbf{W},\mathbf{Z}) = \mathrm{Rm}_M(\mathbf{Z},\mathbf{W},\mathbf{X},\mathbf{Y})$*;*

(2) $\mathrm{Ric}_M(\mathbf{X},\mathbf{Y}) = \mathrm{Ric}_M(\mathbf{Y},\mathbf{X})$*;*

(3) 1st Bianchi identity:

$$\mathrm{Rm}_M(\mathbf{X},\mathbf{Y},\mathbf{Z},\mathbf{W}) + \mathrm{Rm}_M(\mathbf{Y},\mathbf{Z},\mathbf{X},\mathbf{W}) + \mathrm{Rm}_M(\mathbf{Z},\mathbf{X},\mathbf{Y},\mathbf{W}) = 0;$$

(4) 2nd Bianchi identity

$$\nabla_{\mathbf{X}} \mathrm{Rm}_M(\mathbf{Y},\mathbf{Z},\mathbf{W},\mathbf{V}) + \nabla_{\mathbf{Y}} \mathrm{Rm}_M(\mathbf{Z},\mathbf{X},\mathbf{W},\mathbf{V}) + \nabla_{\mathbf{Z}} \mathrm{Rm}_M(\mathbf{Y},\mathbf{X},\mathbf{W},\mathbf{V}) = 0;$$

(5) traced 2nd Bianchi identity on 1-forms: $\mathrm{div}_M \mathrm{Ric}_M = \frac{1}{2} dR_M$.

There are also extrinsic notions of curvature, that describe the interaction of Riemannian submanifolds with their ambient Riemannian manifolds.

Definition A.1.3 (Extrinsic curvature). *Let (M^n, g) be a Riemannian manifold and Σ^{n-1} be a hypersurface with unit normal ν. The second fundamental form of Σ with respect to ν is a symmetric covariant 2-tensor field on Σ defined by*

$$\mathrm{II}_\Sigma(\mathbf{X},\mathbf{Y}) \triangleq \langle \nabla_{\mathbf{X}} \nu, \mathbf{Y} \rangle \text{ for all } \mathbf{X},\mathbf{Y} \in \Gamma(T\Sigma); \tag{A.1.1}$$

where ∇ is the ambient Levi-Civita connection. The mean curvature of Σ with respect to ν is function on Σ defined by

$$H_\Sigma \triangleq \mathrm{div}_\Sigma \nu = \mathrm{Tr}_\Sigma \mathrm{II}_\Sigma, \tag{A.1.2}$$

and the mean curvature vector of Σ is defined by

$$\mathbf{H}_\Sigma \triangleq -H_\Sigma \nu. \tag{A.1.3}$$

Remark A.1.4. *The mean curvature vector $\mathbf{H}_\Sigma$ is independent of our choice of ν.*

The intrinsic curvature of a hypersurface and the intrinsic curvature of ambient space are related explicitly by the Gauss and Codazzi equations:

Proposition A.1.5 (Gauss, Codazzi equations, cf. Petersen [55, Section 4.2]). *Let (M^n, g) be a*

Riemannian manifold and Σ^{n-1} be a hypersurface with unit normal ν and second fundamental form II_Σ with respect to ν. Then:

(1) the Gauss equation says:

$$\mathrm{Rm}_\Sigma(\mathbf{X},\mathbf{Y},\mathbf{Z},\mathbf{W}) = \mathrm{Rm}_M(\mathbf{X},\mathbf{Y},\mathbf{Z},\mathbf{W}) + \mathrm{II}_\Sigma(\mathbf{X},\mathbf{W})\,\mathrm{II}_\Sigma(\mathbf{Y},\mathbf{Z}) \\ - \mathrm{II}_\Sigma(\mathbf{X},\mathbf{Z})\,\mathrm{II}_\Sigma(\mathbf{Y},\mathbf{W}) \text{ for } \mathbf{X},\mathbf{Y},\mathbf{Z},\mathbf{W} \in \Gamma(T\Sigma), \text{ and}$$

(2) the Codazzi equation says:

$$\mathrm{Rm}(\mathbf{X},\mathbf{Y},\mathbf{Z},\nu) = \nabla^\Sigma_{\mathbf{Y}}\,\mathrm{II}_\Sigma(\mathbf{X},\mathbf{Z}) - \nabla^\Sigma_{\mathbf{X}}\,\mathrm{II}_\Sigma(\mathbf{Y},\mathbf{Z}) \text{ for } \mathbf{X},\mathbf{Y},\mathbf{Z} \in \Gamma(T\Sigma).$$

The following lemma will be used frequently in this thesis. It is well-known to the experts, but we prove it here again for the sake of completeness.

Lemma A.1.6 (M.)**.** *Let (M^n, g) be a Riemannian manifold and Σ^{n-1} be a hypersurface. Consider a smooth family of maps $F : \Sigma \times (-\varepsilon, \varepsilon) \to M$. Suppose, additionally, that $\Sigma_t^{n-1} \triangleq F(\Sigma^{n-1}, t)$, $\nabla_{\partial/\partial t}F \perp \Sigma_t$, and $u \triangleq \langle \nabla_{\partial/\partial t}F, \nu\rangle > 0$, where ν denotes a unit normal vector field to Σ_t. Setting $\gamma(t) \triangleq F(\cdot,t)^*g \in \mathrm{Met}(\Sigma^{n-1})$, so that, locally, $F^*g = \gamma(t) + u^2dt^2$. Then:*

(1) the evolution of $\gamma(t)$ is related to the second fundamental form II_{Σ_t} of Σ_t with respect to ν by

$$\frac{\partial}{\partial t}\gamma(t) = 2\,\mathrm{II}_{\Sigma_t}\,u,\ \mathbf{X},\mathbf{X} \in \Gamma(T\Sigma);$$

(2) the evolution of the induced hypersurface volume form $d\upsilon_t \triangleq F(\cdot,t)^*d\upsilon_g$ is related to the mean curvature H_{Σ_t} of Σ_t with respect to ν by

$$\frac{\partial}{\partial t}d\upsilon_t = \frac{1}{2}\,\mathrm{Tr}_{\gamma(t)}\,\frac{\partial}{\partial t}\gamma(t) = H_{\Sigma_t}u;$$

(3) the evolution of the mean curvature H_{Σ_t} is given by the so-called "Jacobi equation:"

$$\frac{\partial}{\partial t}H_{\Sigma_t} = -\Delta_{\Sigma_t}u - \left(\|\,\mathrm{II}_{\Sigma_t}\,\|^2 + \mathrm{Ric}_M(\nu,\nu)\right)u, \text{ and}$$

(4) the ambient scalar curvature is given in terms of the scalar curvature of the foliating surfaces and their extrinsic curvature by

$$R_M = \frac{2}{u}\left(-\Delta_{\Sigma_t}u + \frac{1}{2}R_{\Sigma_t}u\right) - H^2_{\Sigma_t} - \|\,\mathrm{II}_{\Sigma_t}\,\|^2 - \frac{2}{u}\frac{\partial}{\partial t}H_{\Sigma_t}.$$

Proof. We will prove (1), (2), and (4) here, deferring to [12, Section 1.8] for a proof of (3).

Without loss of generality, we may compute these formulas at $t = 0$, where $F(\cdot, 0) = \mathrm{Id}$. Let $p \in \Sigma$. For all $\mathbf{X}$, $\mathbf{Y} \in T_p\Sigma$,

$$\begin{aligned}\left[\frac{\partial}{\partial t}\langle \mathbf{X}, \mathbf{Y}\rangle_{\gamma(t)}\right]_{t=0} &= \left[\frac{\partial}{\partial t}\langle \nabla_{F(\cdot,t)_*\mathbf{X}} F(\cdot,t), \nabla_{F(\cdot,t)_*\mathbf{Y}} F(\cdot,t)\rangle\right]_{t=0} \\ &= \langle \nabla_{\partial/\partial t}\nabla_{\mathbf{X}} F(\cdot,t), \mathbf{Y}\rangle + \langle \mathbf{X}, \nabla_{\partial/\partial t}\nabla_{\mathbf{Y}} F(\cdot,t)\rangle \\ &= \langle \nabla_{\mathbf{X}}\nabla_{\partial/\partial t} F(\cdot,t), \mathbf{Y}\rangle + \langle \mathbf{X}, \nabla_{\mathbf{Y}}\nabla_{\partial/\partial t} F(\cdot,t)\rangle \\ &= \langle \nabla_{\mathbf{X}}(u\nu), \mathbf{Y}\rangle + \langle \mathbf{X}, \nabla_{\mathbf{Y}}(u\nu)\rangle = 2\,\mathrm{II}_\Sigma(\mathbf{X},\mathbf{Y})u. \qquad \text{(A.1.4)}\end{aligned}$$

We may trace (A.1.4) over Σ to get the identity for H_Σ.

The scalar curvature identity is proven as follows. Tracing the Gauss equation (Proposition A.1.5) of $\Sigma_t \subset M$, twice, we get

$$R_{\Sigma_t} = R_M - 2\,\mathrm{Ric}_M(\nu,\nu) + H_{\Sigma_t}^2 - \|\,\mathrm{II}_{\Sigma_t}\,\|^2. \qquad \text{(A.1.5)}$$

Then, we know from the Jacobi equation on the family Σ_t that moves orthogonally that

$$\frac{\partial}{\partial t} H_{\Sigma_t} = -\Delta_{\Sigma_t} u - \left(\|\,\mathrm{II}_{\Sigma_t}\,\|^2 + \mathrm{Ric}_M(\nu,\nu)\right) u. \qquad \text{(A.1.6)}$$

Rearranging (A.1.6) we can get an expression for $\mathrm{Ric}_M(\nu,\nu)$. Plugging it into (A.1.5) we get

$$R_{\Sigma_t} = R_M - \frac{2}{u}\left(-\Delta_{\Sigma_t} u - \|\,\mathrm{II}_{\Sigma_t}\,\|^2 u - \frac{\partial}{\partial t} H_{\Sigma_t}\right) + H_{\Sigma_t}^2 - \|\,\mathrm{II}_{\Sigma_t}\,\|^2,$$

which gives the required result upon rearranging. ∎

Lemma A.1.7 (Conformal metric transformations, cf. Schoen-Yau [63, Chapter 5])**.** *Let (M^n, g) be a Riemannian manifold and $\Sigma^{n-1} \subset M^n$ a hypersurface with unit normal ν. If $n = 2$ (i.e., $\Sigma^1 \subset M^2$ is a curve) and $\widetilde{g} = e^{2f} g$, then*

(1) the Laplace-Beltrami operator of M transforms as: $\Delta_{\widetilde{g}} h = e^{-2f}\Delta_g h$;

(2) the Gauss curvature of M transforms as: $K_{\widetilde{g}} = e^{-2f}(K_g - \Delta_g f)$;

(3) the geodesic curvature of Σ with respect to ν transforms as: $\kappa_{\widetilde{g}} = e^{-f}(\kappa_g + \nabla_\nu^g f)$.

If $n \geq 3$ and $\widetilde{g} = u^{\frac{4}{n-2}} g$, then

(1) the Laplace-Beltrami operator of M transforms as:

$$\Delta_{\widetilde{g}} h = u^{-\frac{4}{n-2}} \left(\Delta_g h + 2\langle \nabla^g \log u, \nabla^g h\rangle_g\right);$$

(2) the scalar curvature of M transforms as:

$$R_{\widetilde{g}} = \frac{4(n-1)}{n-2} u^{-\frac{n+2}{n-2}} \left(-\Delta_g u + \frac{n-2}{4(n-1)} R_g u\right);$$

(3) the mean curvature of Σ with respect to ν transforms as:

$$H_{\widetilde{g}} = u^{-\frac{2}{n-2}} \left(H_g + \frac{2(n-1)}{n-2} u^{\frac{-4}{n-2}} \nabla^g_\nu u\right).$$

Proof. The transformation of the Laplace-Beltrami operator, the Gauss curvature, and the scalar curvature are all well known. (See, for example, [63, Chapter 5].) We compute the transformation of the mean curvature and geodesic curvature, which are less well documented.

Let us first compute the change in mean curvature in the case $n \geq 3$. If $\zeta \in C_c^\infty(\Sigma^{n-1})$, and we apply an infinitesimal normal perturbation to $\Sigma \subset M$ in the direction $\zeta\widetilde{\nu}$, where $\widetilde{\nu}$ is the $\widetilde{g}$-unit normal to $\Sigma \subset M$. Denote the family of hypersurfaces that are traced out by Σ_t, with $\Sigma_0 = \Sigma$, then, because $d\sigma_{\widetilde{g}} = (u^{\frac{4}{n-2}})^{\frac{n-1}{2}} d\sigma_g$, we see that

$$\left[\frac{d}{dt}\sigma_{\widetilde{g}}(\Sigma_t)\right]_{t=0} = \int_\Sigma H_{\widetilde{g}} \zeta \, d\sigma_{\widetilde{g}} = \int_\Sigma H_{\widetilde{g}} \zeta u^{\frac{2(n-1)}{n-2}} \, d\sigma_g.$$

On the other hand, since the change is conformal, we can also realize the $\widetilde{g}$-normal variation Σ_t as a modified g-normal variation, with velocity $\zeta u^{-\frac{2}{n-2}}\nu$. Thus,

$$\begin{aligned}\left[\frac{d}{dt}\sigma_{\widetilde{g}}(\Sigma_t)\right]_{t=0} &= \left[\frac{d}{dt}\int_{\Sigma_t} u^{\frac{2(n-1)}{n-2}} \, d\sigma_g\right]_{t=0} = \int_\Sigma \left[H_g \zeta u^2 + \zeta u^{\frac{-2}{n-2}} \nabla^g_\nu u^{\frac{2(n-1)}{n-2}}\right] d\sigma_g \\ &= \int_\Sigma \left[H_g \zeta u^2 + \frac{2(n-1)}{n-2} \zeta u^{-\frac{2n}{n-2}} \nabla^g_\nu u\right] d\sigma_g.\end{aligned}$$

Since these two expressions have to coincide for all ζ, $H_{\widetilde{g}} = u^{-\frac{2}{n-2}} \left(H_g + \frac{2(n-1)}{n-2} u^{\frac{-4}{n-2}} \nabla^g_\nu u\right)$. We can compute the change in mean curvature when $n = 2$ and $\widetilde{g} = e^{2f} g$ similarly. We would have

$$\left[\frac{d}{dt}\ell_{\widetilde{g}}(\Sigma_t)\right]_{t=0} = \int_\Sigma \kappa_{\widetilde{g}} \zeta \, d\ell_{\widetilde{g}} = \int_\Sigma \kappa_{\widetilde{g}} \zeta e^f \, d\ell_g.$$

Likewise, a $\zeta\widetilde{\nu}$ $\widetilde{g}$-normal movement corresponds to a $\zeta e^{-f}\nu$ g-normal movement, so:

$$\left[\frac{d}{dt}\ell_{\widetilde{g}}(\Sigma_t)\right]_{t=0} = \left[\frac{d}{dt}\int_{\Sigma_t} e^f \, d\ell_g\right]_{t=0} = \int_\Sigma \left[\kappa_g \zeta + \zeta e^{-f}\nabla^g_\nu e^f\right] \, d\ell_g = \int_\Sigma \left[\kappa_g\zeta + \zeta \nabla_\nu f\right] \, d\ell_g.$$

In particular $\kappa_{\widetilde{g}} = e^{-f}(\kappa_g + \nabla^g_\nu f)$, as claimed. □

A.2 Minimal hypersurfaces

We briefly outline some basic facts about minimal hypersurfaces here, with few proofs. For a thorough treatment of these results, which are classical by now, we refer the reader to [12].

Definition A.2.1 (Totally geodesic and minimal hypersurfaces)**.** *A hypersurface Σ^{n-1} in a Riemannian manifold (M^n, g) is said to be totally geodesic if its second fundamental form with respect to a local choice of unit normal vector ν, $\mathrm{I\!I}_\Sigma$, vanishes. More generally, it is said to be minimal if the mean curvature vector $\mathbf{H}_\Sigma$ vanishes.*

The proposition gives a variational characterization of "minimality."

Proposition A.2.2 (cf. Colding-Minicozzi [12, Section 1.1])**.** *A complete hypersurface Σ^{n-1} in a Riemannian manifold (M^n, g) is minimal if and only if for every variation $F_t \in \mathrm{Diff}(M)$, $t \in (-\varepsilon, \varepsilon)$ with $\{x \in M : F_t(x) \neq F_0(x)\} \subseteq K \subset\subset \operatorname{int} M \setminus \partial\Sigma$ for all $t \in (-\varepsilon, \varepsilon)$, and $F_0 = \mathrm{Id}_M$, we have*

$$\left[\frac{d}{dt}\operatorname{area}(F_t(\Sigma \cap K))\right]_{t=0} = 0.$$

In other words, minimal surfaces are critical points of the area functional. The following definition captures those minimal surfaces that are precisely stable critical points of the area functional.

Definition A.2.3 (Stable minimal hypersurfaces)**.** *A complete minimal hypersurface Σ^{n-1} inside a Riemannian manifold (M^n, g) is said to be stable if for every variation $F_t \in \mathrm{Diff}(M)$, $t \in (-\varepsilon, \varepsilon)$ with $\{x \in M : F_t(x) \neq F_0(x)\} \subseteq K \subset\subset \operatorname{int} M \setminus \partial\Sigma$ for all $t \in (-\varepsilon, \varepsilon)$, and $F_0 = \mathrm{Id}_M$, we have*

$$\left[\frac{d^2}{dt^2}\operatorname{area}(F_t(\Sigma \cap K))\right]_{t=0} \geq 0.$$

When the minimal hypersurface admits a globally defined unit normal vector field, we can check for its stability using the so-called **Jacobi operator**:

Proposition A.2.4 (cf. Colding-Minicozzi [12, Section 1.8])**.** *A complete minimal hypersurface*

Σ^{n-1} with unit normal vector ν in a Riemannian manifold (M^n, g) is stable if and only if

$$J_\Sigma \triangleq -\Delta_\Sigma - \| \mathrm{II}_\Sigma \|^2 - \mathrm{Ric}_M(\nu, \nu)$$

is nonnegative definite on $C_c^\infty(\Sigma)$.

Elements of the nullspace of the Jacobi operator are called Jacobi fields of the minimal surface.

Example A.2.5 (Common examples of minimal surfaces). *The following hypersurfaces are all minimal with respect to their ambient space:*

(1) any hyperplane $P^{n-1} \subset \mathbf{R}^n$,

(2) the catenoid $\{(x, y, z) \in \mathbf{R}^3 : x^2 + y^2 = \cosh^2 z\} \subset \mathbf{R}^3$,

(3) any equatorial $\mathbf{S}^{n-1} \subset \mathbf{S}^n$,

(4) the Clifford torus

$$\left\{ (x_1, x_2, x_3, x_4) \in \mathbf{S}^3 \subset \mathbf{R}^4 : x_1^2 + x_2^2 = x_3^2 + x_4^2 = \frac{1}{2} \right\} \subset \mathbf{S}^3,$$

where $\mathbf{S}^3$ here is identified with the unit sphere of $\mathbf{R}^4$.

Of these, equatorial $\mathbf{S}^{n-1}$s are totally geodesic, and hyperplanes are both stable and totally geodesic.

Stable minimal surfaces play a key role in the Schoen-Yau approach to classifying closed three-dimensional manifolds with positive scalar curvature. An important step in this approach is to note the relationship of the second variation operator of the area functional with the ambient scalar curvature:

Proposition A.2.6 (cf. Schoen-Yau [59, Theorem 5.1]). *Let (M^n, g) be a Riemannian manifold, and Σ^{n-1} be a complete minimal hypersurface with unit normal vector field ν. Then, Σ^{n-1} is stable if and only if*

$$\int_\Sigma \left[\|\nabla^\Sigma f\|^2 - \frac{1}{2} \left(R_M - R_\Sigma + \| \mathrm{II}_\Sigma \|^2 \right) f^2 \right] d\sigma \geq 0 \text{ for all } f \in C_c^\infty(\Sigma). \tag{A.2.1}$$

As a corollary to Proposition A.2.6, Schoen-Yau proved the following stringent topological restriction on the classes of manifolds that can admit metrics of positive scalar curvature:

Proposition A.2.7 (Schoen-Yau [59, Theorem 5.1]). *If (M^3, g) is a Riemannian manifold with positive scalar curvature, and Σ^2 is a closed two-sided stable minimal surface in M^3, then Σ^2 is diffeomorphic to $\mathbf{S}^2$ or $\mathbf{RP}^2$.*

Proof. Plug in $f = 1$ to inequality (A.2.1) of Proposition A.2.6. We get

$$\frac{1}{2}\int_{\Sigma}\left[R_M - R_\Sigma + \| \mathrm{I\!I}_\Sigma \|^2\right]\, d\sigma \leq 0.$$

Clearly, $\| \mathrm{I\!I}_\Sigma \|^2 \geq 0$. If, in addition, $R_M > 0$, then we must have

$$\int_\Sigma R_\Sigma \, d\sigma > 0.$$

By the Gauss-Bonnet theorem, $\Sigma^2 \approx \mathbf{S}^2$ or $\mathbf{RP}^2$. □

One way that is used to produce stable minimal hypersurfaces is through direct area minimization. To this end, we will invoke the following standard lemma on a few occasions in this thesis.

Lemma A.2.8 (Federer [21, Theorem 5.4.15])**.** *Let (Ω^n, g), $3 \leq n \leq 7$, be a compact Riemannian manifold whose boundary has an outward pointing mean curvature vector, or is empty. For every $T \in H_{n-1}(\Omega, \mathbf{Z}) \setminus \{[\mathbf{0}]\}$, there exists a smoothly immersed hypersurface $\Sigma \subset \operatorname{int}\Omega$ such that $[\Sigma] - T = [\mathbf{0}] \in H_{n-1}(\Omega, \mathbf{Z})$ and $\operatorname{vol}(\Sigma)$ is minimum among all elements of $H_{n-1}(\Omega, \mathbf{Z})$.*

Appendix B

Elliptic Partial Differential Equations

We compile in this appendix some facts about elliptic partial differential equations that are used throughout this thesis. We refer the reader to [27] and [3] for elliptic theory on Euclidean space and manifolds, respectively.

We will consider so-called Schrödinger differential operators of the form

$$L \triangleq -\Delta_g + V \tag{B.0.1}$$

on Riemannian manifolds (Σ^n, g), where $V \in C^\infty_{\mathrm{loc}}(\Sigma^n) \cap L^\infty(\Sigma^n)$.

B.1 Maximum principles

The classical interior and boundary strong maximum principles will be used often in this thesis, so we state them here for convenience. For proofs, please consult [27, Section 3.2], [3, Chapter 3].

Theorem B.1.1 (Maximum principle, cf. Gilbarg-Trudinger [27, Section 3.2]). *Let (Σ^n, g) be complete, $\Omega \subseteq \operatorname{int} \Sigma$ a bounded, connected, open subset, $u \in C^2_{\mathrm{loc}}(\Omega) \cap C^0(\overline{\Omega})$. Suppose $Lu \leq 0$ in Ω, and that $u|_\Omega$ attains a maximum at $p \in \Omega$. If $u(p) = 0$, then u is constant in Ω.*

If $V \geq 0$ in Ω, we can replace the $u(p) = 0$ assumption by $u(p) \geq 0$.

If $V \equiv 0$ in Ω, we can lift the sign assumption on $u(p)$ altogether.

Lemma B.1.2 (Hopf boundary point lemma, cf. Gilbarg-Trudinger [27, Section 3.2]). *Let (Σ^n, g) be complete, $B \subset \Sigma^n$ be a bounded open ball, and suppose $u \in C^2_{\mathrm{loc}}(B)$. Suppose $Lu \leq 0$ and $u < u(p)$ in B, and that u is continuous at p. If $u(p) = 0$, then*

$$\limsup_{t \downarrow 0} \frac{u(p + t\eta) - u(p)}{t} < 0,$$

where η is the inward pointing normal at p.

If $V \geq 0$ in B, we can replace the $u(p) = 0$ assumption by $u(p) \geq 0$.

If $V \equiv 0$ in B, we can lift the sign assumption on $u(p)$ altogether.

B.2 Dirichlet eigenvalues on compact domains

In what follows, we assume that Ω is a bounded, connected, open subset of $\operatorname{int} \Sigma$ with Lipschitz boundary. The following result is classical and shows that the spectrum of a self-adjoint operator can be realized by an L^2-complete orthonormal set of functions.

Proposition B.2.1 (Existence of eigenfunctions, cf. Aubin [3, Section 4.1]). *There exists a sequence*

$$\lambda_1(L; \Omega) \leq \lambda_2(L; \Omega) \leq \ldots, \ \lim_i \lambda_i(L; \Omega) = \infty,$$

("eigenvalues") and a corresponding sequence of functions $\varphi_1, \varphi_2, \ldots \in C^\infty_{\mathrm{loc}}(\Omega) \cap W^{1,2}_0(\Omega)$ ("eigenfunctions") that form an $L^2(\Omega)$-complete orthonormal set, and satisfy $L\varphi_j = \lambda_j(L; \Omega)\varphi_j$ in Ω for all $j = 1, 2, \ldots$.

The following variational ("min-max") characterization of eigenvalues is important:

Proposition B.2.2 ([27, Theorem 8.37]). *The eigenvalues $\lambda_1(L; \Omega) \leq \lambda_2(L; \Omega) \leq \ldots$ are characterized variationally by:*

$$\lambda_k(L; \Omega) = \min \left\{ \max_{\zeta \in V \setminus \{\mathbf{0}\}} \mathcal{R}[\zeta] : V \subset W^{1,2}_0(\Omega) \text{ is a } k\text{-dimensional subspace} \right\}.$$

Here, $\mathcal{R} : W^{1,2}_0(\Omega) \setminus \{0\} \to \mathbf{R}$ is the Rayleigh quotient, given by

$$\mathcal{R}[\zeta] \triangleq \frac{\mathcal{Q}(\zeta, \zeta)}{\|\zeta\|^2_{L^2(\Omega)}},$$

where in turn, $\mathcal{Q} : W_0^{1,2}(\Omega) \otimes W_0^{1,2}(\Omega) \to \mathbf{R}$ *is the bilinear form*

$$\mathcal{Q}(\zeta, \psi) \triangleq \int_\Omega [\langle \nabla\zeta, \nabla\psi\rangle + V\zeta\psi] \, dv.$$

An immediate corollary of this min-max procedure is:

Corollary B.2.3 (Monotonicity of Dirichlet eigenvalues)**.** *If* $\Omega \subseteq \Omega'$ *are two open, bounded, Lipschitz domains, then* $\lambda_i(L; \Omega') \leq \lambda_i(L; \Omega)$*. Equality holds if and only if* $\Omega = \Omega'$*.*

The following result, which shows that the first eigenspace is one-dimensional, is considered standard. Nonetheless, we outline a proof here because it sets the stage for a more involved argument involved in the proof of Lemma 4.3.8.

Proposition B.2.4 ("$\lambda_1 < \lambda_2$", [27, Theorem 8.38])**.** *The first eigenspace*

$$E_1(L; \Omega) = \{f \in W_0^{1,2}(\Omega) : Lf = \lambda_1(L; \Omega) f \text{ weakly on } \Sigma\}$$

is a one-dimensional subspace of $C^\infty_{\text{loc}}(\Omega) \cap W_0^{1,2}(\Omega)$*, and if* $f \in E_1$*, then* $\{f = 0\} = \emptyset$ *or* Ω*.*

Proof. It is clear that $E_1(L; \Omega) \subset C^\infty_{\text{loc}}(\Omega) \cap W_0^{1,2}(\Omega)$, by elliptic regularity [3, Section 4.1]. Let's first prove the latter claim:

Claim. Let f, $f \not\equiv 0$, be an eigenfunction corresponding to the first eigenvalue, λ_1. Then f has no zeros in Ω.

Proof of claim. We have

$$\mathcal{Q}(|f|, |f|) = \int_\Omega \left[\|\nabla|f|\|^2 + V|f|^2\right] dv \leq \int_\Omega \left[\|\nabla f\|^2 + Vf^2\right] dv = \mathcal{Q}(f, f),$$

which implies, by Proposition B.2.2, that $|f| \in E_1(L; \Omega)$. However, $E_1(L; \Omega) \subset C^\infty_{\text{loc}}(\Omega)$, i.e., $|f| \in C^\infty_{\text{loc}}(\Omega)$. Moreover, $|f| \geq 0$. By the strong maximum principle, $|f|$ cannot attain an interior zero minimum, so f has no zeros. □

Next, we show $\dim E_1(L; \Omega) = 1$. If f, $f' \in E_1(L; \Omega) \setminus \{0\}$, then there exists $c \in \mathbf{R} \setminus \{0\}$ such that $f - cf'$ is still a first eigenfunction, but with a nonempty zero set. By the previous claim, $f = cf'$ in Ω. □

B.3 Dependence on parameters

A standard reference for the material in this section, including a treatment of tame Fréchet spaces, is [30].

Lemma B.3.1 (M., cf. M.-Schoen [40, Lemma A.1])**.** *Let Σ^n be a closed manifold, $\mathcal{D}^d$ a manifold, $g : \mathcal{D}^d \to \Gamma(\mathrm{Met}(\Sigma^n))$ a (Fréchet) smooth family of metrics, and $f : \mathcal{D}^d \to C^\infty(\Sigma)$ (Fréchet) smooth, too, with*

$$\int_\Sigma f(p)(\cdot)\, dv_{g(p)} = 0 \text{ for all } p \in \mathcal{D}.$$

If there exist $p_0 \in \mathcal{D}$ and $u_0 \in C^\infty(\Sigma) \setminus \{0\}$ with $\Delta_{g(p_0)} u_0 = f(p_0)$, then there also exists a (Fréchet) smooth $\sigma : \mathcal{D} \to C^\infty(\Sigma) \setminus \{0\}$ such that $\Delta_{g(p)}\sigma(p) = f(p)$ for all $p \in \mathcal{D}$.

Proof. This follows from the theory of [30, Section II.3]. Since $\mathcal{D}$ is finite dimensional, we know that the map $\mathcal{P} : \mathcal{D} \times C^\infty(\Sigma) \times \mathbf{R} \to C^\infty(\Sigma) \times \mathbf{R}$ given by

$$\mathcal{P}(p, u, t) \triangleq \left(\Delta_{g(p)} u + t, \int_\Sigma u\, dv_{g(p)} \right)$$

is a smooth tame Fréchet map, and that $\mathcal{P}(p, \cdot, \cdot)$ is invertible for all $p \in \mathcal{D}$. The solution operator $\mathcal{S} : \mathcal{D} \times C^\infty(\Sigma) \times \mathbf{R} \to C^\infty(\Sigma) \times \mathbf{R}$ is, therefore, also a smooth tame Fréchet map. Defining, e.g.,

$$\sigma(p) \triangleq \mathcal{S}\left(p, f(p), \int_\Sigma u_0\, dv_{g(p_0)} \right)$$

gives the required result. $\square$

Lemma B.3.2 (M., cf. M.-Schoen [40, Lemma A.1])**.** *Let Σ^n be a closed manifold, $\mathcal{D}^d$ a manifold, $g : \mathcal{D}^d \to \Gamma(\mathrm{Met}(\Sigma^n))$ a (Fréchet) smooth family of metrics, and $V : \mathcal{D} \to C^\infty(\Sigma)$ (Fréchet) smooth, too. If $\lambda(p)$ denotes the first eigenvalue of $u \mapsto L_p u \triangleq -\Delta_{g(p)} u + V(p) u$, $p \in \mathcal{D}$, then $\lambda : \mathcal{D} \to \mathbf{R}$ is smooth, and there exists a (Fréchet) smooth $\sigma : \mathcal{D} \to C^\infty(\Sigma) \setminus \{0\}$ so that $\sigma(p)(\cdot)$ is an eigenfunction of L_p with eigenvalue $\lambda(p)$, for all $p \in \mathcal{D}$.*

Proof. Let us first prove this for any sufficiently small neighborhood $U \subset \mathcal{D}$ that is diffeomorphic to $B_\varepsilon(\mathbf{0}) \subset \mathbf{R}^d$. For $\varepsilon > 0$, $\delta > 0$ small enough, $\sigma(L_p) \cap \{z \in \mathbf{C} : |z - \lambda(\mathbf{0})| < \delta\} = \{\lambda(p)\}$ for all $|p| < \varepsilon$; this is possible because $\sigma(L_p) \subset \mathbf{R}$ (all L_p are self-adjoint), the first eigenvalue is simple, and spectra depend continuously on the operators. For all such p, the following contour integral over $\gamma = \{z \in \mathbf{C} : |z - \lambda(\mathbf{0})| = \delta\}$ makes sense:

$$\Pi(p, u) = \frac{1}{2\pi i} \int_\gamma (\zeta I - L_{g(p)})^{-1} u\, d\zeta.$$

This is the spectral projection map onto $\ker(L_p - \lambda(p)I)$, for all $|p| < \varepsilon$. By elliptic theory [30, Section II.3], $\Pi : B_\varepsilon(\mathbf{0}) \times (C^\infty(\Sigma^n) \otimes_{\mathbf{R}} \mathbf{C}) \to (C^\infty(\Sigma^n) \otimes_{\mathbf{R}} \mathbf{C})$ is a (Fréchet) smooth map.

If $u_{\mathbf{0}} \in \ker(L_{\mathbf{0}} - \lambda(\mathbf{0})I)$ and $u_{\mathbf{0}} \not\equiv 0$, then for a possibly smaller $\varepsilon' < \varepsilon$ (i.e., small enough so that $|p| < \varepsilon' \implies \Pi(p, u_{\mathbf{0}}) \not\equiv 0$), the maps $\Pi(p, u_{\mathbf{0}})$ are all nontrivial $\mathbf{C}$-valued first eigenfunctions of L_p. Recall that $\dim_{\mathbf{C}}(\ker(L_p - \lambda(p)I) \otimes_{\mathbf{R}} \mathbf{C}) = 1$. Therefore, we can rescale (in the field $\mathbf{C}$) at each p, in a smooth fashion, to also get a real-valued (Fréchet) smooth map $\sigma : B_{\varepsilon'}(\mathbf{0}) \to C^\infty(\Sigma) \setminus \{0\}$.

Finally, to prove this for the general case of $\mathcal{D} \subset \mathbf{R}^d$, notice that by virtue of $\dim_{\mathbf{R}} \ker(L_p - \lambda(p)I) = 1$, we can fix a canonical choice of $\sigma(p)$, e.g., by requiring that

$$\int_\Sigma \sigma(p)(\cdot)\, dv_{g(p)} = 1,$$

and the result will follow by patching together all local solutions. $\square$

B.4 Ekeland variational principle

We mention here a theorem of Ekeland's and we include a proof because this result is not found in the standard textbooks.

Theorem B.4.1 (Ekeland variational principle, Ekeland [18, Theorem 1.1])**.** *Let $F : M \to \mathbf{R} \cup \{\infty\}$ be a lower semicontinuous function, which is bounded from below on a complete metric space (M, d). Given ε, $\delta > 0$, and $p \in M$ such that $F(p) \leq \inf_M F + \varepsilon$, there exists $p^* \in M$ such that:*

$$F(p^*) \leq F(p),\ d(p^*, p) \leq \delta,\ \text{and } F(q) > F(p^*) - \varepsilon\delta^{-1} d(q, p^*) \text{ for all } q \in M.$$

Proof. Let $\alpha = \varepsilon\delta^{-1}$, and for each $q \in M$ consider the set

$$C_q \triangleq \{q' \in M : F(q') < F(q) - \alpha d(q, q')\}.$$

Notice, then, our desired conclusion is equivalent to

$$\overline{B}_\delta(p) \cap \{q \in M : F \leq F(p) \text{ and } C_q = \emptyset\} \neq \emptyset.$$

Consider a sequence of points defined inductively as follows: $p_0 \triangleq p$, and

$$p_{i+1} \triangleq \text{ any element of } \left\{ q \in C_{p_i} \text{ with } F(q) - \inf_{C_{p_i}} F \leq \frac{1}{2}\left(F(p_i) - \inf_{C_{p_i}} F\right)\right\}$$

if the set is nonempty; the induction is terminated if this set ever becomes empty. Notice that, by definition, the sequence of sets $\{C_{p_i}\}_{i=0,1,\ldots}$ is shrinking and, in particular, if the induction terminates at p_i, then $p^* \triangleq p_i$ fulfills the desired conclusion.

Suppose, thus, that the induction does not terminate. For each i we have

$$q \in C_{p_i} \implies d(q, p_i) < \alpha^{-1}(F(p_i) - F(q)) \implies \operatorname{diam} C_{p_i} \leq 2\alpha^{-1}\left(F(p_i) - \inf_{C_{p_i}} F\right). \tag{B.4.1}$$

Moreover, by our choice of p_{i+1} and since $C_{p_{i+1}} \subset C_{p_i}$, we also have

$$F(p_{i+1}) - \inf_{C_{p+1}} F \leq F(p_{i+1}) - \inf_{C_{p_i}} F \leq \frac{1}{2}\left(F(p_i) - \inf_{C_{p_i}} F\right),$$

which in view of $F(p_0) \leq \inf_X F + \varepsilon$, gives

$$F(p_i) - \inf_{C_{p_i}} F \leq \varepsilon 2^{-i}.$$

Plugging this into (B.4.1), we find that $\operatorname{diam} C_{p_i} \leq \varepsilon\alpha^{-1}2^{1-i}$ for all i; thus, $\{p_i\}_{i=0,1,\ldots}$ is Cauchy and the result follows by the completeness of (M, d). □

Corollary B.4.2 (Ekeland [18, Theorem 2.2])**.** *Let $F : X \to \mathbf{R} \cup \{\infty\}$ be a lower semicontinuous function that is bounded from below on a Banach space X, and which is Gâteaux-differentiable at all finite-valued points. Given $\varepsilon > 0$, $\delta > 0$, and a $u \in X$ such that $F(u) \leq \inf_X F + \varepsilon$, there exists $u^* \in X$ such that $F(u^*) \leq F(u)$, $\|u - u^*\| \leq \delta$, and $\|\delta F(u^*)\| \leq \varepsilon\delta^{-1}$.*

Proof. Apply Theorem B.4.1, viewing X as a metric space. We get a $u^* \in X$, and for all $v \in X$

$$\delta F(u^*)\{v\} = \lim_{t\to 0} \frac{F(u^* + tv) - F(u^*)}{t} \geq \lim_{t\downarrow 0} \frac{-\varepsilon\delta^{-1}\|tv\|}{t} = -\varepsilon\delta^{-1},$$

as well as $\leq \varepsilon\delta^{-1}$ by instead using $t \uparrow 0$ in the last limit. The result follows. □

Bibliography

[1] L. Ambrosio and X. Cabré. Entire solutions of semilinear elliptic equations in $\mathbf{R}^3$ and a conjecture of De Giorgi. *J. Amer. Math. Soc.*, 13(4):725–739, 2000.

[2] R. Arnowitt, S. Deser, and C. W. Misner. Dynamical structure and definition of energy in general relativity. *Phys. Rev. (2)*, 116:1322–1330, 1959.

[3] T. Aubin. *Some nonlinear problems in Riemannian geometry.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[4] R. Bartnik. The mass of an asymptotically flat manifold. *Comm. Pure Appl. Math.*, 39(5):661–693, 1986.

[5] R. Bartnik. New definition of quasilocal mass. *Phys. Rev. Lett.*, 62(20):2346–2348, 1989.

[6] A. N. Bernal and M. Sánchez. Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes. *Comm. Math. Phys.*, 257(1):43–50, 2005.

[7] L. Bessières, G. Besson, and S. Maillot. Ricci flow on open 3-manifolds and positive scalar curvature. *Geom. Topol.*, 15(2):927–975, 2011.

[8] G. D. Birkhoff. Dynamical systems with two degrees of freedom. *Trans. Amer. Math. Soc.*, 18(2):199–300, 1917.

[9] H. L. Bray. Proof of the Riemannian Penrose inequality using the positive mass theorem. *J. Differential Geom.*, 59(2):177–267, 2001.

[10] J. D. Brown and J. W. York, Jr. Quasilocal energy and conserved charges derived from the gravitational action. *Phys. Rev. D (3)*, 47(4):1407–1419, 1993.

[11] X. Cabré. Uniqueness and stability of saddle-shaped solutions to the Allen-Cahn equation. *J. Math. Pures Appl. (9)*, 98(3):239–256, 2012.

[12] T. H. Colding and W. P. Minicozzi, II. *A course in minimal surfaces*, volume 121 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2011.

[13] C. B. Croke. Area and the length of the shortest closed geodesic. *J. Differential Geom.*, 27(1):1–21, 1988.

[14] M. Cveti ˇ c, G. W. Gibbons, and C. N. Pope. More about Birkhoff's invariant and Thorne's hoop conjecture for horizons. *Classical Quantum Gravity*, 28(19):195001, 22, 2011.

[15] E. De Giorgi. Convergence problems for functionals and operators. In *Proceedings of the International Meeting on Recent Methods in Nonlinear Analysis (Rome, 1978)*, pages 131–188. Pitagora, Bologna, 1979.

[16] M. del Pino, M. Kowalczyk, and F. Pacard. Moduli space theory for the Allen-Cahn equation in the plane. *Trans. Amer. Math. Soc.*, 365(2):721–766, 2013.

[17] M. del Pino, M. Kowalczyk, and J. Wei. On De Giorgi's conjecture in dimension $N \geq 9$. *Ann. of Math. (2)*, 174(3):1485–1569, 2011.

[18] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.

[19] X.-Q. Fan, Y. Shi, and L.-F. Tam. Large-sphere and small-sphere limits of the Brown-York mass. *Comm. Anal. Geom.*, 17(1):37–72, 2009.

[20] A. Farina, L. Mari, and E. Valdinoci. Splitting theorems, symmetry results and overdetermined problems for Riemannian manifolds. *Comm. Partial Differential Equations*, 38(10):1818–1862, 2013.

[21] H. Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[22] D. Fischer-Colbrie. On complete minimal surfaces with finite Morse index in three-manifolds. *Invent. Math.*, 82(1):121–132, 1985.

[23] Pengfei G. and Junfang L. The quermassintegral inequalities for k-convex starshaped domains. *Advances in Mathematics*, 221(5):1725 – 1732, 2009.

[24] N. Garofalo and F.-H. Lin. Monotonicity properties of variational integrals, A_p weights and unique continuation. *Indiana Univ. Math. J.*, 35(2):245–268, 1986.

[25] N. Ghoussoub and C. Gui. On a conjecture of De Giorgi and some related problems. *Math. Ann.*, 311(3):481–491, 1998.

[26] G. W. Gibbons. Birkhoff's invariant and Thorne's Hoop Conjecture. arXiv:0903.1580, 2009.

[27] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[28] M. Gromov and H. B. Lawson, Jr. The classification of simply connected manifolds of positive scalar curvature. *Ann. of Math. (2)*, 111(3):423–434, 1980.

[29] C. Gui. Symmetry of some entire solutions to the Allen-Cahn equation in two dimensions. *J. Differential Equations*, 252(11):5853–5874, 2012.

[30] R. S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)*, 7(1):65–222, 1982.

[31] S. W. Hawking. Black holes in general relativity. *Comm. Math. Phys.*, 25(2):152–166, 1972.

[32] G. Huisken and T. Ilmanen. The inverse mean curvature flow and the Riemannian Penrose inequality. *J. Differential Geom.*, 59(3):353–437, 2001.

[33] R. C. Kirby. *The topology of* 4*-manifolds*, volume 1374 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1989.

[34] M. Kowalczyk and Y. Liu. Nondegeneracy of the saddle solution of the Allen-Cahn equation. *Proc. Amer. Math. Soc.*, 139(12):4319–4329, 2011.

[35] M. Kowalczyk, Y. Liu, and F. Pacard. The space of 4-ended solutions to the Allen-Cahn equation in the plane. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 29(5):761–781, 2012.

[36] Y. Liu, K. Wang, and J. Wei. Global minimizers of the Allen-Cahn equation in dimension $n \geq 8$. arXiv:1606.05315, June 2016.

[37] S. Lu. On Weyl's embedding problem in Riemannian manifolds. arXiv:1608.07539, August 2016.

[38] C. Mantoulidis and P. Miao. Mean curvature deficit and a quasi-local mass. CMSA Lecture Series (to appear), 2016.

[39] C. Mantoulidis and P. Miao. Total mean curvature, scalar curvature, and a variational analog of Brown–York mass. *Communications in Mathematical Physics*, 352(2):703–718, 2017.

[40] C. Mantoulidis and R. Schoen. On the Bartnik mass of apparent horizons. *Classical and Quantum Gravity*, 32(20):205002, 2015.

[41] W. Meeks, III, L. Simon, and S.-T. Yau. Embedded minimal surfaces, exotic spheres, and manifolds with positive Ricci curvature. *Ann. of Math. (2)*, 116(3):621–659, 1982.

[42] P. Miao. Positive mass theorem on manifolds admitting corners along a hypersurface. *Adv. Theor. Math. Phys.*, 6(6):1163–1182 (2003), 2002.

[43] P. Miao. On a localized Riemannian Penrose inequality. *Comm. Math. Phys.*, 292(1):271–284, 2009.

[44] P. Miao, Y. Shi, and L.-F. Tam. On geometric problems related to Brown-York and Liu-Yau quasilocal mass. *Comm. Math. Phys.*, 298(2):437–459, 2010.

[45] L. Modica. A gradient bound and a Liouville theorem for nonlinear Poisson equations. *Comm. Pure Appl. Math.*, 38(5):679–684, 1985.

[46] S. Montiel and A. Ros. Schrödinger operators associated to a holomorphic map. In *Global differential geometry and global analysis (Berlin, 1990)*, volume 1481 of *Lecture Notes in Math.*, pages 147–174. Springer, Berlin, 1991.

[47] L. Nirenberg. The Weyl and Minkowski problems in differential geometry in the large. *Comm. Pure Appl. Math.*, 6:337–394, 1953.

[48] B. O'Neill. *Semi-Riemannian geometry*, volume 103 of *Pure and Applied Mathematics*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1983. With applications to relativity.

[49] F. Pacard and J. Wei. Stable solutions of the Allen-Cahn equation in dimension 8 and minimal cones. *J. Funct. Anal.*, 264(5):1131–1167, 2013.

[50] T. Parker and C. H. Taubes. On Witten's proof of the positive energy theorem. *Comm. Math. Phys.*, 84(2):223–238, 1982.

[51] R. Penrose. Naked Singularities. *Annals of the New York Academy of Sciences*, 224(1):125–134, 1973.

[52] G. Perelman. The entropy formula for the Ricci flow and its geometric applications. arXiv:math/0211159, November 2002.

[53] G. Perelman. Finite extinction time for the solutions to the Ricci flow on certain three-manifolds. arXiv:math/0307245, July 2003.

[54] G. Perelman. Ricci flow with surgery on three-manifolds. arXiv:math/0303109, March 2003.

[55] P. Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics*. Springer, Cham, third edition, 2016.

[56] A. V. Pogorelov. Some results on surface theory in the large. *Advances in Math.*, 1(fasc. 2):191–264, 1964.

[57] A. D. Rendall. The initial value problem for a class of general relativistic fluid bodies. *J. Math. Phys.*, 33(3):1047–1053, 1992.

[58] O. Savin. Regularity of flat level sets in phase transitions. *Ann. of Math. (2)*, 169(1):41–78, 2009.

[59] R. Schoen and S.-T. Yau. Existence of incompressible minimal surfaces and the topology of three-dimensional manifolds with nonnegative scalar curvature. *Ann. of Math. (2)*, 110(1):127–142, 1979.

[60] R. Schoen and S.-T. Yau. On the proof of the positive mass conjecture in general relativity. *Comm. Math. Phys.*, 65(1):45–76, 1979.

[61] R. Schoen and S.-T. Yau. On the structure of manifolds with positive scalar curvature. *Manuscripta Math.*, 28(1-3):159–183, 1979.

[62] R. Schoen and S.-T. Yau. Proof of the positive mass theorem. II. *Comm. Math. Phys.*, 79(2):231–260, 1981.

[63] R. Schoen and S.-T. Yau. *Lectures on differential geometry*. Conference Proceedings and Lecture Notes in Geometry and Topology, I. International Press, Cambridge, MA, 1994. Lecture notes prepared by Wei Yue Ding, Kung Ching Chang [Gong Qing Zhang], Jia Qing Zhong and Yi Chao Xu, Translated from the Chinese by Ding and S. Y. Cheng, Preface translated from the Chinese by Kaising Tso.

[64] Y. Shi and L.-F. Tam. Positive mass theorem and the boundary behaviors of compact manifolds with nonnegative scalar curvature. *J. Differential Geom.*, 62(1):79–125, 2002.

[65] Y. Shi and L.-F. Tam. Quasi-local mass and the existence of horizons. *Comm. Math. Phys.*, 274(2):277–295, 2007.

[66] Y. Shi and L.-F. Tam. Rigidity of compact manifolds and positivity of quasi-local mass. *Classical Quantum Gravity*, 24(9):2357–2366, 2007.

[67] B. Smith. Black hole initial data with a horizon of prescribed intrinsic and extrinsic geometry. In *Complex analysis and dynamical systems IV. Part 2*, volume 554 of *Contemp. Math.*, pages 229–244. Amer. Math. Soc., Providence, RI, 2011.

[68] R. M. Wald. *General relativity*. University of Chicago Press, Chicago, IL, 1984.

[69] K. Wang. A new proof of Savin's theorem on Allen-Cahn equations. arXiv:1401.6480, January 2014.

[70] K. Wang. Some remarks on the structure of finite Morse index solutions to the Allen-Cahn equation in $\mathbb{R}^2$. arXiv:1506.00499, June 2015.

[71] K Wang and J. Wei. Finite Morse index implies finite ends. arXiv:1705.06831, May 2017.

[72] M.-T. Wang and S.-T. Yau. A generalization of Liu-Yau's quasi-local mass. *Comm. Anal. Geom.*, 15(2):249–282, 2007.

[73] E. Witten. A new proof of the positive energy theorem. *Comm. Math. Phys.*, 80(3):381–402, 1981.