Renormalized area and properly embedded minimal surfaces in hyperbolic 3-manifolds

Spyridon Alexakis [*]
Princeton University

Rafe Mazzeo [†]
Stanford University

Abstract

We study the renormalized area functional $\mathcal{A}$ in the AdS/CFT correspondence, specifically for properly embedded minimal surfaces in convex cocompact hyperbolic 3-manifolds (or somewhat more broadly, Poincaré-Einstein spaces). Our main results include an explicit formula for the renormalized area of such a minimal surface Y as an integral of local geometric quantities, as well as formulæ for the first and second variations of $\mathcal{A}$ which are given by integrals of *global* quantities over the asymptotic boundary loop γ of Y. All of these formulæ are also obtained for a broader class of nonminimal surfaces. The proper setting for the study of this functional (when the ambient space is hyperbolic) requires an understanding of the moduli space of all properly embedded minimal surfaces with smoothly embedded asymptotic boundary. We show that this moduli space is a smooth Banach manifold and develop a $\mathbb{Z}$-valued degree theory for the natural map taking a minimal surface to its boundary curve. We characterize the nondegenerate critical points of $\mathcal{A}$ for minimal surfaces in $\mathbb{H}^3$, and finally, discuss the relationship of $\mathcal{A}$ to the Willmore functional.

1 Introduction

There is an interesting nonlinear asymptotic boundary problem in which one seeks a minimal submanifold in hyperbolic space with prescribed asymptotic boundary a submanifold in the sphere at infinity. This was treated conclusively by Anderson [2], [3] in the early 1980's using techniques from geometric measure theory; the solutions he obtains are absolutely volume minimizing with respect to compact variations. One may also pose this problem when the ambient space is a convex cocompact hyperbolic manifold, or even more generally a conformally compact manifold (M^{n+1}, g) (all definitions are reviewed in §2), and it is not hard to extend the existence theory to these settings. Here, however, we focus mostly on the special case of properly embedded minimal surfaces Y^2 in $M = \mathbb{H}^3/\Gamma$ where Γ is a convex cocompact subgroup (a particular case is $M = \mathbb{H}^3$ itself), with boundary curve $\partial Y = \gamma$ an embedded closed curve $\gamma \subset \partial M$. Beyond Anderson's aforementioned work, in this particular setting there is also a rich existence theory of minimal (not necessarily minimizing) surfaces of arbitrary genus by de Oliveira and Soret [30], see also Coskunzer [12], [13].

It turns out that there is a well-defined Hadamard regularization of the area of such a minimal surface, and this renormalized area is our central concern. Roughly speaking,

---

[*]Email: alexakis@math.princeton.edu

[†]Email: mazzeo@math.stanford.edu. Supported by the NSF under Grant DMS-0505709

our goal is to obtain a local formula for this renormalized area, i.e. one involving integrals of local geometric quantities, and then to use this to study the variational theory of the renormalized area functional $\mathcal{A}$. In order to do this properly, we must study the moduli space of all properly embedded minimal surfaces with embedded asymptotic boundary, as this is the natural domain of $\mathcal{A}$. This leads to a subsidiary investigation of the structure of these moduli spaces and the degree theory for the natural map taking a minimal surface to its boundary curve. We also consider the renormalized area of a larger class of nonminimal surfaces. We calculate the first and second variations of $\mathcal{A}$; interestingly, these are expressed as integrals of global quantities over the boundary curve. While we do not touch on all aspects of this variational theory, we are able to characterize the nondegenerate critical points amongst surfaces in hyperbolic space, and give some estimates for the numerical range of $\mathcal{A}$. Finally, we show the relationship of $\mathcal{A}$ to the much-studied Willmore functional $\mathcal{W}$, which suggests that $\mathcal{A}$ is the correct conformally natural generalization of $\mathcal{W}$ to surfaces with boundary.

There are strong motivations from the AdS/CFT correspondence in string theory for studying the renormalized area, and we shall explain some of these below, after describing the mathematical context more carefully.

Our results about minimal surfaces parallel a number of known results about Poincaré-Einstein (PE) spaces, so we describe these together. Let (M, g) be a PE space; this means that M is a manifold with boundary, $g = \rho^{-2}\overline{g}$ where ρ is a boundary defining function for ∂M and $\overline{g}$ is smooth and nondegenerate up to the boundary, and g is Einstein. There is a well-defined conformal class $\mathfrak{c}(g)$ on ∂M, called the conformal infinity of g, which should be regarded as the asymptotic boundary value of g. The space of all PE metrics (with some fixed regularity) on the interior of a given manifold with boundary M is a Banach manifold, and the conformal infinity map from this to the space of conformal structures on ∂M (which also has the structure of a Banach manifold) is Fredholm of degree 0. These facts were proved by Anderson [5], see also Biquard [8] and Lee [25]. Most existence results for PE metrics are perturbative in nature, but Anderson established a scheme to obtain a much broader existence theory when $\dim M = 4$ using degree theory [6]. One key ingredient is the properness of this conformal infinity boundary value map over the preimage of scalar positive conformal classes on ∂M. There are substantial technicalities in making all of this work; recent work of Chang and Yang [10] clarifies some of this.

We first prove an analogous result for properly embedded minimal submanifolds:

Theorem 1.1. *Suppose that $M = \mathbb{H}^3/\Gamma$ is a convex cocompact hyperbolic manifold, and let $\mathcal{M}_k(M)$ be the space of properly embedded minimal surfaces in M of genus k with asymptotic boundary a $\mathcal{C}^{3,\alpha}$ embedded closed (but possibly disconnected) curve in ∂M. Let $\mathcal{E}$ denote the space of all $\mathcal{C}^{3,\alpha}$ closed embedded curves in ∂M. Then both $\mathcal{M}_k(M)$ and $\mathcal{E}$ are Banach manifolds, and the natural map*

$$\Pi : \mathcal{M}_k(M) \longrightarrow \mathcal{E}$$

is a smooth proper Fredholm map of index 0.

These properties of Π imply the existence of a $\mathbb{Z}$-valued degree for it, which yields many refinements of the existence theory for these minimal surfaces. Some consequences will be described in §4.

The proof of most of this uses various well-known tools, hence this can be regarded as a good toy model for the corresponding result about four-dimensional Poincaré-Einstein

spaces. Note that the use of degree theory for the boundary map of minimal surfaces goes back to work of Tromba [38] in the 1970's and White [42] in the 1980's, and indeed those papers provided some of the inspiration for Anderson's proposal to use degree theory in the Einstein setting. A special case of this degree theory, for genus zero surfaces, was developed in [12].

Now return to the PE setting. Assuming the conformal infinity of the PE metric g is sufficiently regular, then g itself has an expansion up to some order at the boundary. When $\dim M = 4$, this has the form

$$g = \frac{dx^2 + h(x)}{x^2}, \qquad h(x) \sim h_0 + x^2 h_2 + x^3 h_3 + \dots; \tag{1.1}$$

here each h_j is a symmetric 2-tensor on ∂M; in particular, h_0 is a metric representing $\mathfrak{c}(g)$ and h_3 is trace- and divergence-free with respect to h_0. All other h_j are determined in terms of these two tensors. Furthermore, x is a special boundary defining function naturally associated to the choice of h_0. The volume form dV_g has a corresponding expansion

$$dV_g \sim \frac{A_0}{x^4} + \frac{A_2}{x^2} + A_4 + \dots;$$

the x^{-3} and x^{-1} terms are absent due to the absence of the h_1 term and the vanishing trace of h_3. The volume of $\{x \geq \epsilon\}$ is obviously finite for each $\epsilon > 0$ and has an expansion as $\epsilon \searrow 0$ of the form

$$\mathrm{Vol}\left(\{x \geq \epsilon\}\right) \sim \frac{\alpha_0}{\epsilon^3} + \frac{\alpha_1}{\epsilon} + \mathcal{V}(M, g) + \dots.$$

The constant term in this expansion is by definition the renormalized volume of (M, g). The key fact, first proved by the physicists Henningson and Skenderis [22], cf. [18] for a careful mathematical treatment, is that this is well-defined independently of the choice of metric $h_0 \in \mathfrak{c}(g)$. The definition of renormalized volume extends to arbitrary dimensions, and they show that it is well defined when $\dim M$ is even; when $\dim M$ is odd, however, it is not well-defined and has a simple transformation law under change of representative h_0. For simplicity here we focus on the four-dimensional case. Using the Einstein condition in the Gauss-Bonnet formula, Anderson [5] noted that

$$\mathcal{V}(M, g) = \frac{4\pi^2}{3} \chi(M) - \frac{1}{6} \int_M |W|^2 \, dV_g; \tag{1.2}$$

here W is the Weyl tensor, and the integral is convergent since $|W|^2$ is pointwise conformally invariant of weight -4. Anderson also computed a formula giving the infinitesimal variation of the renormalized volume in the direction of an infinitesimal Einstein deformation κ:

$$D\mathcal{V}|_g(\kappa) = -\frac{1}{4} \int_{\partial M} \langle \kappa_0, h_n \rangle \, dV_{h_0}, \tag{1.3}$$

in terms of the leading term in the expansion $\kappa \sim \kappa_0 + x\kappa_1 + \dots$. A much easier derivation of this formula is given in [1]. Again, this extends immediately to all even dimensions. It follows from this that (when n is even), $\mathbb{H}^n$, and indeed any convex cocompact hyperbolic quotient $\mathbb{H}^n / \Gamma$, is a critical point of $\mathcal{V}$. The variational problem for renormalized volume remains unstudied. When $\dim M = 4$, $\mathcal{V}$ is closely related to the σ_2 functional of the underlying incomplete metric on M, and there are some interesting rigidity results using it, see [9]. There are also several nicely geometric results about renormalized volume in 3

dimensions [40], [24] (recall that it depends on some choices here, so one has not simply a number but rather a functional on a given conformal class of the boundary surface).

Shortly after [22], and motivated by the same string-theoretic concerns, Graham and Witten [20] proved the existence of a well-defined renormalized area $\mathcal{A}$ for properly embedded minimal submanifolds Y in a PE space where the boundary of Y is also embedded in ∂M. Two dimensions is critical for minimal surfaces in roughly the same way that four dimensions is critical for Einstein metrics, so it is reasonable that the results above about renormalized volume of four-dimensional PE metrics have analogues for properly embedded minimal surfaces, and this is indeed true. Our second main result is an explicit formula for $\mathcal{A}$ and its first and second variations:

Theorem 1.2. *Let $Y \in \mathcal{M}(M)$ have a $\mathcal{C}^{3,\alpha}$ embedded boundary curve γ. Then*

$$\mathcal{A}(Y) = -2\pi\chi(Y) - \frac{1}{2}\int_Y |\widehat{k}|^2 \, dA,$$

where $\widehat{k}$ is the trace-free second fundamental form of Y; the integral is convergent since $|\widehat{k}|^2 \, dA$ is invariant under conformal changes of the ambient metric. Furthermore, if 0 is not in the spectrum of the Jacobi operator L_Y (in which case we say that Y is non-degenerate), and $\dot\phi$ is a Jacobi field on Y (i.e. $L_Y\dot\phi = 0$), which thus corresponds to a one-parameter family of minimal surfaces around Y, then (relative to a normalization which will be explained later), $\dot\phi \sim \dot\phi_0 + x\dot\phi_1 + \ldots$ and

$$D\mathcal{A}|_Y(\phi) = -3\int_\gamma (\dot\phi)_0 \, u_3 \, ds;$$

here u_3 is the coefficient of x^3 in the expansion for the function u which gives a graph parametrization of Y over the vertical cylinder $\gamma \times [0, x_0)$. Furthermore,

$$D^2\mathcal{A}|_Y(\phi, \phi) = -\frac{1}{2}\int_\gamma \dot\phi_0 \, \dot\phi_3 \, ds;$$

as we explain later, this shows that the Hessian of $\mathcal{A}$ is represented by the Dirichlet-to-Neumann operator for the Jacobi operator L_Y. Finally, if $M = \mathbb{H}^3$, the unique non-degenerate critical points of $\mathcal{A}$ are the totally geodesic copies of $\mathbb{H}^2$ (so γ is a round circle).

We now turn to the physical precursors of all of this. Maldacena's pioneering work [27] proposes that in the large t'Hooft coupling regime, the expectation value of the Wilson loop operator corresponding to some closed loop $\gamma \subset \partial M$ should be given by the area of the minimal surface $Y \subset M^{n+1}$ with asymptotic boundary γ. The papers [27] and [14] already point out that one must introduce an area renormalization, which motivated [20].

Quite recently, it has also been suggested ([34], [35]) that this renormalized area be used to measure the entanglement entropy of a particular region in the CFT. More specifically, [34] (see also [23]) proposes an 'area law': for the model $(\mathbb{H}^{n+1}, \mathbb{S}^n)$, the information of a domain $\Omega \subset \mathbb{S}^n$ should correspond to the region in $\mathbb{H}^{n+1}$ enclosed by a minimal submanifold with asymptotic boundary $\partial\Omega$ (which need not be well-defined, of course, since the minimal submanifold is not unique), and in particular, the entanglement entropy of a domain $\Omega \subset \mathbb{S}^2$ should correspond to the renormalized area of the minimal surface $Y \subset \mathbb{H}^3$ with boundary $\gamma = \partial\Omega$ (see formula (1.5) in [34]). This assertion is checked in the lowest dimensional case $n = 1$ in [35], [34], and special examples are also

presented in [23] for $n = 2$ – but the validity of the assertion in higher dimensions is disputed in [36].

Motivated by these proposals, substantial effort has been devoted in several recent physics papers to understanding the geometric features and renormalized area of various simple cases of minimal surfaces in $\mathbb{H}^3$. For example, in [14] the authors compute the renormalized area of totally geodesic planes; in [23] Hirata and Takayanagi study the existence of minimal surfaces with two disconected circles as asymptotic boundaries and also estimate the renormalized area of those surfaces; Maldacena [27] studies the case of a rectangle where the length T of one side approaches infinity. Furthermore, Drukker-Gross-Ooguri, [14] and Polyakov-Rychkov, [33] have sought to check the proposed formula in [27] relating the expectation value of the Wilson loop in CFT with the renormalized area of a minimal surface in AdS. This verification involves calculating the first and second variations of the renormalized area functional with respect to deformations of the loop γ. Since those authors did not have a usable explicit formula for the renormalized area, their calculations required justification for dropping certain divergent terms; in contrast, our local formulæ allow for straightforward calculations.

Since there seems to be active and continuing interest in these proposals relating renormalized area with the expectation values of Wilson loop operators, the loop equation and to entanglement entropy, we hope that our results will facilitate further investigations in this area.

Our paper is structured as follows: in section §2 we present some background material needed for this work, on Poincaré-Einstein metrics, uniformly degenerate elliptic operators and embedded minimal surfaces in convex co-compact hyperbolic 3-manifolds. The local formula for $\mathcal{A}$ is proved in section §3, and certain global aspects of this functional are studied in §5. The intervening §4 develops the moduli space theory of properly embedded minimal surfaces. This provides the correct setting in which to derive the first and second variation formulæ, which appears in §6. §7 characterizes the nondegenerate critical points of $\mathcal{A}$ when the ambient space is hyperbolic 3-space, $\mathbb{H}^3$. Finally, in §8 we discuss the relationship of $\mathcal{A}$ and the Willmore functional.

2   Geometric and analytic preliminaries

We now give precise definitions of the spaces and submanifolds we shall be working with and explain some of their properties. We also discuss some basic results about elliptic operators on these spaces.

2.1   Conformally compact and Poincaré-Einstein spaces and convex co-compact hyperbolic 3-manifolds

A Riemannian manifold (M, g) is called conformally compact if M is the interior of a smooth compact manifold with boundary and $g = \rho^{-2}\overline{g}$ where ρ is a defining function for ∂M and $\overline{g}$ is a metric smooth and nondegenerate up to ∂M. Any such metric is complete and has sectional curvatures tending to $-|d\rho|^2_{\overline{g}}(q)$ upon approach to any point $q \in \partial M$. In particular, if $|d\rho|^2_{\overline{g}}$ is constant along ∂M, we say that (M, g) is asymptotically hyperbolic (AH). To any conformally compact metric g one may associate a conformal

class on ∂M:

$$\mathfrak{c}(g) = \left[\rho^2 g \big|_{T\partial M} \right],$$

which is obviously independent of the choice of defining function ρ. This conformal equivalence class is called the conformal infinity of g.

Any AH metric has a normal form, due to Graham and Lee [19]. Let (M, g) be an AH space and fix any metric h_0 representing the conformal class $\mathfrak{c}(g)$. Then there is a unique defining function x for ∂M, defined in some neighborhood $\mathcal{U}$ of the boundary, which satisfies the two conditions

$$|d \log x|_g^2 \equiv 1, \qquad \overline{g}\big|_{T\partial M} = h_0, \quad \text{where} \ \ \overline{g} = x^2 g.$$

The flow lines for the gradient $\nabla^{\overline{g}} x$ give a product decomposition $\mathcal{U} \cong [0, x_0) \times \partial M$, in terms of which the pullback of the metric g takes the form

$$g = \frac{dx^2 + h(x)}{x^2}, \qquad h(x) \sim h_0 + x h_1 + x^2 h_2 + \dots \tag{2.4}$$

The defining function x associated to the boundary metric h_0 will be called a special boundary defining function (bdf).

A case of particular special interest is when (M, g) is Poincaré-Einstein (PE), which means simply that it is both conformally compact and Einstein. These metrics were introduced by Fefferman and Graham [16] as a way of canonically associating a Riemannian metric on an ambient $(n + 1)$-manifold to a conformal class on an n-manifold, with the goal of finding new conformal invariants on the boundary via Riemannian invariants of the ambient manifold. If the conformal infinity of such a g is smooth, then the family of tensors $h(x)$ in (2.4) has a complete expansion in powers of x (and also powers of $x^{n-1} \log x$ when $n = \dim X$ is odd, $n \geq 5$). The coefficients h_0 and h_{n-1} are formally undetermined, but all other h_j can be expressed as local differential operators applied to these two coefficients; it is thus natural to think of the pair (h_0, h_{n-1}) as the Cauchy data of g.

In this paper we shall be primarily concerned with the three-dimensional case. If (M^3, g) is PE, then M is isometric to a convex cocompact quotient $\mathbb{H}^3/\Gamma$. (Convex cocompact means that Γ is geometrically finite and has no parabolic elements; equivalently, the quotient by Γ of the convex hull (in $\mathbb{H}^3$) of the limit set $\Lambda(\Gamma)$ is compact in M.) The Fefferman-Graham expansion for g simplifies then, and has a special form where only h_0, h_2 and h_4 are nonzero, see [15] and Epstein's appendix in [31]. These coefficients can be calculated in terms of the metric and second fundamental form of any one of the level sets $\{x = \text{const.}\}$, and the special bdf x then has the property that $-\log x$ is the distance function to this level set (up to an additive constant).

2.2 Uniformly degenerate operators

We shall be using results about the mapping and regularity properties for elliptic operators which are uniformly degenerate. The theory here is drawn from [28], but see also [25].

Let X be a manifold with boundary, and suppose that (x, y) is a local chart near some boundary point, where x is a boundary defining function and y restricts to coordinates along the boundary. A differential operator L is called uniformly degenerate if in any such chart it takes the form

$$L = \sum_{j + |\alpha| \leq m} a_{j,\alpha}(x, y)(x \partial_x)^j (x \partial_y)^\alpha.$$

6

We assume that the coefficients are smooth, or at least $\mathcal{C}^{2,\alpha}$ up to ∂X. There is a well-defined uniformly degenerate symbol

$$^0\sigma_m(L)(x, y, \xi, \eta) = \sum_{j+|\alpha|=m} a_{j,\alpha}(x, y)\xi^j\eta^\alpha$$

and L is elliptic in this category of objects if this symbol is invertible for all (x, y) and $(\xi, \eta) \neq 0$. Unlike in the standard interior case, there is a further model which must be studied, called the normal operator, which is defined by

$$N(L) = \sum_{j+|\alpha|\leq m} a_{j,\alpha}(0, y)(t\partial_t)^j(t\partial_v)^\alpha,$$

where (t, v) are linear coordinates on the half-space $\mathbb{R}_t^+ \times \mathbb{R}_v^\ell$, $\ell + 1 = \dim X$. Finally, for any such operator, we define its set of indicial roots to be the values of μ for which $Lx^\mu = \mathcal{O}(x^{\mu+1})$. (This definition must be modified slightly when L is a system.) These values are the roots of the indicial polynomial $\sum_{j\leq m} a_{j0}(0, y)\mu^j$, so (in the scalar case) there are exactly m such values. For simplicity, we now restrict to the case where the degree of L is 2, and list the indicial roots as μ_1 and μ_2.

We shall let these operators act on weighted Sobolev and Hölder spaces of functions. By definition $H_0^k(X)$ consists of functions which lie L^2 along with all derivatives up to order k with respect to the vector fields $x\partial_x$ and $x\partial_y$. Similarly, $\Lambda_0^{k,\alpha}$ denotes the Hölder space where the derivatives and difference quotients are measured with respect to these same vector fields. If E is any function space, then $x^\mu E$ denotes the set of functions $x^\mu v$ where $v \in E$.

The basic result we need is the following:

Proposition 2.1. *Let L be a uniformly degenerate operator of degree 2 on the compact manifold with boundary X, and suppose that L is uniformly degenerate elliptic. If $N(L)$: $t^{\mu-1/2}H_0^2(dtdv) \longrightarrow t^{\mu-1/2}L^2(dtdv)$ is an isomorphism for one value of $\mu \in (\mu_1, \mu_2)$, then it is an isomorphism for every $\mu \in (\mu_1, \mu_2)$, and for all such μ,*

$$L : x^\mu\Lambda_0^{k+2,\alpha}(X) \longrightarrow x^\mu\Lambda_0^{k,\alpha}(X)$$

is Fredholm, with nullspace contained in $x^{\mu_2}\Lambda_0^{\ell,\alpha}(X)$ for every ℓ. If $N(L)$ is only surjective as a map $t^{\mu-1/2}H_0^2(dtdv) \to t^{\mu-1/2}L^2(dtdv)$ but its nullspace is nontrivial, then L itself still has closed range of finite codimension, but an infinite dimensional kernel.

The proof is contained in [28].

2.3 Properly embedded minimal surfaces with embedded asymptotic boundary

As explained in the introduction, there is a rich existence theory for properly embedded minimal or area-minimizing surfaces in convex cocompact hyperbolic 3-manifolds. Something not treated in Anderson's original investigations is the boundary regularity. One expects that a properly embedded minimal surface Y is as regular as its asymptotic boundary curve γ. This problem and its generalization to higher dimensional minimal codimension one submanifolds was investigated by Lin [26], Hardt and Lin [21] and Tonegawa [39]. The higher codimension case has apparently not been treated at all, but is in fact not so difficult using the theory of uniformly degenerate elliptic operators; we

shall come back to this in a later paper. In general dimensions and codimensions, if γ is smooth then any corresponding minimal Y with $\partial Y = \gamma$ is polyhomogeneous at the boundary, i.e. has an expansion in powers of any defining function for M restricted to Y; when $\dim Y$ is even, only positive integer powers appear, while if $\dim Y$ is odd, then powers of $x^k \log x$ also appear; all of this is completely analogous to the situation for PE metrics. The case of importance here, however, is covered by the various papers cited above:

Proposition 2.2. *Let γ be $\mathcal{C}^{k,\alpha}$ embedded curve in ∂M, where $M = \mathbb{H}^3 / \Gamma$ is convex cocompact, $k \in \mathbb{N}$, $0 < \alpha < 1$. Then Y is $\mathcal{C}^\infty$ in the interior of M and $\mathcal{C}^{k,\alpha}$ up to $\gamma = \overline{Y} \cap \partial M$.*

We discuss some features of the proof in order to bring out some consequences. This result is local in γ, so we may as well suppose that $M = \mathbb{H}^3$ and focus on the behaviour of Y near some fixed point $p \in \gamma$. Using the upper half-space model with coordinates $y \in \mathbb{R}^2$, $x > 0$, place p at the origin and choose a local arc-length parametrization $\gamma(s)$ for γ (with respect to the standard Euclidean metric on $\mathbb{R}^2$). Let Γ denote the vertical cylinder over γ, i.e. $\Gamma = \{(y, x) \in \mathbb{R}^2 \times \mathbb{R}^+ : y \in \gamma\}$; thus near the origin, $\Gamma = \{(\gamma(s), x)\}$.

Choose two smooth families of minimal hemispheres, i.e. totally geodesic copies of $\mathbb{H}^2$, which lie completely inside and outside of γ, respectively, and which are tangent to γ, and let $\Gamma^\pm$ be the envelopes of these families. These are smooth mean-convex surfaces tangent to Γ along γ, and it is straightforward to use them as barriers to deduce that Y must lie in the open set between Γ^- and Γ^+. It follows that Y is vertical along γ, or equivalently, that its unit normal with respect to the Euclidean metric on the upper half-space is tangent to $\mathbb{R}^2 = \{x = 0\}$ along γ. We now write Y as a horizontal graph over Γ. More specifically, if $\overline{N} = \overline{N}(s)$ is the unit normal (again with respect to the Euclidean metric) at a point of Γ, then there is a scalar function $u(s, x)$ and a neighbourhood $\mathcal{U}$ of the origin so that

$$Y \cap \mathcal{U} = \{F(s, x) := (\gamma(s) + u(s, x)\overline{N}(s), x) : |s| < \epsilon, x < \epsilon\}.$$

The argument above implies that $u(s, 0) = \partial_s u(s, 0) = 0$.

The regularity of Y along γ is equivalent to that of this function u, and the key point is that u is a solution of a uniformly degenerate elliptic partial differential equation $\mathcal{F}(u) = 0$ corresponding to the minimality of Y, which we derive now. The function F induces a coordinate chart on Y; let the indices 1 and 2 refer to the s and x coordinates, respectively. Letting $\overline{T} = \gamma'(s)$, then

$$F_s = (1 - \kappa u)\overline{T} + u_s \overline{N}, \qquad F_x = u_x \overline{N} + \partial_x,$$

where κ is the curvature of γ. For convenience below, write $w = 1 - \kappa u$. The inward pointing $\overline{g}$ unit normal is equal to

$$\overline{\nu} = \frac{F_x \times F_s}{|F_x \times F_s|} = J^{-1}(-u_s\overline{T} + w\overline{N} - u_x w \partial_x), \qquad J := \sqrt{u_s^2 + w^2(1 + u_x^2)}. \tag{2.5}$$

The coefficients of the first fundamental form and its inverse are

$$(\overline{g}_{ij}) = \begin{bmatrix} w^2 + u_s^2 & u_x u_s \\ u_x u_s & 1 + u_x^2 \end{bmatrix}, \qquad \text{and} \qquad (\overline{g}^{ij}) = \frac{1}{J^2} \begin{bmatrix} 1 + u_x^2 & -u_x u_s \\ -u_x u_s & w^2 + u_s^2 \end{bmatrix}.$$

Next, we compute that

$$F_{ss} = (w_s - \kappa u_s)\overline{T} + (u_{ss} + \kappa w)\overline{N}, \quad F_{sx} = -\kappa u_x \overline{T} + u_{xs}\overline{N}, \quad F_{xx} = u_{xx}\overline{N},$$

8

$$(\overline{k}_{ij}) = -\frac{1}{J}\begin{bmatrix} w(u_{ss} + \kappa w) - u_s(w_s - \kappa u_s) & wu_{xs} + \kappa u_x u_s \\ wu_{xs} + \kappa u_x u_s & wu_{xx} \end{bmatrix}.$$

Finally, use the general formula $k_{ij} = e^\phi(\overline{k}_{ij} + \partial_{\overline{\nu}}\phi\,\overline{g}_{ij})$ relating the second fundamental forms of Y of two conformally related metrics $g = e^{2\phi}\overline{g}$. Here $\phi = -\log x$ and $\overline{\nu}$ is as in (2.5), so the matrix (k_{ij}) is equal to

$$-\frac{1}{Jx}\begin{bmatrix} w(u_{ss} + \kappa w) - u_s(w_s - \kappa u_s) - x^{-1}u_x w(w^2 + u_s^2) & wu_{xs} + \kappa u_x u_s - x^{-1}u_x^2 u_s w \\ wu_{xs} + \kappa u_x u_s - x^{-1}u_x^2 u_s w & wu_{xx} - x^{-1}u_x w(1 + u_x^2) \end{bmatrix}.$$

The equation of minimality, i.e. that $g^{ij}k_{ij} = H = 0$, is then given by the expression

$$\begin{aligned} \mathcal{F}(u) &:= (1 + u_x^2)\,[w(u_{ss} + \kappa w) - u_s(w_s - \kappa u_s)] - 2u_x u_s\,(wu_{xs} + \kappa u_x u_s) \\ &+ w(w^2 + u_s^2)u_{xx} - 2\tfrac{wu_x}{x}\left(u_s^2 + w^2(1 + u_x^2)\right) = 0. \end{aligned} \tag{2.6}$$

The coefficient $1/x$ in this last term makes this a degenerate elliptic equation.

Assume that γ is at least $\mathcal{C}^3$; we compute the first few coefficients in the expansion of $u(s,x)$ as $x \searrow 0$. Set $u \sim u_2(s)x^2 + u_3(s)x^3 + \dots$ (since we already know that u vanishes to second order). Inserting this into $\mathcal{F}(u) = 0$ yields that $u_2(s) = \frac{1}{2}\kappa(s)$, but $u_3(s)$ is formally undetermined by the equation. In other words, this coefficient must depend globally on Y. Just as in the Fefferman-Graham expansion for PE metrics, all higher terms in the expansion for u are determined by γ and u_3 and their derivatives, so we regard (γ, u_3) as the Cauchy data for the minimal surface Y. Using the unique continuation theorem from [29], it is straightforward to show that if Y_1 and Y_2 are two minimal surfaces with the same Cauchy data (γ, u_3) (even locally), then $Y_1 \equiv Y_2$. This global coefficient u_3 plays a central role in our work.

As a side remark for the moment, consider $\mathcal{C}^3$ surfaces with boundary $\overline{Y} \subset \overline{M}$ with $\partial Y \subset \partial M$, which intersect ∂M orthogonally (this makes sense since $\overline{M}$ has a conformal structure). Any such Y can still be represented near the boundary as a normal graph over the vertical cylinder Γ over its boundary curve γ, and the graph function still vanishes to second order. It is no longer necessarily true that $u_2 = \frac{1}{2}\kappa$. The second fundamental form now satisfies

$$(k_{ij}) = \begin{bmatrix} \frac{1}{x}(2u_2 - \kappa) + 3u_3 + \mathcal{O}(x) & -2u_2' + \mathcal{O}(x^2) \\ -2u_2' + \mathcal{O}(x^2) & -3u_3 + \mathcal{O}(x) \end{bmatrix};$$

note that we now have only $|k|_g = \mathcal{O}(x)$ unless $2u_2 = \kappa$ in which case $|k|_g = \mathcal{O}(x^2)$.

Finally, since the Jacobian term $J = |F_s \times F_x| = 1 + \mathcal{O}(x^2)$, we see that in these coordinates, the area form equals

$$dA = \frac{1 + \mathcal{O}(x^2)}{x^2}\,dsdx.$$

Writing $Y_\epsilon = Y \cap \{x \geq \epsilon\}$, then by definition, the renormalized area of Y is the constant term in the expansion

$$\int_{Y_\epsilon} dA = \frac{\text{length}(\gamma_\epsilon)}{\epsilon} + \mathcal{A}(Y) + \mathcal{O}(\epsilon). \tag{2.7}$$

In order for this to be interesting, we must show that $\mathcal{A}(Y)$ is well-defined, independently of the choice of special bdf x. This was done by Graham and Witten [20]; their key observation, which is particularly simple in this low-dimensional setting, is that if h_0

and $\widehat{h}_0 = e^{2\chi_0} h_0$ are two representatives of the conformal class $\mathfrak{c}(g)$, corresponding to special bdf's x and $\widehat{x}$, respectively, then $\widehat{x} = e^\chi x$, where $\chi(x, y) = \mathcal{O}(x^2)$. This means that $\widehat{x} = x + \mathcal{O}(x^2)$, and hence in the new coordinate system $(\widehat{s}, \widehat{x})$ on Y, one still has $dA = \widehat{x}^{-2}(1 + \mathcal{O}(\widehat{x}^2))\, d\widehat{s}d\widehat{x}$. From this, the claim about well-definedness of $\mathcal{A}(Y)$ is immediate.

The only property about Y needed for this argument to work is that it is at least $\mathcal{C}^2$ and meets ∂M orthogonally. Thus even in this broader setting there is still a well-defined notion of renormalized area of Y. To maintain the distinction, we shall denote this extended renormalized area functional by $\mathcal{R}$ rather than $\mathcal{A}$ when the surface Y is not minimal.

3   A formula for renormalized area

We now express the renormalized area of a properly embedded minimal surface Y in M in terms of its Euler characteristic and an integral of local invariants. In fact, since it is not much more complicated to do so, we find an expression for the renormalized area when Y lies in an arbitrary Poincaré-Einstein space of any dimension and is not necessarily minimal, but still meets ∂M orthogonally.

Proposition 3.1. *Let (M^{n+1}, g) be a PE space and $\gamma \subset \partial M$ a $\mathcal{C}^{3,\alpha}$ embedded curve, and suppose that $Y^2 \subset M$ is a properly embedded minimal surface with asymptotic boundary γ, an embedded closed curve in ∂M. Then the renormalized area $\mathcal{A}$ of Y is equal to*

$$\mathcal{A}(Y) = -2\pi\chi(Y) - \frac{1}{2}\int_Y |\widehat{k}|^2\, dA + \int_Y W_{1212}\, dA, \qquad (3.8)$$

where $\widehat{k}$ is the trace-free second fundamental form of Y and W_{1212} is the Weyl curvature of g evaluated on any orthonormal basis for TY. In particular, the integrals on the right are convergent. If Y is any properly embedded surface which extends to be a $\mathcal{C}^2$ surface with boundary in $\overline{M}$ intersecting the boundary orthogonally, then (with the convention that $H = (\operatorname{tr} k)/2$), the renormalized area is equal to

$$\mathcal{A}(Y) = -2\pi\chi(Y) + \frac{1}{2}\int_Y \left(2|H|^2 - |\widehat{k}|^2\right) dA + \int_Y W_{1212}\, dA.$$

Proof. We begin with some preliminary observations and calculations.

First, denote by $R_{ijk\ell}$ and $(R_Y)_{ijk\ell}$ the components of the curvature tensor of g and of the induced metric on Y, respectively. The Ricci curvature of g satisfies $R_{ij} = -ng_{ij}$, and from the standard decomposition of the curvature tensor of an Einstein metric, the components of the Weyl tensor for g are given by

$$W_{ijk\ell} = R_{ijk\ell} + g_{ik}g_{j\ell} - g_{i\ell}g_{jk}. \qquad (3.9)$$

Fix a point $p \in Y$ and choose an oriented orthonormal basis $\{e^1, \ldots, e^{n+1}\}$ for T_pM such that e^1 and e^2 are an oriented basis for T_pY. Now, denoting by k_{ij}^s, $i, j = 1, 2$, $s = 3, \ldots, n+1$, the components of the second fundamental form of Y at p, the Gauss-Codazzi equations become

$$R_{1212} = (R_Y)_{1212} - \sum_{s=3}^{n+1}(k_{11}^s k_{22}^s - k_{12}^s k_{12}^s) = (R_Y)_{1212} - |H|^2 + \frac{1}{2}|\widehat{k}|^2.$$

To check this last equality, simply note that for each s, $k_{11}^s + k_{22}^s = 2H^s$, so $k_{ii}^s = \widehat{k}_{ii}^s + H^s$ and $k_{ij}^s = \widehat{k}_{ij}^s$ for $i \neq j$, and hence

$$
\begin{aligned}
\sum_{s=3}^{n+1} (k_{11}^s k_{22}^s - k_{12}^s k_{12}^s) &= \sum_{s=3}^{n+1} \left((\widehat{k}_{11}^s + H^s)(\widehat{k}_{22}^s + H^s) - (\widehat{k}_{12}^s)^2 \right) \\
&= \sum_{s=3}^{n+1} (H^s)^2 - \tfrac{1}{2}((\widehat{k}_{11}^s)^2 + (\widehat{k}_{22}^s)^2 + 2(\widehat{k}_{12}^s)^2).
\end{aligned}
$$

Combined with (3.9), this gives

$$
(R_Y)_{1212} + \frac{1}{2}|\widehat{k}|^2 - |H|^2 - W_{1212} = -1. \tag{3.10}
$$

This equation holds at each point. The first term on the left is simply the Gauss curvature K of Y; for simplicity, we continue to write W_{1212} for the third term on the left, noting that it is independent of orthonormal frame.

Now integrate over $Y_\epsilon = Y \cap \{x \geq \epsilon\}$ to obtain

$$
\int_{Y_\epsilon} K \, dA - \frac{1}{2} \int_{Y_\epsilon} \left(2|H|^2 - |\widehat{k}|^2 \right) dA - \int_{Y_\epsilon} W_{1212} \, dA = - \int_{Y_\epsilon} dA.
$$

By the Gauss-Bonnet theorem, since $\chi(Y_\epsilon) = \chi(Y)$ for ϵ small enough,

$$
\int_{Y_\epsilon} K \, dA = 2\pi\chi(Y) - \int_{\gamma_\epsilon} \kappa_\epsilon \, ds,
$$

where κ_ϵ is the geodesic curvature of the boundary $\gamma_\epsilon := \partial Y_\epsilon$ in Y_ϵ and ds is the length element with respect to the metric induced by g. Altogether we get

$$
\int_{Y_\epsilon} dA = -2\pi\chi(Y) + \int_{\gamma_\epsilon} \kappa_\epsilon \, ds + \frac{1}{2} \int_{Y_\epsilon} \left(2|H|^2 - |\widehat{k}|^2 \right) dA + \int_{Y_\epsilon} W_{1212} \, dA.
$$

To proceed further, we use the formula that as $\epsilon \searrow 0$,

$$
\int_{\gamma_\epsilon} \kappa_\epsilon \, ds = \frac{\text{length}\,(\gamma_\epsilon)}{\epsilon} + \mathcal{O}(\epsilon). \tag{3.11}
$$

Deferring the proof of this for a moment, using the basic definition of renormalized area via Hadamard regularization in (2.7), we find that

$$
\mathcal{A}(Y) = -2\pi\chi(Y) + \lim_{\epsilon \to 0} \left(\frac{1}{2} \int_{Y_\epsilon} \left(2|H|^2 - |\widehat{k}|^2 \right) dA + \int_{Y_\epsilon} W_{1212} \, dA \right). \tag{3.12}
$$

In order to show that the second and third terms on the right have limits as $\epsilon \searrow 0$, recall the transformation law

$$
\widehat{k}_{ij}(e^{2\phi}g) = e^\phi \widehat{k}_{ij}(g),
$$

for the trace-free second fundamental form $\widehat{k}(g)$ under the conformal change of ambient metric from g to $e^{2\phi}g$ (this is true no matter the dimension or codimension of the submanifold Y). When $\dim Y = 2$,

$$
|\widehat{k}(e^{2\phi}g)|^2_{e^{2\phi}g} \, dA_{e^{2\phi}g} = |\widehat{k}(g)|^2_g \, dA_g.
$$

Similarly, the components of the Weyl tensor transform as

$$
W_{ijk\ell}(e^{2\phi}g) = e^{-2\phi}W_{ijk\ell}(g) \implies W_{1212}(e^{2\phi}g) \, dA_{e^{2\phi}g} = W_{1212}(g) \, dA_g.
$$

11

Thus these two potentially worrisome terms do have a limit. Similarly, even when Y is not minimal, by the calculations in §2, the mean curvature H is $\mathcal{O}(x)$, so its integral has a limit too.

It remains to prove (3.11). Denote by $\overline{\kappa}_\epsilon$ the geodesic curvature of γ_ϵ with respect to the metric $\overline{g} = x^2 g$ and $\overline{n}$ the interior $\overline{g}$-unit normal to ∂Y_ϵ in Y_ϵ. Since $u \sim u_2 x^2 + u_3 x^3 + \mathcal{O}(x^{3+\alpha})$, it follows that $\overline{n} = (1 + \mathcal{O}(x^2))\partial_x + V$, where $\overline{g}(V, \partial_x) = 0$. Now, geodesic curvature also transforms nicely under conformal re-scalings: $\kappa_\epsilon = \epsilon(\overline{\kappa}_\epsilon + \partial_{\overline{n}} \log x)$. Since $\overline{g}(F_{ss}, \partial_x) = 0$ at $x = 0$, we deduce $\overline{\kappa}_\epsilon = \mathcal{O}(\epsilon)$, hence $\kappa_\epsilon = 1 + \mathcal{O}(\epsilon^2)$; recalling too that $ds = \epsilon^{-1} d\overline{s}$, we obtain finally

$$\int_{\gamma_\epsilon} \kappa_\epsilon \, ds = \frac{\text{length } (\gamma_\epsilon)}{\epsilon} + \mathcal{O}(\epsilon),$$

as claimed. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The expression $\frac{1}{2} \int_{Y_\epsilon} (2H^2 - |\widehat{k}|^2) \, dA$ is finite only when Y intersects ∂M orthogonally. Indeed, using the notation and formulæ from §2.3 again, suppose that Y is written as a normal graph over the vertical cylinder Γ over the boundary curve γ, but do not assume that $u_x(s, 0) \equiv 0$. Now $H = 2u_x(s, 0) + \mathcal{O}(x)$, as follows from the formulæ $H = x(\overline{H} + 2\partial_{\overline{\nu}}(\log x))$ and $\partial_{\overline{\nu}} x = \overline{g}(\overline{\nu}, \partial_x) = u_x(s, 0) + \mathcal{O}(x)$. Recalling again that $|\widehat{k}|^2 \, dA$ is conformally invariant, we see that

$$\frac{1}{2} \int_{Y_\epsilon} (2H^2 - |\widehat{k}|^2) \, dA = \frac{4|u_x(s, 0)|^2}{\epsilon} + \mathcal{O}(\log \epsilon)$$

does not have a limit as $\epsilon \searrow 0$ unless $u_x(s, 0) \equiv 0$. This is consistent with the fact that the definition of renormalized area $\mathcal{A}(Y)$ via Hadamard regularization is independent of choice of special bdf x only when this same condition is satisfied.

4 The moduli spaces $\mathcal{M}_k(M)$

Fix the convex cocompact hyperbolic 3-manifold (M, g) and an integer $k \geq 0$. We define $\widetilde{\mathcal{M}}_k(M)$ to be the space of all properly embedded surfaces of genus k which extend to $\overline{M}$ as $\mathcal{C}^{3,\alpha}$ submanifolds with boundary and which intersect ∂M orthogonally, and $\mathcal{M}_k(M)$ the subspace of all such surfaces which are minimal. In this section we study the structure of these moduli spaces, which are the natural domains for the renormalized area functional, as well as some properties of the natural map Π which assigns to any such Y its asymptotic boundary $\partial Y = \gamma$, which is a $\mathcal{C}^{3,\alpha}$ closed (but possibly disconnected) embedded curve in ∂M.

It is a standard fact that the space of all $\mathcal{C}^{3,\alpha}$ surfaces $Y \subset M$ with boundary γ lying in ∂M is a Banach manifold. The space $\widetilde{\mathcal{M}}_k(M)$ defined above is clearly a closed submanifold (of infinite codimension). Similarly, the space $\mathcal{E}$ of all $\mathcal{C}^{3,\alpha}$ closed embedded (but not necessarily connected) curves $\gamma \subset \partial M$ is also a Banach manifold. The corresponding structure for the smaller space of minimal surfaces is also true.

Proposition 4.1. *For each k, $\mathcal{M}_k(M)$ is a Banach manifold.*

Proof. Fix any $Y \in \mathcal{M}_k(M)$ and assume for the moment that $\partial Y = \gamma$ is actually a $\mathcal{C}^\infty$ embedded curve in ∂M. We construct a coordinate neighbourhood around Y in $\mathcal{M}_k$ which in the generic (nondegenerate) setting is modelled on a small ball around 0 in the space of Jacobi fields for the minimal surface operator on Y which are $\mathcal{C}^{3,\alpha}$ up to ∂Y;

this ball in turn is identified with a small ball in the space of $\mathcal{C}^{3,\alpha}$ normal vector fields along γ. We make this nondegeneracy condition explicit below.

To set this up, let ν be the unit normal (with some fixed choice of orientation) along Y. If ϕ is any scalar function on Y which is small in $\mathcal{C}^{3,\alpha}$, we can define a new surface

$$Y_{0,\phi} = \{\exp_p(\phi(p)\nu(p)) : p \in Y\},$$

which we call a normal graph over Y.

The mean curvature of $Y_{0,\phi}$ is computed by a nonlinear elliptic second order operator $\mathcal{F}(\phi)$. The precise expression of this operator is rather complicated, but its linearization has the familiar form

$$D\mathcal{F}|_{\phi=0} := L_Y = \Delta_Y + |A_Y|^2 - 2;$$

here A_Y is the second fundamental form of Y and Δ_Y is its Laplacian with respect to the induced metric.

This Jacobi operator, L_Y, is an elliptic uniformly degenerate operator of order 2. Its normal operator is

$$N(L_Y) = t^2 \partial_t^2 + t^2 \partial_v^2 - 2$$

since the second fundamental form A_Y vanishes at ∂Y; the leading (second order) term is just the Laplacian on the hyperbolic plane, so $N(L_Y) = \Delta_{\mathbb{H}^2} - 2$. The indicial roots are $\mu_1 = -1$, $\mu_2 = 2$, hence solutions of $L_Y u = 0$ satisfy $u \sim a(y)x^{-1} + \ldots$ or $u \sim a(y)x^2 + \ldots$. Note that since the g- and $\overline{g}$-unit normals are related by $\nu = x\overline{\nu}$, in the case where u blows up as $x \searrow 0$, the product $u\nu = (xu)\overline{\nu}$ behaves like $(a(y) + \mathcal{O}(x))\overline{\nu}$, or in other words, the solutions growing at this rate are the ones which are bounded (but not blowing up) at $x = 0$ with respect to $\overline{g}$, and hence correspond to moving the boundary curve γ nontrivially. In this $\overline{g}$ normalization, the decaying Jacobi fields vanish like x^3, which should be no surprise. In any case, it follows directly from self-adjointness and integration by parts that

$$N(L_Y) : t^\mu H_0^2(\mathbb{H}^2; t^{-2}dtdv) = t^{1+\mu} H_0^2(\mathbb{H}^2; dtdu) \longrightarrow t^{1+\mu} L^2(\mathbb{H}^2; dtdu)$$

is invertible when $\mu = 0$. By Proposition 2.1, this is true for any $-1 < \mu < 2$ and for any μ in this range,

$$L_Y : x^\mu \Lambda_0^{2,\alpha} \longrightarrow x^\mu \Lambda_0^{0,\alpha} \tag{4.13}$$

is Fredholm of index zero. We call the minimal surface Y *nondegenerate* if the nullspace K_μ of this mapping contains only 0 for any $\mu \in (-1, 2)$; in this case, (4.13) is surjective. In general, its cokernel is canonically identified with K_μ in the following sense. First note that by Proposition 2.1 again, $K_\mu \subset x^2 \Lambda_0^{2,\alpha}(Y)$ (indeed, if γ is smooth, any $u \in K_\mu$ is polyhomogeneous, i.e. has full tangential regularity), so we may as well drop the subscript μ. Next, if $f \in x^\mu \Lambda_0^{0,\alpha}(Y)$ lies in the range of (4.13), $f = L_Y w$, then obviously $\int_Y uf \, dA_Y = \int_Y uL_Y w \, dA_Y = 0$ for all $u \in K$. (Note that this integral makes sense since $\mu > -1$.) However, this gives precisely the correct number of linear conditions, so this necessary condition is also sufficient.

To study $\mathcal{M}_k(M)$, we must consider a broader class of deformations of Y where the boundary curve γ also varies. Let $\overline{\nu} = x^{-1}\nu$ be the unit normal to Y with respect to the conformally compactified metric $\overline{g} = x^2 g$. This vector field extends smoothly to $\overline{Y}$, and its restriction to $\gamma = \partial Y$ is the unit normal $\overline{N}$ to this curve in ∂M with respect to h_0. Any nearby curve can be written as a normal graph

$$\gamma_\psi = \{\exp_p(\psi(p)\overline{N}(p)) : p \in \gamma\}$$

(where now exp is with respect to h_0). We now define an extension operator $\mathcal{E}$ which assigns to any small ψ a surface $Y_{\psi,0}$ which is 'approximately minimal' and which has $\partial Y_{\psi,0} = \gamma_\psi$. To do this, let u be the graph function for Y over the cylinder Γ. We define a new graph function u_ψ in some neighbourhood $\{x < \epsilon\}$ of the boundary such that $u_\psi(s,0) = \psi(s)$, and $\partial_x^j u_\psi(s,0)$, $j = 1, 2$ is determined by the formal expansion of solutions for $\mathcal{F}$; $\partial_x^3 u_\psi(s,0)$ could be chosen freely, but we set it equal to $u_3(s)$. Now let $U_\psi = \chi u_\psi + (1 - \chi)u$ where χ is a cutoff function which equals 1 near $x = 0$. It is not hard to check that $\mathcal{F}(U_\psi) \in x^\mu \Lambda^{1,\alpha}$ for some $0 < \mu < 2$. The extension $\mathcal{E}$ can be chosen to depend smoothly on ψ. We then have that

$$D\mathcal{E}|_0\,(\hat{\psi}) = w$$

is a function on Y which satisfies $w \sim x^{-1}\hat{\psi}$ as $x \searrow 0$ and $L_Y w = \mathcal{O}(x^\mu)$ for some $\mu \in (0, 2)$.

Finally, perturb $Y_{\psi,0}$ to a normal graph over it using the unit normal for $Y_{\psi,0}$ and as graph function any small $\phi \in x^\mu \Lambda_0^{2,\alpha}(Y)$. The resulting surface will be denoted $Y_{\psi,\phi}$, and we write its mean curvature as $\mathcal{F}(\psi, \phi)$. Thus if $\mathcal{B}$ is a small neighbourhood of the origin in $\mathcal{C}^{3,\alpha}(\gamma) \times x^\mu \Lambda_0^{3,\alpha}(Y)$, then

$$\mathcal{F} : \mathcal{B} \longrightarrow x^\mu \Lambda_0^{1,\alpha}(Y) \tag{4.14}$$

is a smooth mapping.

A neighbourhood of Y in $\mathcal{M}_k(M)$ is identified with the space of solutions to $\mathcal{F}(\psi, \phi) = 0$, and so may be studied by the implicit function theorem. Note that

$$D\mathcal{F}|_{(0,0)}\,(\hat{\psi}, \hat{\phi}) = L_Y(D\mathcal{E}(\hat{\psi}) + \hat{\phi}).$$

When Y is nondegenerate, $D_2\mathcal{F}|_{(0,0)} = L_Y$ on $x^\mu \Lambda_0^{2,\alpha}(Y)$ is already surjective; this yields the existence of a smooth map $\mathcal{G}$ defined in a neighbourhood of 0 in $\mathcal{C}^{3,\alpha}(\gamma)$ to $x^\mu \Lambda_0^{2,\alpha}(Y)$ such that $\mathcal{F}(\psi, \mathcal{G}(\psi)) \equiv 0$, and so that all elements of the nullspace of $\mathcal{F}$ near $(0,0)$ are of this form.

In the degenerate case, we must show that by allowing $\hat{\psi}$ to vary over some suitable finite dimensional subspace of infinitesimal deformations of γ, we can still obtain a surjective map. If this were to fail, then there would exist a nontrivial $u \in K$ such that for all $\hat{\psi}$ and $\hat{\phi}$, $L_Y(D\mathcal{E}(\hat{\psi}) + \hat{\phi}) \perp u$. Write $\eta = D\mathcal{E}(\hat{\psi})$. Then

$$0 = \int_Y L_Y(\eta + \hat{\phi})u = \int_\gamma n \cdot \nabla(\eta + \hat{\phi})u - (\eta + \hat{\phi})n \cdot \nabla u = -2\int_\gamma \hat{\psi}u_0,$$

which implies that u_0 (the leading coefficient of x^2 in the expansion of u) is orthogonal to every $\hat{\psi}$, which is impossible. This proves that $\mathcal{F}$ is always surjective as a function of both (ψ, ϕ), and hence finally that $\mathcal{M}_k(M)$ is a smooth Banach manifold in a neighbourhood of Y. $\qquad\square$

Proposition 4.2. *The natural map*

$$\Pi : \mathcal{M}_k(M) \longrightarrow \mathcal{E}(\partial M)$$

given by $\Pi(Y) = \partial Y$ *is Fredholm with index 0.*

(Recall that this means that if $Y \in \mathcal{M}_k(M)$, then $D\Pi_Y$ is a Fredholm map from $T_Y \mathcal{M}_k(M)$ to $T_{\Pi(Y)}\Gamma$ with the dimensions of its kernel and cokernel equal to one another.)

14

Proof. Let K denote the nullspace of L_Y acting on functions of the form $D\mathcal{E}|_0(\hat{\psi}) + \hat{\phi}$, $\hat{\psi} \in \mathcal{C}^{3,\alpha}(\gamma)$, $\hat{\phi} \in x^\mu \Lambda_0^{2,\alpha}(Y)$. If Y is nondegenerate, then K contains no elements of the form $(0, \hat{\phi})$, so $D\Pi_Y$ is an isomorphism. If Y is degenerate, however, then the proof above shows that if $\ell = \dim(K \cap x^\mu \Lambda_0^{2,\alpha}(Y))$, so that $\dim \ker D\Pi_Y = \ell$, then we can make L_Y surjective by supplementing $x^\mu \Lambda_0^{2,\alpha}(Y)$ with an ℓ-dimensional space H of functions of the form $D\mathcal{E}_0(\hat{\psi})$ (and we may even assume that each $\hat{\psi}$ is $\mathcal{C}^\infty$). Let H' be any choice of complement of H in $\mathcal{C}^{3,\alpha}(\gamma)$. The implicit function function theorem shows that there exists a smooth map $\mathcal{G}$ from H' to $H \oplus x^\mu \Lambda_0^{3,\alpha}(Y)$ such that $\mathcal{F}(\hat{\psi}, \mathcal{G}(\hat{\psi})) \equiv 0$, so the codimension of the range of $D\Pi_Y$ is ℓ too. Hence $\text{ind}(D\Pi_Y) = 0$, as claimed. $\qquad\square$

The final general result about these moduli spaces is contained in the

Proposition 4.3. Π *is a proper mapping.*

Proof. We must show that if γ_j is a sequence of elements in $\mathcal{E}$ such that $\gamma_j \to \gamma$ in $\mathcal{C}^{3,\alpha}$, and if $Y_j \in \mathcal{M}_k(M)$ has $\partial Y_j = \gamma_j$, then (possibly after passing to a subsequence) Y_j converges to a properly embedded minimal surface Y with genus k and $\partial Y = \gamma$.

Let Γ be the vertical cylinder over γ, and let u_j be the horizontal graph function corresponding to the surface Y_j. A priori, the function u_j is only defined on some vertical strip where $x < \epsilon_j$. The first step is to show that ϵ_j can be chosen independently of j. The only thing which prevents these graphs from existing on a uniform strip would be if the u_j did not have a uniform gradient bound, or in other words, that there exists a sequence (s_j, x_j) with s_j in the parameter interval for γ and $x_j \searrow 0$, and such that $|\nabla u_j(s_j, x_j)| = 1$, say (any positive number would do), and $|\nabla u_j(s_j, x_j)| < 1$ for all s and for $x < x_j$. The gradient and norm here are with respect to the Euclidean metric. Perform a hyperbolic rescaling by a factor $\frac{1}{x_j}$, centered at the point $(\gamma(s_j), 0)$, and then a translation and rotation to move $(s_j, 0, 0)$ to the origin in $\mathbb{R}^2$ and to make the rescaled curve γ_j tangent to the y^1-axis. The result is a minimal surface $\widetilde{Y}_j$ in the upper half-space, defined in a ball of expanding radius tending to infinity, which passes through $(0, 0, 0)$ in the boundary, and which can be expressed as a horizontal graph $y^2 = F_j(y^1, x)$ over some large ball in the vertical (y^1, x)-plane. By construction, $F_j(0, 1) = \frac{u_j(s_j, x_j)}{x_j}$; by Rolle's theorem, this is bounded by $|\partial_x u_j(s_j, x_j')|$ for some $x_j' < x_j$, hence by construction $|F_j(0, 1)| \leq 1$ and $\frac{\partial F_j}{\partial x}(0, 1) = 1/x_j$.

Passing to a subsequence, as $j \to \infty$ this minimal surface converges to a complete minimal surface $\widetilde{Y} \subset \mathbb{H}^3$ whose boundary is the limit of rescalings of γ, i.e. a straight line, and which can be expressed as a horizontal graph $y^2 = F(y^1, x)$ over all of $\mathbb{R}_{y^1} \times \mathbb{R}_x^+$ with $|F(0, 1)| \leq 1$. However, by construction, the tangent space of $\widetilde{Y}$ is not vertical at the point $(0, 1, F(0, 1))$, which contradicts the fact that the unique minimal surface in hyperbolic space with boundary a straight line is a totally geodesic plane.

This argument proves that the graph functions u_j are defined on a uniform interval $[0, \epsilon]$, and moreover that the boundary curves at height $x = \epsilon$ are also converging in $\mathcal{C}^{3,\alpha}$ (in fact, in $\mathcal{C}^\infty$ by interior elliptic estimates). Notice that this already proves that no handles can slide off to infinity, provided the boundary curves remain uniformly smooth enough. Let $Y_{j,\epsilon} = Y_j \cap \{x \geq \epsilon\}$. This is now a sequence of compact minimal surfaces with boundary in the convex set $\{x \geq \epsilon\} \subset M$. The proof will be finished if we can prove that these surfaces have a convergent subsequence. This in turn follows from the results of Anderson [4] and White [41]. In order to apply their results, it suffices to show that the genera of the $Y_{j,\epsilon}$ remain bounded, which is obvious by definition, and that the areas of these surfaces are also bounded. This follows from the Gauss-Bonnet theorem: since

15

each Y_j is minimal, its Gauss curvature satisfies $K \leq -1$, which implies that

$$\text{Area}\,(Y_{j,\epsilon}) \leq \int_{Y_{j,\epsilon}} (-K)\,dA = -2\pi\chi(Y_j) + \int_{\partial Y_{j,\epsilon}} \kappa\,ds.$$

Now, the first term is fixed, so we must show that the second term is bounded. But this is immediate from standard elliptic estimates applied to the graph function u_j for Y_j in the annulus $\epsilon/2 < x < 2\epsilon$ since ϵ is now fixed. □

As explained in [40], it is important to work with a slightly different regularity condition: we shall replace $\mathcal{C}^{k,\alpha}$ by the closure in this space of $\mathcal{C}^\infty$. This smaller subspace is separable, whereas $\mathcal{C}^{k,\alpha}$ is not, so with this new regularity restriction (which we shall not comment on further) both $\widetilde{\mathcal{M}}_k(M)$ and $\mathcal{E}(\partial M)$ are separable Banach manifolds.

Using all of these facts, we may now define the degree of Π by

$$\deg(\Pi) = \sum_{Y \in \Pi^{-1}(\gamma)} (-1)^{n(Y)},$$

where γ is a regular value of Π, so each $Y \in \Pi^{-1}(\gamma)$ is nondegenerate, and where $n(Y)$ denotes the number of negative eigenvalues of $-L_Y$. This degree is a well-defined invariant on each component of $\mathcal{E}(\partial M)$ (once we have fixed the integer k and the component of $\mathcal{M}_k(M)$ mapping to that isotopy class of boundary curves).

For example, when $M = \mathbb{H}^3$, and γ is any convex curve, then by the maximum principle, there is exactly one properly embedded minimal surface Y with $\partial Y = \gamma$, and necessarily, its genus is 0. This proves that when M is the entire hyperbolic space, then $\deg(\Pi_0) = 1$ while $\deg(\Pi_k) = 0$ for $k > 0$, on the component of $\mathcal{E}$ containing connected curves. This has some interesting consequences. For example, Anderson [3] displayed a connected curve which bounds a minimal surface of genus $k > 0$; by genericity, we can assume that this curve is regular for Π, and since the degree equals zero, we obtain the existence of yet another element in $\mathcal{M}_k(\mathbb{H}^3)$ with boundary equal to this same curve. On the other hand, de Oliveira and Soret [30] construct *stable* properly embedded minimal surfaces in $\mathbb{H}^3$ with arbitrary genus, where the boundary curve has any prescribed number of components. Here too, for any given boundary curve, we conclude the existence of at least one other element of $\mathcal{M}_k$ with that boundary curve which is unstable. It would be interesting to compute the degree of Π precisely in some of these other cases.

5   Area minimization and renormalized area

We now investigate the role of locally area minimizing surfaces in the study of renormalized area for minimal and nonminimal properly embedded surfaces.

5.1   Renormalized area of absolute minimizers

Proposition 5.1. *Let γ be a $\mathcal{C}^{3,\alpha}$ embedded curve in ∂M which bounds in M. Suppose that Y_1 and Y_2 are two properly embedded minimal surfaces with $\partial Y_1 = \partial Y_2 = \gamma$. If Y_1 is area minimizing in M, then $\mathcal{A}(Y_1) \leq \mathcal{A}(Y_2)$, and equality holds if and only if Y_2 is also an area minimizer.*

Proof. Fix a special boundary defining function x and set $Y_{j,\epsilon} = Y_j \cap \{x \geq \epsilon\}$ and $\gamma_{j,\epsilon} = \partial Y_{j,\epsilon}$. The functions u_1 and u_2 for these two surfaces agree up to order three, so in terms of any local coordinate s on γ, $|u_1(s,x) - u_2(s,x)| \leq Cx^3$, with corresponding

estimates for the first 3 derivatives. If S_ϵ denotes the region between $\gamma_{1,\epsilon}$ and $\gamma_{2,\epsilon}$ in $\{x = \epsilon\}$, this gives

$$\text{Area}\,(S_\epsilon) = \mathcal{O}(\epsilon).$$

Recalling that $\text{Area}\,(Y) = \epsilon^{-1}\text{length}\,(\partial Y) + \mathcal{A}(Y) + \mathcal{O}(\epsilon)$, we obtain

$$\text{Area}\,(Y_{2,\epsilon}) - \text{Area}\,(Y_{1,\epsilon}) = \mathcal{A}(Y_2) - \mathcal{A}(Y_1) + \mathcal{O}(\epsilon),$$

since $\partial Y_1 = \partial Y_2$.

Thus if $\mathcal{A}(Y_1) > \mathcal{A}(Y_2)$, then the new (nonsmooth) surface $Y'_{1,\epsilon} = Y_{2,\epsilon} \cup S_\epsilon$ would eventually have area smaller than $Y_{1,\epsilon}$. Indeed, by the inequalities above,

$$\text{Area}\,(Y'_{1,\epsilon}) \leq \text{Area}\,(Y_{2,\epsilon}) + C_1\epsilon \leq \text{Area}\,(Y_{1,\epsilon}) + \mathcal{A}(Y_2) - \mathcal{A}(Y_1) + C_2\epsilon < \text{Area}\,(Y_{1,\epsilon})$$

for ϵ small enough. This contradicts the fact that Y_1 is area-minimizing. $\qquad\square$

5.2 The renormalized area spectrum

The result in this last subsection suggests the consideration of the set-valued function

$$\mathcal{E} \ni \gamma \xrightarrow{\mathcal{S}_k} \{\mathcal{A}(Y) : Y \in \mathcal{M}_k(M) \ \ \partial Y = \gamma\},$$

which we call the renormalized area spectrum (of degree k). Properness of the boundary map ensures that $\mathcal{S}_k(\gamma)$ is always a compact set. Proposition 5.1 implies that if Y_0 is any absolutely area minimizing surface with $\partial Y_0 = \gamma$, then $\underline{\mathcal{A}}(\gamma) := \mathcal{A}(Y_0)$ is a lower bound for $\mathcal{S}_k(\gamma)$ for *every* k. On the other hand, trivially by (3.1), this set is bounded above by $-2\pi\chi(Y) = 2\pi(2k + \ell - 2)$, where ℓ is the number of components of γ. In other words,

$$\mathcal{S}_k(\gamma) \subset [\underline{\mathcal{A}}(\gamma), 2\pi(2k + \ell - 2)]. \tag{5.15}$$

Note furthermore that the upper limit is never attained unless there exists a totally geodesic minimal surface Y with $\partial Y = \gamma$, which never happens unless γ is a round circle. The lower bound is attained if and only if there exists a genus k absolute area minimizer with boundary γ.

5.3 Minimizers of the extended renormalized area functional

In section 4.2 of [23], Hirata and Takayanagi assert that minimizers of the extended renormalized area functional amongst all surfaces with a given boundary are necessarily area minimizing surfaces with this same asymptotic boundary. Furthermore, Polyakov has communicated to us his suggestion (based on his work [32]) that the area-minimizers amongst all surfaces with a given boundary and a given genus must also minimize a functional that depepnds on a quadratic expression of the extrinsic curvature. Both these assertions follow easily from our techniques. Recall that we are denoting the extension of $\mathcal{A}$ to this setting by $\mathcal{R}$.

Proposition 5.2. *Let γ be a $C^{3,\alpha}$ closed curve in ∂M which bounds in M. Then the infimum of $\mathcal{R}(Y)$ where Y ranges over the set of all $C^{3,\alpha}$ surfaces with $\partial Y = \gamma$ which intersect ∂M orthogonally is attained only by absolutely area-minimizing surfaces.*

Proof. Let Y be any $\mathcal{C}^{3,\alpha}$ surface which intersects ∂M at $\partial Y = \gamma$. We must show that if $\mathcal{R}(Y) \leq \mathcal{R}(Y')$ for all other such surfaces Y', then Y is absolutely area-minimizing. Fix a decreasing sequence $\epsilon_j \searrow 0$ and let $Y_j = Y \cap \{x \geq \epsilon_j\}$ and $\gamma_j = \partial Y_j$. Let Y'_j denote an area minizing surface with $\partial Y'_j = \gamma_j$. Following the original existence proof by Anderson, possibly after passing to a subsequence, we may assume that

$$Y'_j \to Y' \tag{5.16}$$

where Y' is properly embedded and area-minimizing with $\partial Y' = \gamma$. Standard results imply that the convergence is in $\mathcal{C}^\infty$ in the interior, and also $\mathcal{C}^{3,\alpha}$ up to the boundary. This may be proved by an argument very similar to that used in establishing properness in Proposition 4.3. Indeed, if the convergence were not $\mathcal{C}^1$, we would be able to take a suitable rescaling and obtain a limiting surface which has boundary a straight line in $\mathbb{R}^2 = \partial \mathbb{H}^3$ but which is not a totally geodesic plane, which would be a contradiction. Knowing this, if the convergence were not $\mathcal{C}^{3,\alpha}$, we could again rescale and extract a limiting surface which converges to a totally geodesic plane, but not smoothly, which also contradicts standard convergence results for minimal surfaces.

Denote the genus of Y' by k, so Y' minimizes $\mathcal{A}$ in $\cup_s \mathcal{M}_s(M)$ amongst all minimal surfaces with the same boundary and of arbitrary genus.

We now claim that

$$\text{Area}\,(Y_j) \geq \text{Area}\,(Y'_j) = \frac{1}{\epsilon_j}\text{length}(\gamma) + \mathcal{A}(Y') + o(1). \tag{5.17}$$

The first inequality is by definition. As for the second equality, we proceed in steps. Since $Y'_j \to Y'$ in $\mathcal{C}^{3,\alpha}$, there exists $\rho > 0$ so that each of the annuli $Y'_j \setminus Y'_{j,\rho} := Y'_j \cap \{x \leq \rho\}$ and $Y' \setminus Y'_\rho := Y' \cap \{x \leq \rho\}$ is a normal graph over some portion of the vertical cylinder Γ of height ρ above γ. In particular, suppose that $Y'_j \setminus Y'_{j,\rho}$ is the graph of a function u_j, defined on some band $\Gamma_j = \{\epsilon_j \leq x \leq \rho\}$ and $Y' \setminus Y'_\rho$ is the graph of a function u defined on the band $\{0 \leq x \leq \rho\}$. Clearly the areas of the portions where $x \geq \rho$ converge, i.e. $\text{Area}\,(Y'_{j,\rho}) \to \text{Area}\,(Y'_\rho)$. Thus if we denote by $Y'j, \rho, \epsilon_j$ the portion of Y' that lies below the hyperplane $x = \rho$ and above the hyperplane $x = \epsilon_j$ it suffices to show that

$$\lim_{j \to \infty} \left(\text{Area}\,(Y'_{j,\rho,\epsilon_j} - \text{Area}\,(Y'_j \setminus Y'_{j,\rho}) \right) = 0.$$

This follows by direct computation. First write $u_j(x,s) = x^2 \tilde{u}_j(x,s)$, $u(x,s) = x^2 \tilde{u}(x,s)$, so that $\tilde{u}_j \to \tilde{u}$ in $\mathcal{C}^1$. Denote by J_u, J_{u_j} the Jacobians for parametrizations of these surfaces in the (s,x) coordinate charts as in §2.3. Using the formulæ from that section we compute that, with constants independent of j,

$$\left| \left(\text{Area}\,(Y'_{j,\rho,\epsilon_j} - \text{Area}\,(Y'_j \setminus Y'_{j,\rho}) \right) \right| = \left| \int_{\epsilon_j}^{\rho} \int_{\gamma} \frac{J_u - J_{u_j}}{x^2}\, ds\, dx \right|$$

$$\leq C \int_{\epsilon_j}^{\rho} \frac{1}{x^2} \int_{\gamma} \left| u_s^2 - u_{j,s}^2 - ((1 - \kappa u)^2(1 + u_x^2) - (1 - \kappa u_j^2)(1 + u_{j,x}^2)) \right|\, ds\, dx \tag{5.18}$$

$$\leq C' \int_{\epsilon_j}^{\rho} \int_{\gamma} |\tilde{u}_j - \tilde{u}| + |\tilde{u}_{j,s} - \tilde{u}_s| + |\tilde{u}_{j,x} - \tilde{u}_x|\, ds\, dx \to 0$$

as $j \to \infty$. This proves (5.17). Subtracting $\epsilon_j^{-1}\text{length}\,(\gamma)$ from (5.17) and passing to the limit gives

$$\mathcal{R}(Y) \geq \mathcal{R}(Y') = \mathcal{A}(Y').$$

18

The second step of the proof is to show that if $\mathcal{R}(Y) = \mathcal{R}(Y')$ then Y must be an area-minimizer. By Proposition 5.1, it suffices to show that Y is minimal. If it is not, then its mean curvature H is nonvanishing in some open set. We can then perturb Y locally, in some small ball $Y \cap B_r(p)$, to a new surface Y'' which is also smooth, but has smaller area in this ball. Clearly the areas of the truncations to $x \geq \epsilon$ satisfy $\text{Area}(Y''_\epsilon) < \text{Area}(Y'_\epsilon)$ for ϵ small, which in turn implies that $\mathcal{R}(Y) < \mathcal{R}(Y')$, contrary to what we proved above. $\square$

6 First and second variations of renormalized area

We now begin the variational analysis of the renormalized area functional $\mathcal{A}$ on each of the moduli spaces $\mathcal{M}_k(M)$, as well as for its extension to the unconstrained spaces $\widetilde{\mathcal{M}}_k(M)$. The first variation formula for $\mathcal{A}$ on $\mathcal{M}_k$ is formally analogous to the corresponding first variation formula for the renormalized volume of Poincaré-Einstein metrics in even dimensions; this formula appears in a paper by Anderson [5], see also Albin [1] for a simpler approach. A formula for the second variation of renormalized volume does not seem to have been computed in the Poincaré-Einstein setting.

6.1 The first variation

We shall compute the first variation of $\mathcal{A}$ at any $Y \in \mathcal{M}_k(M)$; slightly more generally, we compute the first variation of the extended functional $\mathcal{R}$ at any $Y \in \widetilde{M}_k(M)$. Actually, we compute $D\mathcal{R}$ only applied to compactly supported perturbations, and show this is not well-defined for arbitrary variations in $\widetilde{M}_k(M)$.

As before, fix a special boundary defining function x on M, and for any $Y \in \mathcal{M}_k(M)$ let u_3 denote the free third order term in the expansion of the graph function u of Y with respect to x over its vertical cylinder.

Theorem 6.1. *Fix any $Y \in \mathcal{M}_k(M)$ with $\partial Y = \gamma$. Let Y_t denote a smooth one-parameter curve in $\widetilde{\mathcal{M}}_k(M)$, $\partial Y_t = \gamma_t$, with $Y_0 = Y$ (so the surface Y_0 is minimal but the surfaces Y_t, $t \neq 0$, need not be). Write Y_t as a normal graph Y_ϕ over $Y_0 = Y$ via*

$$Y_t = \{F_t(p) = \exp_p(\phi_t(p)\nu(p)) : p \in Y\}$$

where $\phi_0 = 0$.

Setting $\dot{\phi} = \frac{d}{dt}\phi_t\big|_{t=0}$, then $\dot{\phi} \sim x^{-1}\dot{\phi}_0 + \ldots$, and we have

$$\frac{d}{dt}\bigg|_{t=0} \mathcal{A}(Y_t) = -3\int_\gamma \dot{\phi}_0 \, u_3 \, ds.$$

On the other hand, if $Y \in \widetilde{\mathcal{M}}_k(M)$ and $\phi_t = 0$ outside some compact set $K \subset Y$, then:

$$\frac{d}{dt}\bigg|_{t=0} \mathcal{A}(Y_t) = 2\int_Y H\dot{\phi}dA$$

Proof. By virtue of (3.8), we have

$$\frac{d}{dt}\bigg|_{t=0} \mathcal{A}(Y_t) = \frac{1}{2}\frac{d}{dt}\bigg|_{t=0}\left(\int_{Y_t} (\text{tr}\, k_t)^2 - |k_t|^2 \, dA_t\right), \tag{6.19}$$

where k_t and dA_t denote the second fundamental form and area form on Y_t ($\text{tr}\, k_t$ denotes the trace of the second fundamental form of Y_t).

19

In order to compute this, introduce the following notation. Fix any point $p \in Y$ and choose an orthonormal moving frame $\{e_1, e_2\}$ on Y near p, as well as a unit normal vector field ν, so that e_1 and e_2 are the principal directions and $\nabla_{e_i} e_j = -\kappa_i \delta_{ij}\nu$ at p (the κ_i are the principal curvatures of Y). Define $\eta_i(t) = (F_t)_*(e_i)$, so $\eta_i(0) = e_i$ but $\{\eta_1(t), \eta_2(t)\}$ is no longer orthonormal when $t \neq 0$. We define

$$g_{ij}(t) = \langle \eta_i(t), \eta_j(t)\rangle \qquad \text{and} \qquad k_{ij}(t) = -\langle \nabla_{\eta_i(t)} \eta_j(t), \nu(t)\rangle,$$

where $\nu(t)$ is the unit vector orthogonal to both $\eta_1(t)$ and $\eta_2(t)$ with $\nu(0) = \nu = e_3$. Finally, write $T = F_*\partial_t$, so $T = \dot{\phi}\nu$ when $t = 0$.

In the following, all terms are to be computed eventually at p. We first compute that

$$\nabla_T \eta_i|_{t=0} = \nabla_T \nabla_{\eta_i} F|_{t=0} = \nabla_{e_i} \nabla_T F|_{t=0} = \nabla_{e_i}\left(\dot{\phi}\nu\right) = \dot{\phi}\kappa_i e_i + \dot{\phi}_i \nu,$$

which yields immediately

$$T g_{ij}(t)|_{t=0} = \langle \dot{\phi}\kappa_i e_i + \dot{\phi}_i \nu, e_j\rangle + \langle e_i, \dot{\phi}\kappa_j e_j + \dot{\phi}_j \nu\rangle = 2\dot{\phi} k_{ij},$$

since $k_{ij} = \kappa_i \delta_{ij}$ in this frame. This also implies that $T g^{ij}(t)|_{t=0} = -2\dot{\phi} k^{ij}$. It is also standard that

$$\frac{d}{dt}\bigg|_{t=0} F^*(dA_t) = \dot{\phi}(\operatorname{tr} k_{ij})\, dA_0.$$

Finally,

$$T k_{ij}(t)|_{t=0} = -T\langle \nabla_{\eta_i} \eta_j, \nu\rangle = -\left(\langle \nabla_T \nabla_{\eta_i} \eta_j, \nu\rangle + \langle \nabla_{\eta_i} \eta_j, \nabla_T \nu\rangle\right)\big|_{t=0}.$$

By assumption, $\nabla_{e_i} e_j$ is orthogonal to Y, while $\nabla_\nu \nu$ is tangential, so the second term on the right vanishes. By definition of the curvature tensor, the other term equals

$$-\langle \nabla_{\eta_i} \nabla_T \eta_j + \nabla_{[T,\eta_i]} \eta_j + R(T, \eta_i)\eta_j, \nu\rangle = -\langle \nabla_{e_i}(\dot{\phi}\kappa_j e_j + \dot{\phi}_j \nu), \nu\rangle - \dot{\phi}\langle R(\nu, e_i)e_j, \nu\rangle.$$

Observe that

$$[T, \eta_i] = \nabla_T e_i - \nabla_{e_i}(\dot{\phi}\nu) = \dot{\phi}\kappa_i e_i + \dot{\phi}_i \nu - \dot{\phi}_i \nu - \dot{\phi}\kappa_i e_i = 0,$$

so expanding out yields that

$$T k_{ij}|_{t=0} = \dot{\phi}\kappa_j^2 \delta_{ij} - \dot{\phi}_{ij} + R_{3i3j} = -(\nabla^2 \dot{\phi})_{ij} + \dot{\phi}(k \circ k)_{ij} + R_{3i3j}.$$

Putting these formulæ together, we compute that

$$\begin{aligned}
\frac{d}{dt}\bigg|_{t=0} \frac{1}{2}\int_{Y_t} (\operatorname{tr} k_t)^2 - |k_t|^2\, dA_t &= -\int_Y \left(-\nabla^2 \dot{\phi}_{ij}\right) k^{ij} + \Delta\dot{\phi}\operatorname{tr} k\, dA \\
&+ \int_Y (\operatorname{tr} k)^3 - 3|k|^2(\operatorname{tr} k) + 2\operatorname{tr}(k \circ k \circ k)\, dA + 2\int_Y \dot{\phi}(R_{3i3}{}^i \operatorname{tr} k - R_{3i3j}k^{ij})\, dA.
\end{aligned} \tag{6.20}$$

(Here $\operatorname{tr}(k \circ k \circ k) = g^{ip}g^{jq}g^{\ell r}k_{ir}k_{pj}k_{q\ell}$.) A simple calculation using the principal curvature decomposition shows that

$$(\operatorname{tr} k)^3 - 3|k|^2(\operatorname{tr} k) + 2\operatorname{tr}(k \circ k \circ k) = 0,$$

while since M is hyperbolic,

$$R_{3i3}{}^i \operatorname{tr} k - R_{3i3j}\, k^{ij} = -\operatorname{tr} k.$$

This proves

$$\frac{d}{dt}\bigg|_{t=0} \frac{1}{2} \int_{Y_t} (\operatorname{tr} k_t)^2 - |k_t|^2 \, dA_t = -\int_Y -\nabla^2 \dot\phi_{ij} k^{ij} + \Delta \dot\phi \operatorname{tr} k \, dA + 2\int_Y \dot\phi \operatorname{tr} k \, dA. \quad (6.21)$$

The next step is to evaluate $\int_Y \left(-\nabla^2 \dot\phi_{ij} k^{ij} + \Delta \dot\phi \operatorname{tr} k\right) \, dA$. To do this, we integrate by parts on $Y_\epsilon = Y \cap \{x \geq \epsilon\}$ to get

$$-\int_{Y_\epsilon} \left(\langle \nabla^2 \dot\phi, k\rangle - \Delta\dot\phi \operatorname{tr} k\right) \, dA =$$
$$\int_{Y_\epsilon} \dot\phi_i \left(\nabla^j k_{ij} - \nabla^i k_j{}^j\right) dA + \int_{\gamma_\epsilon} \left(\langle \nabla\dot\phi, n\rangle \operatorname{tr} k - k(\nabla\dot\phi, n)\right) ds, \quad (6.22)$$

where $\gamma_\epsilon = \partial Y_\epsilon$ and n is the g-unit normal in Y_ϵ to γ_ϵ. The contracted Codazzi equation states that

$$\nabla^j k_{ij} = \nabla_i \operatorname{tr} k + \operatorname{Ric}_{\nu i},$$

but in fact $\operatorname{Ric}_{\nu i} = -2g_{\nu i} = 0$ since the i index refers to a vector tangent to Y. Hence only the boundary terms remain. If $Y \in \widetilde{M}_k(M)$ then since $\dot\phi$ is compactly supported, these boundary terms vanish. If $Y \in \mathcal{M}_k(M)$, however, then $\operatorname{tr} k = 0$ so the first part of the integrand in the boundary integral vanishes and we only have to evaluate the second one.

To do this, revert to the (s, x) coordinates introduced in §2. In terms of these, the expression becomes

$$\int_{\gamma_\epsilon} \dot\phi_i g^{ij} k_{j\ell} n^\ell \, ds.$$

To calculate this, we first note that the unit normal $n = n^1 \partial_s + n^2 \partial_x$ has coefficients which satisfy

$$n^1 = \mathcal{O}(x^3), \quad n^2 = 1 + \mathcal{O}(x^2).$$

Thus we may as well set $\ell = 2$ and $n^2 = 1$. Note also that $g^{ij} = x^2 \overline{g}^{ij}$ and $ds = x^{-1} d\overline{s}$, where $\overline{s}$ is arclength on γ with respect to $\overline{g}$. (This differs slightly from our earlier convention.) The terms in the integrand are thus

$$\left(\dot\phi_1 x^2 (\overline{g}^{11} k_{12} + \overline{g}^{12} k_{22}) + \dot\phi_2 x^2 (\overline{g}^{21} k_{12} + \overline{g}^{22} k_{22})\right) d\overline{s}.$$

Finally, recall that $\dot\phi \sim x^{-1}\dot\phi_0$ and use the expansions for the k_{ij} and g_{ij} to deduce that this reduces to

$$-3\int_{\gamma_\epsilon} \dot\phi_0 u_3 \, d\overline{s} + \mathcal{O}(\epsilon).$$

This completes the proof. $\qquad\square$

From this formula we deduce the

Corollary 6.1. *Suppose that $Y \in \mathcal{M}_k(M)$ is a critical point for $\mathcal{A}$. If Y is a nondegenerate element in $\mathcal{M}_k$ (i.e. it has no decaying Jacobi fields), then the coefficient u_3 in its boundary expansion vanishes identically. In general, if Y is degenerate, then we may only conclude that u_3 is orthogonal to the finite dimensional space spanned by leading coefficients ψ_0 of Jacobi fields $\psi \sim \psi_0 x^{-1} + \psi_1 x^0 + \ldots$ along Y. This is equivalent to the statement that u_3 lies in the (finite dimensional) span of all leading coefficients ψ_3 of decaying Jacobi fields $\psi \sim \psi_3 x^2 + \ldots$ on Y.*

If $Y \in \widetilde{M}_k(M)$ is critical for $\mathcal{R}$, then Y is necessarily minimal, and $u_3 = 0$ regardless of whether or not Y is degenerate in $\mathcal{M}_k(M)$.

We remark that it would be quite interesting to know if there exist degenerate critical points of $\mathcal{A}$.

6.2 The second variation

We now derive the second variation formula for $\mathcal{A}$ at a general element $Y \in \mathcal{M}_k$. The calculation is more involved than for the first variation formula, but is actually fairly elementary since it relies primarily on some detailed trigonometric calculations.

Before stating the result, we set up the basic notation. We use the same language as in Theorem 6.1. Let Y_t be any 1-parameter family of surfaces in $\mathcal{M}_k(M)$, with corresponding family of boundary curves γ_t in ∂M. The calculations are local, so we assume that $M = \mathbb{H}^3$ and we work in standard upper half-space coordinates. Let $\overline{N}_t$ denote the inward-pointing unit normal to γ_t with respect to the Euclidean metric on $\mathbb{R}^2$. Let (s, v) be Fermi coordinates around γ_0, i.e. corresponding to the map $(s, v) \mapsto \gamma_0(s) + v\overline{N}_0(s)$. Then γ_t can be written as a normal graph, $v = \psi_t(s)$, where $\psi_0 \equiv 0$ and $\dot{\psi}\overline{N}_0$ is the normal variation vector field. Similarly, introduce Fermi coordinates for the Euclidean metric around Y_0, using the inward unit normal $\overline{\nu}$; the cylindrical coordinates (s, x) parametrize Y_0, and we write Y_t as a normal graph $v = \phi_t(s, x)$ using the $\overline{g}$-unit normal $\overline{\nu}$. The variation vector field is $\dot{\phi}\overline{\nu}$, so $\dot{\phi}$ is a solution of the Jacobi equation $L_{Y_0}\dot{\phi} = 0$. The regularity theory for this equation shows that it has an expansion

$$\dot{\phi}(s, x) \sim \dot{\phi}_0(s) + \dot{\phi}_2(s)x^2 + \dot{\phi}_3(s)x^3 + \dots,$$

where $\dot{\phi}_0(s) = \dot{\psi}(s)$. If Y_0 is nondegenerate, the coefficient of x^3 is determined by a global Dirichlet-to-Neumann map in terms of the leading coefficient, which we write as

$$\dot{\phi}_3 = \frac{1}{6} \left.\partial_x^3\right|_{x=0} \dot{\phi} = \mathcal{D}_{Y_0}\dot{\phi}_0.$$

Theorem 6.2. *The second variation of $\mathcal{A}$ along the family Y_t is given by*

$$D^2\mathcal{A}\big|_{Y_0}(\dot{\phi}, \dot{\phi}) = -3\int_{\gamma_0} \dot{\phi}_0\dot{\phi}_3 \, d\overline{s},$$

or equivalently, when Y_0 is nondegenerate,

$$-3\int_{\gamma_0} \dot{\phi}_0\mathcal{D}_{Y_0}\dot{\phi}_0 \, d\overline{s}.$$

Proof. Let us begin by setting up some notation. Write each Y_t as a horizontal graph over the vertical cylinder Γ_t on the curve γ_t, i.e. via $(\gamma_t(s), x) + u_t(s, x)\overline{N}_t$. Through the remainder of this proof, subscripts refer only to t-dependence but not derivatives, and we often write $\dot{w}$ for $\partial_t w$. Set

$$u_t(s, x) = \frac{1}{2}\kappa_t(s)x^2 + u_{3,t}(s)x^3 + \mathcal{O}(x^{3+\alpha}).$$

The unit tangent and normal vectors of γ_t are

$$\overline{T}_t = \frac{1}{\sqrt{(\partial_s\phi_t)^2 + (1 - \kappa_0\phi_t)^2}}\left((1 - \kappa_0\phi_t)\overline{T}_0 + \partial_s\phi_t\overline{N}_0\right),$$

$$\overline{N}_t = \frac{1}{\sqrt{(\partial_s\phi_t)^2 + (1 - \kappa_0\phi_t)^2}}\left(-\partial_s\phi_t\overline{T}_0 + (1 - \kappa_0\phi_t)\overline{N}_0\right),$$

so the component along $\overline{N}_t$ of the variation vector field $\dot\phi\,\overline{N}_0$ is

$$\langle \overline{N}_t, \dot\phi\overline{N}_0 \rangle = \dot\phi_t \frac{1 - \kappa_0\phi_t}{\sqrt{(\partial_s\phi_t)^2 + (1 - \kappa_0\phi_t)^2}} = \dot\phi_t \left(1 + (\partial_s\phi_t/(1 - \kappa_0\phi_t)^2)\right)^{-1/2}.$$

This is equal to $\dot\phi_t(1 + \mathcal{O}(t^2))$. Hence the first variation of $\mathcal{A}$ at Y_t is given by

$$\frac{d}{dt}\mathcal{A}(Y_t) = -3\int_{\gamma_t} u_{3,t}\dot\phi_t(1 + \mathcal{O}(t^2))\,d\bar{s}.$$

Differentiating once more at $t = 0$ yields

$$\frac{d^2}{dt^2}\mathcal{A}(Y_t)\Big|_{t=0} = -3\int_{\gamma_0}\left(\dot{u}_{3,0}\,\dot\phi_0 + u_{3,0}\ddot\phi_0\right)d\bar{s},$$

The second term in the integrand is an artifact of the parametrization. Indeed, recall that for the composition of any two functions, $A(s)$ and $s = Y(t)$, $(A(Y(t))'' = A''(Y_0)(Y')^2 + A'(Y_0)Y''$, but the Hessian of A corresponds only to the first term on the right. In other words,

$$D^2\mathcal{A}\big|_{Y_0}(\dot\phi, \dot\phi) = -3\int_{\gamma_0}\dot{u}_{3,0}\dot\phi_0\,d\bar{s}. \tag{6.23}$$

The key point is to relate $\dot\phi_3$ to $\dot{u}_{3,0}$. The long calculation that follows involves expressing ϕ_t in terms of u_t and then calculating $\partial_t\partial_x^3$ at $(x,t) = (0,0)$. The relationship between these two terms arises as follows. Fix any $P \in Y_0$ with x coordinate $x(P) = x_0$ sufficiently small. Let $\tilde\lambda$ be the line emanating from $P \in Y_0$ in the direction $\overline\nu(P)$. Denote by $\tilde{P}$ and $\tilde{R}$ its points of intersection with Γ_t and Y_t, respectively. Note that assuming γ_0 is convex near P, then moving from P along the line $\tilde\lambda$, one encounters first $\tilde{P}$ then $\tilde{R}$; we orient coordinates so that all quantities are positive in this situation. By definition, $\phi_t(P)$ is the distance from Y_0 to Y_t along $\tilde\lambda$, and it can be decomposed as the sum of two signed distances, from P to $\tilde{P}$ and from $\tilde{P}$ to $\tilde{R}$, which we call $\tilde\ell$ and $\tilde\chi$, respectively:

$$\phi_t = |P\tilde{R}| = |P\tilde{P}| + |\tilde{P}\tilde{R}| = \tilde\ell + \tilde\chi.$$

We show how to relate each of the two quantities $\tilde\ell$ and $\tilde\chi$ to u_t in turn.

First consider $\tilde\ell$. Let x_0 be the x coordinate of the point P and α the angle between $\tilde\lambda$ and the plane $x = x_0$. From (2.5) the orthogonal projection of $\overline\nu(P)$ onto this plane is spanned by $-\partial_s u\,\overline{T}_0 + (1 - \kappa_0 u_0)\overline{N}_0$, and

$$\sin\alpha = \frac{\partial_x u\,(1 - \kappa_0 u_0)}{\sqrt{(\partial_s u)^2 + (1 - \kappa_0 u_0)^2(1 + (\partial_x u_0)^2)}} \quad\Longrightarrow\quad \alpha = \kappa_0 x + 3u_{3,0}x^2 + \mathcal{O}(x^3).$$

Now let λ^1 denote the line in the plane $x = x_0$ from P in the direction of this projection of $\overline\nu$, and $P^1 = \lambda^1 \cap \Gamma_t$. Setting $\ell^1 = |PP^1|$ (interpreted as a signed distance), since $PP^1\tilde{P}$ is a right triangle, we have

$$\tilde\ell = \frac{\ell^1}{\cos\alpha} = \ell^1(1 + \frac{1}{2}\alpha^2 + \ldots) = \ell^1(1 + \frac{1}{2}\kappa_0^2 x^2 + 3\kappa_0 u_{3,0}x^3 + \ldots).$$

We shall prove below that

$$\ell^1\big|_{x=0} = \psi_t, \quad \partial_x\ell^1\big|_{x=0} = 0, \quad\text{and}\quad \partial_x^3\ell^1\big|_{x=0} = 6u_{3,0} + \mathcal{O}(t^2), \tag{6.24}$$

and hence, after some calculation,

$$\partial_t \partial_x^3 \tilde{\ell}\Big|_{x=t=0} = 18\kappa_0 u_{3,0}\dot{\phi}_0. \tag{6.25}$$

To prove (6.24), we examine ℓ^1 more closely. Consider the (very thin) triangle T in the plane $x = x_0$ with sides the segments of the line λ^1, the line $\bar{\lambda}$ emanating from P with direction $\overline{N}_0$, and the line tangent to γ_t at the point of intersection $\bar{\lambda} \cap \gamma_t$. Let ℓ^2, $\bar{\ell}$ and $\hat{\ell}$ denote the lengths of these sides, respectively, and let β and ω denote the angles between $\lambda^1 \cap \bar{\lambda}$ and $\bar{\lambda} \cap \gamma_t$. Note that

$$
\begin{aligned}
\cos\beta &= \frac{1}{\sqrt{(\partial_s u_0)^2 + (1-\kappa_0 u_0)^2}} \langle -\partial_s u \overline{T}_0 + (1-\kappa_0 u_0)\overline{N}_0, \overline{N}_0 \rangle \\
&= \left(1 + (\partial_s u_0/(1-\kappa_0 u_0))^2\right)^{-1/2} = 1 - \tfrac{1}{2}(\tfrac{1}{2}\partial_s\kappa_0\, x^2 + \partial_s u_{3,0}\, x^3 + \ldots)^2 + \ldots
\end{aligned}
$$

$$\implies \beta = \frac{1}{2}\partial_s\kappa_0\, x^2 + \partial_s u_{3,0}\, x^3 + \ldots. \tag{6.26}$$

In addition, since $\pi/2 - \omega$ is the angle between $\overline{N}_0$ and $\overline{N}_t$,

$$\cos\left(\frac{\pi}{2} - \omega\right) = 1 - \frac{1}{2}(\partial_s\phi_t + \mathcal{O}(t^2))^2 + \ldots \Rightarrow \omega = \frac{\pi}{2} - \partial_s\phi_t + \mathcal{O}(t^2).$$

The law of sines for this triangle is the set of equalities

$$\frac{\hat{\ell}}{\sin\beta} = \frac{\ell^2}{\sin\omega} = \frac{\bar{\ell}}{\sin(\pi - \beta - \omega)}.$$

Since ω is bounded away from 0, the first equality and (6.26) imply that $\hat{\ell} = \mathcal{O}(x^2)$, which in turn shows that $\ell^2 = \ell^1 + \mathcal{O}(x^4)$. Thus for computing third derivatives, we may as well work with ℓ^2. The second equality yields

$$\ell^2 = \bar{\ell}\,\frac{\sin\omega}{\sin(\pi - \beta - \omega)} = \frac{\bar{\ell}}{\cos\beta}\left(1 - \tan\beta\cot\omega + \mathcal{O}(x^4)\right).$$

Recalling that $\bar{\ell} = u_0 + \psi_t$, and using the expressions for β and ω above, we obtain

$$
\begin{aligned}
\ell^2 &= (\psi_t + u_0)\left(1 + \mathcal{O}(x^4)\right)\left(1 - (\tfrac{1}{2}\partial_s\kappa_0 x^2 + \ldots)(\partial_s\phi_t + \mathcal{O}(t^2))\right) \\
&= \psi_t + \tfrac{1}{2}\kappa_0 x^2 + u_{3,0}x^3 + \mathcal{O}(t^2) + \mathcal{O}(x^4),
\end{aligned}
$$

which yields (6.24).

We turn to the computation of $\tilde{\chi}$. The idea is much the same. Let (s_0, x_0) denote the cylindrical coordinates of P, i.e. as a horizontal graph over the cylinder Γ_0, and (s_0', x_0') the cylindrical coordinates of the point $\tilde{R} = \tilde{\lambda} \cap Y_t$, when written as a horizontal graph over Γ_t. Since $\tilde{R} = P + \phi_t(s,x)\bar{\nu}$ and $\bar{\nu} = \overline{N}_0 - \kappa_0 x \partial_x + \mathcal{O}(x^2)$, we obtain

$$s_0' = s_0 + \mathcal{O}(x^2), \qquad x_0' = x_0 - \kappa_0 x\phi_t + \mathcal{O}(x^2). \tag{6.27}$$

We shall compute the horizontal displacement $u_t(s_0', x_0')$ of Y_t from Γ_t in the plane $x = x_0'$. For this, let $\tilde{\gamma}_t = \Gamma_t \cap \{x = x_0'\}$, which is just the vertical translate of γ_t into this plane. Abusing notation slightly, write $\tilde{P}$ for the point on $\tilde{\gamma}_t$ which is directly below the intersection of $\tilde{\lambda} \cap \Gamma_t$ (which had been called $\tilde{P}$ above). Finally, let $\tilde{Q}$ denote the point on $\tilde{\gamma}_t$ which is nearest to $\tilde{R}$, i.e. so that $\tilde{R} = \tilde{Q} + c\overline{N}_t$ for some constant c.

Consider the thin triangle $\tilde{P}\tilde{Q}\tilde{R}$. The sides $\tilde{P}\tilde{R}$ and $\tilde{R}\tilde{Q}$ have lengths $\tilde{\chi}\cos\alpha$ and $u_t(s_0', x_0')$, respectively. Let η and ζ be the angles at the vertices $\tilde{P}$ and $\tilde{Q}$. We find that

$$\eta = \frac{\pi}{2} + \mathcal{O}(t) + \mathcal{O}(x^2), \qquad \zeta = \frac{\pi}{2} + \mathcal{O}(x^2).$$

By the law of sines again, $u_t(s_0', x_0')/\sin\eta = \tilde{\chi}/\sin\zeta$, so

$$\tilde{\chi}\cos\alpha = u_t(s_0', x_0')(1 + \mathcal{O}(x^4))(1 + \mathcal{O}((t + x^2)^2)) = u_t(s_0', x_0') + \mathcal{O}(t^2) + \mathcal{O}(x^4).$$

Expanding in a Taylor series, using (6.27) and recalling that $u_t = \mathcal{O}(x^2)$, we find that

$$
\begin{aligned}
u_t(s_0', x_0') &= u_t(s_0, x_0) + \partial_s u_t(s_0, x_0)\mathcal{O}(x^2) + \partial_x u_t(s_0, x_0)(-\kappa_0 x\phi_t + \mathcal{O}(x^2)) + \mathcal{O}(t^2 + x^4) \\
&= u_t(s_0, x_0) - \kappa_0\kappa_t\phi_t x^2 - 3\kappa_0 u_{3,t}\phi_t x^3 + \mathcal{O}(t^2 + x^4) \\
&= (\tfrac{1}{2}\kappa_t - \kappa_0\kappa_t\phi_t)x^2 + (u_{3,t} - 3\kappa_0 u_{3,t}\phi_t)x^3 + \dots;
\end{aligned}
$$

also, as before, $1/\cos\alpha = 1 + \mathcal{O}(x^2)$, so altogether we obtain

$$\partial_t\partial_x^3\tilde{\chi}\big|_{t=x=0} = 6\dot{u}_{3,0} - 18\kappa_0 u_{3,0}\dot{\psi}$$

We have now proved that

$$\partial_t\partial_x^3\phi_t\big|_{x=t=0} = \partial_t\partial_x^3(\tilde{\ell} + \tilde{\chi})\Big|_{x=t=0} = 6\mathcal{D}_{Y_0}(\dot{\phi}_0) = 6\dot{u}_{3,0} + 18\kappa_0 u_{3,0}\dot{\psi}.$$

Inserting this into (6.23) yields the formulæ in the statement of the theorem. $\qquad\square$

7 Critical points of $\mathcal{A}$ in $\mathbb{H}^3$

In this section we prove that the only nondegenerate critical points of renormalized area for proper minimal surfaces in all of $\mathbb{H}^3$ are the totally geodesic planes, the boundary curves of which are circles. The proof requires a preliminary geometric lemma about osculating circles of plane curves, which is perhaps of independent interest, and then proceeds via a refined version of the asymptotic maximum principle.

Recall that the osculating circle C at a point $p \in \gamma$ is a circle which makes second order contact with γ at that point; its curvature, the inverse of its radius, is therefore the same as that for γ at this point of intersection. By inscribed we simply mean that C remains entirely within the closure of one of the two components of $S^2 \setminus \gamma$.

Proposition 7.1. *Let γ be a $\mathcal{C}^2$ embedded loop in $S^2 = \partial\mathbb{H}^3$ and Ω one component of $S^2 \setminus \gamma$. Then there exists a point $p \in \gamma$ so that the osculating curve of γ at p lies in $\overline{\Omega}$.*

Remark 7.1. *As the proof makes clear, there may be many such points. Of course, we could equally well have replaced Ω by the other component of $S^2 \setminus \gamma$.*

Proof. We begin with a few elementary observations. First, the entire question is invariant under Möbius transformations, hence we may freely apply such transformations to reduce the problem to one that is easier to visualize. Second, if C' is any circle inscribed in Ω which is locally maximal (in the sense that for any continuous family of inscribed circles $C'(\epsilon)$ with $C'(0) = C'$, the family of radii $r'(\epsilon)$ reaches a local maximum at $\epsilon = 0$) then necessarily either C' is tangent to γ at two or more distinct points, or else C' is tangent to γ at a single point and is the osculating circle there. The reason is that if there is only one point of contact, $p = C' \cap \gamma$, then the curvature of the circle $1/r'$ is greater than or equal to $\kappa(p)$, the curvature of γ at p. If this inequality is strict and there are

no other points of contact, then we could increase the radius of C' slightly while keeping it inside Ω.

Thirdly, and slightly more complicated, we claim that if C' arises as a limit of inscribed circles C_j' such that each $C_j' \cap \gamma$ contains at least two points $P_j \neq Q_j$ and $\mathrm{dist}\,(P_j, Q_j) \to 0$ as $j \to \infty$, then the limit C' is necessarily an inscribed osculating circle. To see this, choose for each j a Möbius transformation F_j which carries C_j' into a fixed straight line in $\mathbb{R}^2$, say the y_1-axis. We also suppose that $\mathrm{dist}\,(F_j(P_j), F_j(Q_j)) = \mathrm{dist}\,(P_j, Q_j)$, so that F_j does not diverge, that F_j carries Ω to the lower half-plane, and that $\lim F_j(P_j) = \lim F_j(Q_j)$ is the origin. Each curve $F_j(\gamma)$ lies in the upper half-plane and is tangent to the y_1-axis at two points which are converging to the origin. Clearly, the curve must be a graph over the axis between these two points for j large enough, say of some function f_j. By the intermediate value theorem, there is a sequence of points $t_j \to 0$ such that $f_j''(t_j) = 0$. Taking a limit, we see that the limiting curve is flat to second order at the origin and lies entirely in the closed upper half-plane, which proves the claim.

We can now proceed with the proof. Let $\mathcal{K}$ denote the set of pairs $(P, Q) \in \gamma \times \gamma$, $P \neq Q$ such that there is an inscribed circle which is tangent to γ at these two points (and possibly other points as well). By the second remark above, if this set were empty, then there would have to be an inscribed osculating circle already. So assume $\mathcal{K} \neq \emptyset$. We claim that the closure of $\mathcal{K}$ must intersect the diagonal, which by the third remark above would produce an inscribed osculating circle: If it did not, then $\mathcal{K}$ would be compact in $\gamma \times \gamma$, and hence there would be a point $(P', Q') = (\gamma(t_1), \gamma(t_2))$ such that the difference in parameter values $|t_2 - t_1|$ is minimal (for some fixed parametrization of the curve). Let C' be the corresponding circle. Conformally transform so that C' is the y_1-axis. Then the images of P' and Q' lie on this axis and the transformed curve lies entirely on or above the axis. If the image of P' is the left-most point of tangency of the curve with the y_1-axis, fix another point P'' just to the right for which the horizontal component of the downward pointing normal is positive. There is a maximal radius for which a circle tangent to the curve at P'' remains in the component of the lower half-plane, and this circle is obviously tangent to the curve at another point Q'' to the left of the image of Q'. This shows the existence of another pair (P'', Q'') strictly between the pair (P', Q') for which there is an inscribed circle tangent at these two points, which is a contradiction. This finishes the proof. $\square$

Now we turn to the main result of this section.

Theorem 7.1. *Let $Y \in \mathcal{M}_k(\mathbb{H}^3)$ be nondegenerate and suppose that $\partial Y = \gamma$ is $\mathcal{C}^{3,\alpha}$ and connected, and Y is a critical point for $\mathcal{A}$. Then ∂Y is a round circle and Y is a totally geodesic disk.*

Remark 7.2. *Using Corollary 6.1, the proof below actually shows that if Y is critical for the extended renormalized area functional $\mathcal{R}$, then the same conclusion holds.*

Proof. First, by the results in the previous section, we know that Y must be a minimal surface and also that the formally undetermined term u_3 in its expansion must vanish.

Using the result about osculating circles, and applying a conformal transformation, we reduce to the case where γ is a closed curve in $\mathbb{R}^2$ which lies entirely in the closed upper half-plane, and which is tangent to second order to the y_1-axis at the origin. We now write some neighbourhood of $0 \in Y$ as a horizontal graph over the (y_1, x)-plane, i.e. $Y = \{y_2 = u(y_1, x)\}$ for $|y_1| < \delta$, $x < \delta$. Set $s = y_1$ for simplicity. This function satisfies

the minimal surface equation, which in this coordinate system takes the form

$$\mathcal{F}(u) = (1 + u_x^2)u_{ss} - 2u_s u_x u_{sx} + (1 + u_s^2)u_{xx} - \frac{2(1 + u_x^2 + u_s^2)}{x}u_x = 0. \tag{7.28}$$

(Note that this is simply the specialization of (2.6) when the base of the cylinder Γ is flat, so $\kappa = 0$ and $w = 1$.) Furthermore, since Y lies to one side of the (s, x) plane, $u \geq 0$ everywhere, and it has the asymptotic expansion

$$u(s, x) \sim a(s)x^2 + \mathcal{O}(x^{3+\alpha}). \tag{7.29}$$

The absence of the x^3 term is because Y is critical for $\mathcal{A}$; furthermore, $a(0)$ is the one half the curvature of γ at 0, and hence because it osculates the line there, $a(0) = 0$.

Now choose any $\beta \in (0, \alpha)$ and define $u_c = u - cx^{3+\beta}$ for $c > 0$. By the various properties above, if we fix δ then choose c sufficiently small, the function $u_c \geq 0$ on all four sides of the rectangle $|s| \leq \delta$, $0 \leq x \leq \delta$. However, using (7.29) and the fact that $\beta < \alpha$, we also have that $u_c(0, x) < 0$ for x sufficiently small. This means that the minimum of u_c is achieved somewhere strictly inside this rectangle, and of course at that point (s_c, x_c), we have $\mathcal{F}(u_c)(s_c, x_c) \geq 0$.

On the other hand, we compute that

$$\mathcal{F}(u - cx^{3+\beta}) = \mathcal{F}(u) - L_u(cx^{3+\beta}) + \mathcal{O}(cx^{6+2\beta});$$

the first term on the right vanishes, while the operator in the second term is the linearization of $\mathcal{F}$ at u,

$$\begin{aligned}
L_u v &= (1 + u_x^2)v_{ss} - 2u_s u_x v_{sx} + (1 + u_s^2)v_{xx} + 2(u_x u_{ss} - u_s u_{sx})v_x + v_s \\
&\quad + 2(u_s u_{xx} - u_x u_{sx})v_s - \frac{2(1 + u_x^2 + u_s^2)}{x}v_x - \frac{4u_x^2}{x}v_x - \frac{4u_s u_x}{x}v_s.
\end{aligned}$$

By (7.29) and scale-invariant Schauder estimates, $x^{-2}|u| + x^{-1}|\nabla u| + |\nabla^2 u| \leq K$. Hence altogether

$$\mathcal{F}(u_c) \leq -c\beta(3 + \beta)x^{1+\beta} + Kcx^{2+\beta},$$

where K is independent of c and δ. Choosing δ sufficiently small, we can ensure that $\mathcal{F}(u_c) < 0$ everywhere in this box, which contradicts that it is positive at (s_c, x_c). This proves the theorem. $\qquad\square$

8   Connections with the Willmore functional

It is not particularly surprising that the functional $\mathcal{A}$ is connected with the Willmore functional, which by definition is the total integral of the square of mean curvature. In this final section we explore some of these relationships.

Fixing a special bdf x, then the expansion of $\overline{g} = x^2 g$ has only even powers, so its natural extension to a $\mathbb{Z}_2$-invariant metric on the double of M across its boundary is smooth; furthermore, any surface $Y \in \widetilde{M}_k(M)$ can also be doubled to a closed $\mathcal{C}^{2,1}$ surface. We denote these doubles by $2M$ and $2Y$, respectively.

Proposition 8.1. *If $Y \in \mathcal{M}_k(M)$, then*

$$\mathcal{A}(Y) = -\frac{1}{2}\mathcal{W}(2Y).$$

27

Proof. We have already noted that $|\widehat{k}|^2 \, dA_g = |\widehat{\overline{k}}|^2 \, dA_{\overline{g}}$. Next, observe that

$$\left|\widehat{\overline{k}}\right|^2 = 2 \left|\frac{\overline{\kappa}_1 - \overline{\kappa}_2}{2}\right|^2 = 2 \left(\overline{H}^2 - \overline{\kappa}_1 \overline{\kappa}_2\right),$$

where the $\overline{\kappa}_j$ are the principal curvatures of Y with respect to $\overline{g}$. Thus finally

$$
\begin{aligned}
\mathcal{A}(Y) &= -2\pi\chi(Y) - \int_Y \left(\overline{H}^2 - \overline{\kappa}_1 \overline{\kappa}_2\right) dA_{\overline{g}} \\
&= -2\pi\chi(Y) - \frac{1}{2}\int_{2Y} \left(\overline{H}^2 - \overline{\kappa}_1 \overline{\kappa}_2\right) dA_{\overline{g}} = -2\pi\chi(Y) - \frac{1}{2}\mathcal{W}(2Y) + 2\pi\chi(Y),
\end{aligned}
$$

since $\chi(2Y) = 2\chi(Y)$. $\qquad\square$

On the other hand, it does not seem to be the case that the extended functional $\mathcal{R}$ on $\widetilde{\mathcal{M}}_k(M)$ has any simple connection with the Willmore functional. Nonetheless, this result recasts the program of finding extremals of $\mathcal{A}$ in a different and more classical light. Let us conclude by explaining this a bit further.

The conformal invariance of the Willmore functional makes its variational theory quite challenging. The existence of minimizers of the Willmore functional for closed surfaces in $\mathbb{R}^3$ with genus less than or equal to some fixed constant is proved in [37], and this result is sharpened in [7], where it is shown how to prevent a 'drop of genus' in these minimization arguments. We have just shown that the problem of finding extrema, and in particular, maxima, for $\mathcal{A}$ on some given moduli space $\mathcal{M}_k(M)$ is equivalent to finding constrained extrema (in particular, minima) of $\mathcal{W}$ on $2M$ with respect to the metric $\overline{g}$, within the *restricted* class of surfaces which are invariant under the $\mathbb{Z}_2$ involution and which are minimal with respect to g. If it were possible to adapt the arguments from [37] to this ambiently curved setting, we could prove the existence of such extrema. This may well be subtle, and the strengthened result from [7] may not be available, even when $M = \mathbb{H}^3$. Indeed, consider that we have proved that there are no nondegenerate critical points for $\mathcal{A}$ on $\mathcal{M}_k(\mathbb{H}^3)$ when $k > 0$. This indicates that any extremizing sequence Y_j, e.g. one for which $\mathcal{A}(Y_j)$ tends to the supremum, probably does not converge to a surface of the same genus.

References

[1] P. Albin *Renormalizing Curvature Integrals on Poincare-Einstein Manifolds*, arXiv:math/0504161.

[2] M. Anderson *Complete Minimal Varieties in hyperbolic space,* Invent. Math. **69**, (1982), 477-494.

[3] M. Anderson *Complete minimal hypersurfaces in hyperbolic n-manifolds,* Comment. Math. Helv. **58**, (1983), 264-290.

[4] M. Anderson *Curvature estimates for minimal surfaces in 3-manifolds*, Ann. scient. Éc. Norm. Sup. **4** no. 18 (1985), 89-105.

[5] M. Anderson L^2 *curvature and volume renormalization of AHE metrics on 4-manifolds,* Math. Research Lett. **8** (2001), 171-188.

[6] M. Anderson *Topics in conformally compact Einstein metrics*, CRM Proc. Lecture Notes, (2006), 1–26.

[7] M. Bauer and E. Kuwert, *Existence of minimizing Willmore surfaces of prescribed genus*, Int. Math. Res. Not. **10** (2003), 553–576.

[8] O. Biquard *Asymptotically symmetric Einstein metrics*, translated by S Wilson. SMF/AMS Texts and Monographs **13**, (2006), orig. in French as Asterisque **265**, (2000).

[9] S.Y.-A. Chang, J. Qing and P. Yang, *On the renormalized volumes for conformally compact Einstein manifolds* to appear in Proceeding conference in Geometric Analysis, Moscow 2005.

[10] S.Y.-A. Chang, S. Chen and P. Yang, in preparation.

[11] P.T. Chrusćiel, E. Delay, J.M. Lee, D. Skinner *Boundary regularity of conformally compact Einstein metrics* Differential Geometry **69** (2005) 111-136.

[12] B. Coskunuzer, *Minimal planes in hyperbolic space*, Comm. Anal. Geom. 12 (2004) No. 4, 821-836.

[13] B. Coskunuzer, *Generic uniqueness of least area planes in hyperbolic space*, Geom. and Top. 10 (2006), 401-412.

[14] N. Drukker, D. Gross, H. Ooguri *Wilson Loops and Minimal Surfaces* Phys. Rev. D (3) **60** (1999), no. 12.

[15] C. Epstein, *Envelopes of horospheres and Weingarten surfaces in hyperbolic 3-space.* Unpublished manuscript (1983), available at http://www.math.upenn.edu∼cle/papers/index.html

[16] C. Fefferman, C.R. Graham *Conformal Invariants*, Élie Cartan et les mathématiques d'Aujourd'hui, Astérisque, 95-116 (1985).

[17] C. Fefferman, C.R. Graham *The ambient metric*, arXiv:0710.0919.

[18] C.R. Graham *Volume and Area Renormalizations for Conformally Compact Einstein Metrics* Proc. of 19th Winter School in Geometry and Physics, Srni, Czech Rep., (1999).

[19] C.R. Graham, J. M. Lee, *Einstein metrics with prescribed conformal infinity on the ball*, Adv. Math. **87** (1991), 186-225.

[20] C.R. Graham, E. Witten *Conformal Anomaly Of Submanifold Observables In AdS/CFT Correspondence* Nucl.Phys. B *546* (1999) 52-64.

[21] R. Hardt, F. Lin *Regularity at infinity for area minimizing hypersurfaces in hyperbolic space,* Invent. Math. **88**, (1987) 217-224.

[22] M. Henningson and K. Skenderis *T*he holographic Weyl anomaly J. High Energy Phys. **no. 7** (1998).

[23] T. Hirata, T. Takayanagi *AdS/CFT and Strong Subadditivity of Entanglement Entropy* J. High Energy Phys. 2007, no. 2, 042.

[24] K. Krasnov and J.M. Schlenker *On the renormalized volume of hyperbolic 3-manifolds*, Comm. Math. Phys **279** no. 3 (2008), 637-668.

[25] J. Lee, *Fredholm operators and Einstein metrics on conformally compact manifolds*, Mem. Amer. Math. Soc. **183** (2006), no. 864.

[26] F.H. Lin *On the Dirichlet problem for the minimal graphs in hyperbolic space*, Invent. Math. **96**, 593-612 (1989).

[27] J. Maldacena *Wilson loops in Large N field theories*, Phys. Rev. Lett. **80**, 4859 (1998).

[28] R. Mazzeo *Elliptic theory of differential edge operators, I*, Comm.Par.Diff. Eq. **16** No. 10 (1991) 1616 - 1664.

[29] R. Mazzeo *Unique continuation at infinity and embedded eigenvalues for asymptotically hyperbolic manifolds*, Amer. Jour. Math. **113** No. 1 (1991), pp. 25-45.

[30] G. de Oliveira and M. Soret, *Complete minimal surfaces in hyperbolic space*, Math. Annalen 311 (1998), 397-419.

[31] S.J. Patterson and P. Perry, *The divisor of Selberg's zeta function for Kleinian groups.* Appendix A by Charles Epstein. Duke Math. J. **106** (2001), no. 2, 321–390.

[32] A. M. Polyakov *Fine structure of strings* Nuclear Phys. B **268** (1986), no. 2, 406–412.

[33] A. M. Polyakov, V.S. Rychkov, *Loop dynamics and AdS/CFT correspondence* Nuclear Phys. B **594** (2001), no. 1-2, 272–286.

[34] S. Ryu, T. Takayanagi *Holographic derivation of entanglement entropy from AdS/CFT*, Phys. Rev. Lett. **96**, 181602 (2006).

[35] S. Ryu, T. Takayanagi *Aspects of holographic entanglement entropy*, arXiv:hep-th/0605073.

[36] A. Schwimmer, S. Theisen *Entanglement Entropy, Trace Anomalies and Holography* arXiv:0802.1017.

[37] L. Simon *Existence of surfaces minimizing the Willmore functional*, Comm. Anal. Geom. **1** (1993), no. 2, 281-326.

[38] A. Tromba, *Degree theory on oriented infinite dimensional varieties and the Morse number of minimal surfaces spanning a curve in $\mathbb{R}^n$. Part I No. 4*, Trans. Am. Math. Soc. 290 (1985), 385-413.

[39] Y. Tonegawa, *Existence and regularity of constant mean curvature hypersurfaces in hyperbolic space*, Math. Z. **221**, (1996) 591-615.

[40] X. Wang, *Holography and the geometry of certain convex cocompact hyperbolic 3-manifolds* arXiv:math/0210164

[41] B. White, *Curvature estimates and compactness theorems in 3-manifolds for surfaces that are stationary for parametric elliptic functionals*, Invent. Math. 88 (1987), 243-256.

[42] B. White, *The space of m-dimensional surfaces that are stationary for a parametric elliptic functional*, Ind. Univ. Math. Jour. 36 (1987) No. 3, 567-602.