Energy Growth in Minimal Surface Bubbles

John Douglas Moore
Department of Mathematics
University of California
Santa Barbara, CA, USA 93106
e-mail: moore@math.ucsb.edu

November 16, 2007

Abstract

This article is concerned with conformal harmonic maps $f : \Sigma \to M$ from a closed connected surface Σ into a compact Riemannian manifold M of dimension at least four, studied via a perturbative approach based upon the α-energy of Sacks and Uhlenbeck. It gives an estimate on the rate of growth of energy density in the bubbles of minimax α-energy critical points as $\alpha \to 1$, when the bubbles are at a distance at least $L_0 > 0$ from the base. It also describes additional techniques hopefully useful for the development of a partial Morse theory for closed parametrized minimal surfaces (or harmonic surfaces) in compact Riemannian manifolds.

1 Introduction

The Morse theory of geodesics in Riemannian manifolds is a highly successful application of the techniques of global analysis to calculus of variations for nonlinear ODE's. The analogous nonlinear PDE's are the equations for harmonic maps and minimal surfaces. Does there exist a partial Morse theory for closed two-dimensional minimal surfaces in compact Riemannian manifolds? If so, what does it look like?

We treat closed minimal surfaces from the parametrized point of view; thus we regard such minimal surfaces as critical points for the Dirichlet energy function, which we study by means of the α-energy perturbation first introduced by Sacks and Uhlenbeck. The main goal of this article is to give an estimate on energy growth within bubbles of α-energy critical points as $\alpha \to 1$, an estimate motivated by its potential applications to constructing a partial Morse theory for closed two-dimensional minimal surfaces in curved compact ambient manifolds.

We will describe this estimate in §1.2 and will describe the contents of the remainder of the article in §1.3. But first we review the key features of the Morse theory of smooth closed geodesics Bott [2], and describe conditions under which the estimate of §1.2 might be applied.

1.1 Towards a partial Morse theory of parametrized minimal surfaces

If M is a compact Riemannian manifold, we can define the action

$$J : \mathrm{Map}(S^1, M) \to \mathbb{R} \quad \text{by} \quad J(\gamma) = \frac{1}{2}\int_{S^1} |\gamma'(t)|^2 dt, \tag{1}$$

where $\mathrm{Map}(S^1, M)$ denotes a suitable completion of the space of smooth maps from S^1 to M. When we take the L^2_1 completion, this space is denoted by $L^2_1(S^1, M)$ and it is a Hilbert manifold. The Sobolev inequalities give an inclusion $L^2_1(S^1, M) \subset C^0(S^1, M)$, which is well-known to be a homotopy equivalence. Moreover, J is a smooth real-valued function on this Hilbert manifold, and the critical points of J are exactly the smooth closed geodesics in M. Since J satisfies Condition C of Palais and Smale, it is possible to prove existence of minimax critical points by the method of steepest descent, following the orbits of the gradient of $-J$ on the manifold $L^2_1(S^1, M)$.

From here, the development of a Morse theory of smooth closed geodesics proceeds as follows: One shows that for a generic choice of metric on the compact manifold M, all nonconstant smooth closed geodesics lie on one-dimensional nondegenerate critical submanifolds, each such submanifold being an orbit for the action of the group $G = S^1$ of symmetries of J. One then establishes compactness: It follows from Condition C that the number of such submanifolds on which $J \leq J_0$, for any choice of bound J_0, is finite. An analysis of the orbits of the gradient flow for $-J$ then provides (equivariant) Morse inequalities for generic metrics.

Once one has the Morse inequalities for generic metrics, a more refined analysis often provides geometric results for nongeneric metrics. Thus a theorem of Gromoll and Meyer [5] shows that if a compact manifold M has finite fundamental group and satisfies the condition that the Betti numbers of the free loop space $\mathrm{Map}(S^1, M)$ are not bounded, then there exist infinitely many prime smooth closed geodesics for arbitrary choice of Riemannian metric on M. (A geodesic is prime if it is not a multiple cover of a geodesic of smaller length.)

A partial Morse theory for closed parametrized minimal surfaces in a compact Riemannian manifold (M, g) should proceed via the same steps, and should have similar applications. It is helpful to have a problem in the back of our mind to organize development and measure progress:

Basic Problem. Given a closed connected surface Σ, determine conditions on the topology of a compact manifold M with finite fundamental group that ensure that there exist infinitely many imbedded minimal surfaces within M which are diffeomorphic to Σ, when M is given a generic Riemannian metric.

Unlike in the theory of closed geodesics, there are many possible topologies for Σ, and one can expect to obtain quite different answers to the basic problem depending upon the genus g of Σ, and whether Σ is orientable or not. (By the genus of a nonorientable closed connected surface, we mean the genus of its oriented double cover.)

To study the Basic Problem, we consider parametrized minimal surfaces, which we regard as critical points for the *energy* function

$$E : \mathrm{Map}(\Sigma, M) \times \mathrm{Met}(\Sigma) \to \mathbb{R}, \quad \text{defined by} \quad E(f,h) = \frac{1}{2}\int_\Sigma |df|_h^2 dA_h, \tag{2}$$

where $\mathrm{Met}(\Sigma)$ is the space of Riemannian metrics on Σ, and the norm $|\cdot|_h$ and area element dA_h are calculated with respect to $h \in \mathrm{Met}(\Sigma)$. The energy is invariant under Weyl rescalings $h \mapsto e^\lambda h$, where $\lambda : \Sigma \to \mathbb{R}$ is a smooth function, and each element of $\mathrm{Met}(\Sigma)$ is equivalent via such a rescaling to a unique element in the subspace $\mathrm{Met}_0(\Sigma)$ of constant curvature metrics of total area one. Thus we lose nothing in restricting E to $\mathrm{Map}(\Sigma, M) \times \mathrm{Met}_0(\Sigma)$. This restriction is invariant under an obvious action of the group $\mathrm{Diff}_0(\Sigma)$ of diffeomorphisms isotopic to the identity, so E descends to a map on the quotient

$$E : \frac{\mathrm{Map}(\Sigma, M) \times \mathrm{Met}_0(\Sigma)}{\mathrm{Diff}_0(\Sigma)} \longrightarrow \mathbb{R}.$$

The quotient $\mathrm{Met}_0(\Sigma)/\mathrm{Diff}_0(\Sigma)$ can be regarded as the Teichmüller space $\mathcal{T}$ of marked conformal structures on Σ (even if Σ is not orientable). The proof via harmonic maps of Teichmüller's theorem that $\mathcal{T}$ is a finite-dimensional ball ([20], Chapter II) provides a section for the projection

$$\frac{\mathrm{Map}(\Sigma, M) \times \mathrm{Met}_0(\Sigma)}{\mathrm{Diff}_0(\Sigma)} \longrightarrow \frac{\mathrm{Met}_0(\Sigma)}{\mathrm{Diff}_0(\Sigma)},$$

showing that the first quotient is diffeomorphic to $\mathrm{Map}(\Sigma, M) \times \mathcal{T}$. Thus we can regard the energy as a function

$$E : \mathrm{Map}(\Sigma, M) \times \mathcal{T} \to \mathbb{R}, \quad \text{defined by} \quad E(f,\omega) = E(f,h), \tag{3}$$

when h is any metric within the conformal equivalence class $\omega \in \mathcal{T}$.

If Σ is oriented and $\mathrm{Diff}_+(\Sigma)$ denotes the group of orientation-preserving diffeomorphisms, then $\Gamma = \mathrm{Diff}_+(\Sigma)/\mathrm{Diff}_0(\Sigma)$ is called the mapping class group. The action of the mapping class group Γ on $\mathrm{Map}(\Sigma, M) \times \mathcal{T}$ preserves the energy E, which therefore descends once again to a function on the quotient:

$$E : \mathcal{M}(\Sigma, M) \to \mathbb{R}, \quad \text{where} \quad \mathcal{M}(\Sigma, M) = \frac{\mathrm{Map}(\Sigma, M) \times \mathcal{T}}{\Gamma}. \tag{4}$$

This quotient $\mathcal{M}(\Sigma, M)$ projects to the moduli space $\mathcal{R} = \mathcal{T}/\Gamma$ of conformal structures on Σ. This moduli space $\mathcal{R}$ is trivial when $\Sigma = S^2$, is homeomorphic to $\mathbb{C}$ when Σ is T^2, and is an orbifold with a more complicated topology when Σ has genus $g \geq 2$. It is a partial Morse theory for (4) that should be the analog of the Morse theory of smooth closed geodesics.

When Σ has genus zero or one, the energy E is also invariant under the action of the identity component G of the Lie group of complex automorphisms of Σ, which is $PSL(2,\mathbb{C})$ when $\Sigma = S^2$ and $S^1 \times S^1$ when Σ is a torus. We call

G the group of symmetries of E. We say that a minimal surface is *prime* if it is not a nontrivial cover (possibly branched) of a minimal surface of lower energy, possibly nonorientable. It is the number of G-orbits of prime minimal surfaces in $\mathcal{M}(\Sigma, M)$ that we would like to estimate in order to give an answer to the Basic Problem when the genus of Σ is small.

For proving existence of critical points via steepest descent, however, the right completion of $\mathrm{Map}(\Sigma, M)$ is with respect to the L^2_1 norm, and this completion barely fails to lie within $C^0(\Sigma, M)$ via the Sobolev imbedding theorem. We believe that the approach most likely to yield an understanding of the minimax critical points for E needed for partial Morse inequalities is the perturbative approach adopted by Sacks and Uhlenbeck [17], [18]. Sacks and Uhlenbeck define the *α-energy*, for $\alpha > 1$ as the function

$$E_\alpha : \mathrm{Map}(\Sigma, M) \times \mathrm{Met}(\Sigma) \to \mathbb{R}$$

$$\text{given by} \quad E_\alpha(f, h) = \frac{1}{2} \int_\Sigma (1 + |df|_h^2)^\alpha dA_h, \quad (5)$$

where $|df|_h$ and dA_h are calculated with respect to the metric h on Σ.

Unlike the usual energy, the α-energy depends on the choice of Riemannian metric h on Σ, not just the underlying conformal structure. However, if we restrict E_α to $\mathrm{Map}(\Sigma, M) \times \mathrm{Met}_0(\Sigma)$, then just as before this restriction descends to a map on the quotient,

$$E_\alpha : \mathrm{Map}(\Sigma, M) \times \mathcal{T} \longrightarrow \mathbb{R}, \quad (6)$$

and this map approaches $E + (1/2)$ as $\alpha \to 1$. We can therefore regard E_α as a perturbation of E which is defined on exactly the same space. We say that a critical point for E_α is an *α-minimal surface*.

The advantage to the perturbation is that the completion of $\mathrm{Map}(\Sigma, M)$ needed to study E_α via steepest descent is with respect to the $L^{2\alpha}_1$-norm, and this completion is strong enough that $L^{2\alpha}_1(\Sigma, M)$ lies within the space $C^0(\Sigma, M)$ of continuous maps when $\alpha > 1$, the inclusion being a homotopy equivalence. The function (6) is C^2 on the Banach manifold $L^{2\alpha}_1(\Sigma, M) \times \mathcal{T}$, and an argument provided by Sacks and Uhlenbeck shows that the critical points for E_α are C^∞. Moreover, the function

$$E_{\alpha,\omega} : L^{2\alpha}_1(\Sigma, M) \to \mathbb{R}, \quad E_{\alpha,\omega}(f) = E_\alpha(f, \omega), \quad (7)$$

satisfies Condition C of Palais and Smale, and an extension of Morse theory to Banach manifolds [23] allows one to establish Morse inequalities for a small perturbations $E'_{\alpha,\omega}$ of $E_{\alpha,\omega}$ such that all critical points are nondegenerate. (The rational topology of the space $C^0(\Sigma, M)$ can often be estimated by Sullivan's theory of minimal models, and using an argument of Gromov [6], one could show that the number of critical points for $E'_{\alpha,\omega}$ grows exponentially with energy in many cases.)

Using the perturbation $E'_{\alpha,\omega}$, one would hope to obtain partial Morse inequalities for harmonic maps in the limit as $\alpha \to 1$. However, for parametrized minimal surfaces, we need to allow the conformal structure in (6) to vary, and a minimax sequence for E_α might approach the boundary of Teichmüller space $\mathcal{T}$, preventing E_α from satisfying Condition C when Σ has genus at least one. This corresponds to a change in the topology of Σ. We can sometimes simplify matters by restricting E_α to a subspace of $\mathcal{M}(\Sigma, M)$ on which it does satisfy Condition C. The discussion in §4 of [19] provides motivation for one such restriction. We say that a component C of $\mathrm{Map}(T^2, M)$ has *rank two* if

$$f \in C \quad \Rightarrow \quad f_\sharp : \pi_1(T^2) \to \pi_1(M) \text{ maps onto a noncyclic abelian subgroup,}$$

and let $\mathrm{Map}^{(2)}(T^2, M)$ denote the union of all such components.

Note that the mapping class group Γ preserves $\mathrm{Map}^{(2)}(T^2, M)$, so E_α induces a map

$$E_\alpha : \mathcal{M}^{(2)}(T^2, M) \longrightarrow \mathbb{R}, \quad \text{where} \quad \mathcal{M}^{(2)}(T^2, M) = \frac{\mathrm{Map}^{(2)}(T^2, M) \times \mathcal{T}}{\Gamma}.$$

Moreover, if $f \in \mathrm{Map}^{(2)}(T^2, M)$,

$$f \circ \phi = f \quad \text{for some} \quad \phi \in \Gamma \quad \Rightarrow \quad \phi = \text{identity}.$$

Thus the mapping class group $SL(2, \mathbb{Z})$ acts freely on $\mathrm{Map}^{(2)}(T^2, M) \times \mathcal{T}$. Thus if $\mathrm{Map}^{(2)}_\alpha(T^2, M)$ denotes the completion with respect to the $L^{2\alpha}_1$ norm, then $\mathrm{Map}^{(2)}_\alpha(T^2, M)$ will be a smooth Banach manifold.

As we will show in §5, the restricted map

$$E_\alpha : \mathcal{M}^{(2)}_\alpha(T^2, M) \longrightarrow \mathbb{R} \tag{8}$$

does indeed satisfy Condition C: if $[f_i, \omega_i]$ is a sequence of points in $\mathcal{M}^{(2)}(\Sigma, M)$ on which E_α is bounded and for which $\|dE_\alpha([f_i, \omega_i])\| \to 0$, and if for each i, $(f_i, \omega_i) \in \mathrm{Map}(T^2, M) \times \mathcal{T}$ is a representative for $[f_i, \omega_i]$, then there are elements $\phi_i \in \Gamma$ such that a subsequence of $(f_i \circ \phi_i, \phi_i^* \omega_i)$ converges to a critical point for E_α on $\mathrm{Map}(T^2, M) \times \mathcal{T}$. Just as for a generic perturbation of $E_{\alpha,\omega}$, one can establish Morse inequalities for a generic perturbation E'_α of E_α. Even better, one could construct a Morse-Witten chain complex for E'_α, as described implicitly in [11] or more explicitly in [21]. Thus we have a very well-behaved critical point theory for the function $E'_\alpha : \mathcal{M}^{(2)}_\alpha(T^2, M) \to \mathbb{R}$, and we can use it to investigate critical points of E in the limit as $\alpha \to 1$.

Our choice of domain $\mathcal{M}^{(2)}(T^2, M)$ eliminates the necessity to consider branched covers of spheres or degeneration of conformal structure, and we have isolated the difficulty that sequences of critical points for E_α may develop bubbles as $\alpha \to 1$.

A lower bound on distance between minimal tori and minimal two-spheres gives some control on the bubbling process. Choose a bound E_0 on energy, and let

$$\begin{aligned} \mathcal{M}(\Sigma, M)^{E_0} &= \{[f, \omega] \in \mathcal{M}(\Sigma, M) : E([f, \omega]) \leq E_0\}, \\ \mathcal{M}^{(2)}(T^2, M)^{E_0} &= \{[f, \omega] \in \mathcal{M}^{(2)}(T^2, M) : E([f, \omega]) \leq E_0\}. \end{aligned} \tag{9}$$

We can consider the following hypothesis on the union $\mathcal{N}$ of some components of $\mathcal{M}(\Sigma, M)^{E_0}$:

$$\text{Critical points for } E \text{ within } \mathcal{N} \text{ lie at distance} \\ \text{at least } L_0 > 0 \text{ from minimal two-spheres of energy } < E_0. \quad (10)$$

(We believe that this hypothesis holds when M has a generic metric, there exist only finitely many G-orbits of prime parametrized minimal surfaces in $\mathcal{M}(S^2, M)^{E_0}$ and in $\mathcal{M}^{(2)}(T^2, M)^{E_0}$, and the ambient manifold M has dimension at least five.) When this hypothesis is satisfied, bubbling sequences of α_m-minimal surfaces of genus at least one within $\mathcal{N}$ with $\alpha_m \to 1$ must develop necks which have lengths at least L_0, in sharp contrast to the zero-distance bubbling that occurs in limits of sequences of conformal harmonic maps [14]. (In symplectic geometry, zero-distance bubbling is key to understanding moduli spaces of J-holomorphic curves; see §4.7 of [9].)

The main result of this article is that under hypothesis (10), there is a positive constant c (depending on L_0 and a bound on total α-energy) such that when α is sufficiently close to one, bubbles are concentrated within disks of radius

$$\leq r(\alpha, c) = e^{-b(\alpha,c)}, \quad \text{where} \quad b(\alpha, c) = \frac{c}{\sqrt{\alpha - 1}}, \tag{11}$$

the constant c depending on an upper bound on α-energy and a lower bound on the distance to bubbles. Here the radius is measured with respect to the background metric of constant curvature and total area one on Σ. Estimate (11) implies that energy density of minimax α-minimal surfaces must grow at a specific rate within bubbles as $\alpha \to 1$.

On the other hand, if M has finite fundamental group, an estimate of Gromov [6] provides an upper bound on the lengths of necks, which gives an upper bound on energy density within bubbles. Indeed, Gromov argues that if M is a compact Riemannian manifold with finite fundamental group and γ is a geodesic of length $L(\gamma)$ joining two points of M, then there is a constant c depending upon the Riemannian metric such that

$$L(\gamma) \leq c[(\text{Morse index of } \gamma) + 1]. \tag{12}$$

In other words, the Morse index grows at least linearly with length. It follows that if in a sequence $\{f_m\}$ of critical points for E_{α_m} the Morse index is bounded above by a constant λ_1, then the lengths of any necks that form in the limit as $\alpha_m \to 1$ are bounded above by a constant $L_1 = c(\lambda_1 + 1)$.

If M has finite fundamental group and (10) holds, we have both upper and lower bounds on the lengths of necks, and our Main Theorem will yield upper and lower bounds on the energy density within bubbles.

1.2 Statement of the Main Theorem

Before stating the Main Theorem, we discuss the Parker-Wolfson bubble tree [15], which was developed further by Parker [14] and Chen and Tian [3].

Throughout this article, we assume that the Riemannian metric on M has been normalized so that its sectional curvatures are bounded above by one. It then follows from the Gauss equation and the Gauss-Bonnet Theorem that if $g : S^2 \to M$ is minimal,

$$E(g) = \int_\Sigma dA \geq \int_\Sigma K dA = 4\pi, \quad \text{and hence} \quad E(g) \geq 4\pi. \tag{13}$$

Let Σ be a closed oriented surface of genus at least one. We consider a sequence $\{(f_m, \omega_m)\}$ of critical points for E_{α_m} with $m \to \infty$ such that $\alpha_m \to 1$, $\{\omega_m\}$ is bounded, and $f_m : \Sigma \to M$ has bounded energy and bounded Morse index. Under these conditions, $\{(f_m, \omega_m)\}$ has a subsequence (denoted by $\{(f_m, \omega_m)\}$) such that:

1. The sequence $\{\omega_m\}$ converges to an element $\omega_\infty \in \mathcal{T}$.
2. The sequence $\{f_m : \Sigma \to M\}$ converges in C^k for all k on compact subsets of the complement of a finite subset $\{p_1, \ldots, p_l\}$ of "bubble points" in Σ to a map $f_\infty : \Sigma \to M$ which is harmonic for ω_∞. We will check in §3.4 that the map f_∞ is conformal, and hence a parametrized minimal surface.
3. The energy densities $e(f_m)$ converge as distributions to the energy density $e(f_\infty)$ plus a sum of constant multiples of the Dirac delta function,
$$e(f_m) \to e(f_\infty) + \sum_{i=1}^{l} c_i \delta(p_i).$$
4. The restrictions of f_m to a family of suitably rescaled disks centered near the bubble points converge in a suitable sense to a finite family of minimal two-spheres, as explained in more detail below.

It follows from our normalization condition on the metric on M and (13) that the energy c_i lost at each bubble point is at least 4π.

For each m, we divide the Riemann surface Σ into several regions. First, there is a base
$$\Sigma_{0;m} = \Sigma - (D_{1;m} \cup \cdots \cup D_{l;m}),$$
where each $D_{i;m}$ is a metric disk of small radius (which approaches zero as $m \to \infty$) centered at the bubble point $p_i \in \Sigma$. Each disk $D_{i;m}$ is further decomposed into a union $D_{i;m} = A'_{i;m} \cup B'_{i;m}$, where $A'_{i;m}$ is an annular neck region and $B'_{i;m}$ is a bubble region, a smaller disk which is centered at p_i. As explained in [14], we can arrange that the radius of $D_{i;m}$ is $\leq$ (constant)$/m$, the radius of $B'_{i;m}$ is $\leq$ (constant)$/m^3$, and the integral of energy density

$$\int_{A'_{i;m}} e(f_m) dA = C_R,$$

where C_R is a renormalization constant, which we take to be $\leq 2\pi$, small enough to prevent bubbling in $A'_{i;m}$.

We conformally expand $D_{i;m}$ to a disk of unit radius with standard polar coordinates (r, θ), noting that the expansion of $B'_{i;m}$ has radius going to zero like $1/m^2$. Given a ball $\hat{B}$ of radius $1/3$ centered at some point along the circle $r = 1/2$ in the expanded disk, we can apply the ϵ-Regularity Theorem which we will review later (see (27) in § 2.1), or the Main Estimate 3.2 in [17], to show that $r|df| \leq \epsilon_1$ on $\hat{B}$, where ϵ_1 is a constant that can be made arbitrarily small by suitable choice of normalizing constant C_R. Thus $r|df| \leq \epsilon_1$ at any point in the region $1/6 < r < 5/6$. In terms of the coordinates (u, θ). where $e^{-u} = r$, this estimate can be expressed as

$$\left[\left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial \theta}\right)^2\right] < \epsilon_1^2, \quad \text{for} \quad -\log\frac{5}{6} < u < -\log\frac{1}{6}. \tag{14}$$

By a conformal expansion from the disk of radius $1/6$ to the disk of radius $5/6$ we can get a similar estimate on the region $1/30 < r < 5/6$. Continuing in this fashion we get an estimate on the entire annular region $2(\text{radius of } B'_{i;m}) < r < 5/6$. By a slight contraction of the neck, we can redefine neck and bubble so that this estimate holds over the entire neck $A'_{i;m}$.

Estimate (14) implies that we have a bound on the length of each curve $r = (\text{constant})$, and thus we can ensure that each neck is mapped by f_m to a small neighborhood of a smooth curve in M.

Let $B_{i;m}$ be a disk with the same center as $B'_{i;m}$ but with m times the radius and let $A_{i;m} = D_{i;m} - B_{i;m}$. The disk $B_{i;m}$ is then expanded to a disk of radius m by means of the obvious conformal contraction $T_{i;m} : D_m(0) \to B_{i;m}$, where $D_m(0)$ is the disk of radius m in $\mathbb{C}$. A subsequence of

$$g_{i;m} = f_m \circ T_{i;m} : D_m(0) \to M, \qquad m = 1, 2, \ldots$$

then converges uniformly in C^k on compact subsets of $\mathbb{C}$, or on compact subsets of $\mathbb{C} - \{p_{i,1}, \ldots, p_{i,l_i}\}$, where $\{p_{i,1}, \ldots, p_{i,l_i}\}$ is a finite set of new bubble points, to a harmonic map of bounded energy. The limit extends to a harmonic map of the two-sphere, by the Sacks-Uhlenbeck removeable singularity theorem [17], Theorem 3.6.

There may exist new bubble points $\{p_{i,1}, \ldots, p_{i,l_i}\}$ in $\mathbb{C} = S^2 - \{\infty\}$, in which case the process can be repeated. Around each new bubble point $p_{i,j}$, we construct a small disk $D_{i,j;m}$ which is further subdivided into an annular region $A_{i,j;m}$ and a smaller disk $B_{i,j;m}$ on which bubbling will occur. We construct a conformal contraction $T_{i,j;m} : D_m(0) \to B_{i,j;m}$ and a subsequence of

$$g_{i;j;m} = f_m \circ T_{i,j;m} : D_m(0) \to M, \qquad m = 1, 2, \ldots$$

will converge once again to a harmonic two-sphere with one or several punctures, which can be filled in as before. Thus for each f_m in the sequence, we may have several *level-one* bubble regions $B_{i;m}$, each of which may contain several *level-two* bubble regions $B_{i,j;m}$. This can be repeated: each of the level-two bubble regions may contain several *level-three* bubble regions, and so forth, but the process terminates after finitely many steps, yielding what Parker and

Wolfson call a *bubble tree*, the vertices being harmonic maps (a base minimal surface f_∞ and several minimal two-spheres) and the edges corresponding to the parametrized maps from the annular regions into M.

Some of the harmonic two-spheres obtained by this process may have zero energy (all their energy bubbles away) in which case they are called *ghost bubbles*, but ghost bubbles always have at least one bubble point in addition to ∞. Moreover, if a ghost bubble has only one bubble point in addition to ∞, it can be eliminated and the two adjoining annuli can be amalgamated into one. This might be done many times, depending on our choice of renormalization constant. We assume that this process of amalgamation has been carried through, so each ghost bubble has at least two bubble points in addition to ∞.

Thus we find that for each sufficiently large $m \in \mathbb{N}$, Σ is a disjoint union of the base $\Sigma_{0;m}$, annular regions $A_{i_1,\ldots,i_k;m}$, also called *necks*, and bubble regions with disks around bubble points of higher level deleted,

$$B_{i_1,\ldots,i_k;m} - \bigcup_j D_{i_1,\ldots,i_k,j;m}. \tag{15}$$

The annular regions and the regions (15) corresponding to ghost bubbles comprise the *thin part* of the map f_m, while the base and the regions (15) corresponding to nonconstant harmonic two-spheres make up the *thick part*.

Since the bubble tree is finite, after passing to a subsequence, we can arrange that each level k, each of the sequences

$$g_{i_1,\ldots,i_k;m} = f_m \circ T_{i_1,\ldots,i_k;m} : D_m(0) \to M, \qquad m = 1, 2, \ldots \tag{16}$$

converges uniformly on compact subsets of $\mathbb{C}$ minus a finite number of points to a limiting harmonic two-sphere $g_{i_1,\ldots,i_k}$. By (13), nonconstant harmonic two-spheres always have energy at least 4π. Thus a bound on the energy gives a bound on the number of leaves in the bubble tree, and hence a bound on the total number of edges in the bubble tree.

The restrictions of f_m to $A_{i_1,\ldots,i_k;m} = D_{i_1,\ldots,i_k;m} - B_{i_1,\ldots,i_k;m}$ yield the *neck maps*. For each multi-index $(i_1, \ldots, i_k)$, let

$$\epsilon_{i_1,\ldots,i_k;m} = (\text{radius of } D_{i_1,\ldots,i_k;m}), \quad \eta_{i_1,\ldots,i_k;m} = (\text{radius of } B_{i_1,\ldots,i_k;m}),$$

the radii being measured with respect to the canonical constant curvature metric of total area one on Σ. Choose normal coordinates (x, y) about the center of $D_{i_1,\ldots,i_k;m}$, let (r, θ) denote the corresponding polar coordinates, and let $u = -\log r$. If Σ is a torus, the canonical metric on $D_{i_1,\ldots,i_k;m}$ can be expressed as

$$ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\theta^2 = e^{-2u}(du^2 + d\theta^2),$$

while for general choice of Σ, this expression holds approximately to a high degree of precision. In terms of the coordinates, we can define a map

$$\psi_{i_1,\ldots,i_k;m} : [a, b] \times S^1 \longrightarrow A_{i_1,\ldots,i_k;m} \quad \text{by} \quad \psi(u, \theta) = e^{-u+i\theta}$$
$$\text{where} \quad a = -\log \epsilon_{i_1,\ldots,i_k;m} \quad \text{and} \quad b = -\log \eta_{i_1,\ldots,i_k;m}.$$

The restrictions of f_m to $A_{i_1,\ldots,i_k;m}$ can then be reparametrized to yield maps on conformally long cylinders,

$$h'_{i_1,\ldots,i_k;m} = f_m \circ \psi_{i_1,\ldots,i_k;m} : [a,b] \times S^1 \longrightarrow M.$$

These can then be linearly rescaled to yield maps $h_{i_1,\ldots,i_k;m}$ on $[0,1] \times S^1$ which converge to geodesics $\gamma_{i_1,\ldots,i_k}$ in M, the endpoints lying in either the base or one of the bubbles.

The minimal base map f_∞, the harmonic two-spheres $g_{i_1,\ldots,i_k}$ and the geodesics $\gamma_{i_1,\ldots,i_k}$ comprise a harmonic map from a *stratified Riemann surface* into M, as described by Chen and Tian [3].

Main Theorem. *Let M be a compact Riemannian manifold which satisfies the condition that all sectional curvatures are ≤ 1 and let Σ be a closed surfaces of genus $g \geq 1$. Suppose that $\{(f_m, \omega_m)\}$ is a sequence of critical points for E_{α_m}, converging to a bubble tree as described above. Suppose that $m \mapsto B_{i_1,\ldots,i_k;m}$ is a sequence of bubble disks of radius $\epsilon_{i_1,\ldots,i_k;m}$ such that the energy of the restriction of f_m to $B_{i_1,\ldots,i_k;m}$ approaches E_b, and that $m \mapsto A_{i_1,\ldots,i_k;m}$ is the corresponding sequence of annuli such that the reparametrizations of f_m on the $A_{i_1,\ldots,i_k;m}$'s converge to a geodesic $\gamma_{i_1,\ldots,i_k}$ of nonzero length $L_{i_1,\ldots,i_k}$. Then*

$$\lim_{m\to\infty} \frac{\eta_{i_1,\ldots,i_k;m}/\epsilon_{i_1,\ldots,i_k;m}}{e^{-c_0/\sqrt{\alpha_m - 1}}} = 1, \quad \text{where} \quad c_0 = \sqrt{\frac{\pi}{2}} \frac{L_{i_1,\ldots,i_k}}{\sqrt{E_b}}. \tag{17}$$

Remark 1. We will sometimes refer to this theorem as the Scaling Theorem. We can express (17) somewhat roughly as

$$\frac{\eta_{i_1,\ldots,i_k;m}}{\epsilon_{i_1,\ldots,i_k;m}} \sim e^{-c_0/\sqrt{\alpha_m - 1}}, \quad \text{where} \quad c_0 = \sqrt{\frac{\pi}{2}} \frac{L_{i_1,\ldots,i_k}}{\sqrt{E_b}}.$$

In the case where a single bubble forms, the average energy density within the bubble is approximately

$$\frac{E_b}{\pi\eta_1^2} \sim \frac{E_b}{\pi} e^{2c_0/\sqrt{\alpha_m - 1}}, \quad \text{with} \quad c_0 = \sqrt{\frac{\pi}{2}} \frac{L}{\sqrt{E_b}},$$

where L is the length of the neck between base and bubble and E_b is the energy of the bubble. Thus the Scaling Theorem gives an explicit estimate on how rapidly the energy density within bubbles must grow as $\alpha \to 1$.

Remark 2. One can formulate a version of the Main Theorem for sequences $\{f_m\}$ of critical points for $E_{\alpha_m \omega}$, where ω is a fixed element of the Teichmüller space $\mathcal{T}$. In this case the map on the base is harmonic but not necessarily conformal. The proof is virtually identical to the one we give below.

To put the Main Theorem in context, it is helpful to consider how dilations influence the α-energy. Consider a map from the η-disk,

$$f : D_\eta \to \mathbb{R}^N, \quad \text{where} \quad D_\eta = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq \eta^2\},$$

which takes the boundary ∂D_η to a point. We can expand this to a map on the unit disk,

$$f_\eta : D_1 \to \mathbb{R}^N, \qquad f_\eta(x,y) = f(\eta x, \eta y).$$

Since $|df_\eta(x,y)| = \eta|df(\eta x, \eta y)|$,

$$\int_{D_\eta} |df|^{2\alpha} dxdy = \left(\frac{1}{\eta}\right)^{2(\alpha-1)} \int_{D_1} |df_\eta|^{2\alpha} dxdy. \tag{18}$$

It follows from (18) that as a nonconstant map from the disk $f : D_1 \to \mathbb{R}^N$ is rescaled to a disk of radius η, with $\eta \to 0$, the highest order term in the α-energy increases like

$$(1/\eta)^{2(\alpha-1)} = e^{-2(\alpha-1)\log\eta} \quad \text{as } \alpha \to 1.$$

In order to keep the α-energy approximately constant in the rescaled ball as $\alpha \to 1$, we would need to rescale by

$$\eta(\alpha) = e^{-c_0/(\alpha-1)}, \quad \text{so} \quad -\log\eta(\alpha) = \frac{c_0}{\alpha - 1} \tag{19}$$

where c_0 is a positive constant. Note that $(-\log\eta)/2\pi$ can be regarded as the conformal invariant of the annulus $D_1 - D_\epsilon$.

When there is a positive lower bound on the distance between base and bubbles (10), such as when we consider the function E_α on $\mathcal{N} = \mathcal{M}^{(2)}(T^2, M)^{E_0}$ as described before, it takes some energy to construct an α-energy parametrization of a geodesic connecting bubble to base, and critical points for α-energy must achieve a balance between energy in the bubbles and energy in the necks. Achieving this balance forces the estimates on energy density given by the Main Theorem. Note that the radius of the bubble region is actually somewhat larger than the radius (19) which would keep energy in the bubble constant.

1.3 Contents of the remainder of the article

Let us give a brief overview of the remainder of the article. In §2, we review techniques used for studying critical points for E_α and describe a replacement procedure for α-harmonic maps that we hope will be of some independent interest. Recall that the Morse theory of smooth closed geodesics can be treated in two quite different ways: either via calculus on infinite-dimensional manifolds or via a finite-dimensional approximation with broken geodesics (as described in § 16 of [10]). Our replacement procedure enables us to combine the best features of the two approaches when studying the perturbed energy E_α.

To understand the direct limit of the Morse-Witten complexes of the perturbation E'_α of E_α as $\alpha \to 1$, we need to understand the relationship between the length of a neck image within M and the conformal parameter of the corresponding annular domain. We do this in §3.3, using the replacement procedure mentioned above, and obtain additional results on the structure of the bubble tree. For example, we note that the estimate on neck energy together with Gromov's estimate (12) implies that no energy is lost in the necks of minimax sequences of α-energy critical points with $\alpha \to 1$, under the assumption that $\pi_1(M)$ is finite. (Chen and Tian prove this for sequences of α-energy minimizing critical points in [3].) We also check that the base harmonic map $f_\infty : \Sigma \to M$ in the bubble tree is conformal, and hence minimal, when each f_m is a critical point for E_{α_m}.

The proof of the Main Theorem is presented in §4. It is relatively easy to prove the corresponding statement for the function

$$E^0_{\alpha,\omega} : L^{2\alpha}_1(\Sigma, M) \to M, \qquad E^0_{\alpha,\omega}(f) = \frac{1}{2}\int_\Sigma |df|^{2\alpha} dA,$$

because it has nicer scaling properties. (One can check that this function satisfies Condition C, but its critical points are not necessarily smooth.) After establishing the easier assertion it turns out to be relatively straightforward to establish the general case by Taylor series approximation.

Finally, in a brief appendix (§**??**) we return to the discussion which we began in §1.1, and complete the argument that $E'_\alpha : \mathcal{M}^{(2)}(T^2, M)^{E_0} \to \mathbb{R}$ satisfies Condition C. When E_0 is sufficiently small that at most one bubble can form, we show that the critical locus for E'_α divides into two disjoint sets, when α is sufficiently close to one. Although there is no reason to suspect that these disjoint sets are invariant under the flow for an arbitrary pseudogradient for $-E'_\alpha$, we do have considerable flexibility in choice of pseudogradient or of "gradient-like" vector field.

It is our intention to exploit this flexibility when we apply the Main Theorem in a sequel.

2 Local stability and replacement

The goal of this section is to provide an extension of a local stability result of Jäger and Kaul [7] from harmonic to α-harmonic maps, and describe its application to a replacement procedure for α-harmonic maps.

2.1 Background on α-harmonic maps

Suppose that h is a Riemannian metric on the compact connected surface Σ. Critical points of the function

$$E_{\alpha,h} : \mathrm{Map}(\Sigma, M) \longrightarrow \mathbb{R}, \qquad E_{\alpha,h}(f) = \frac{1}{2}\int_\Sigma (1 + |df|^2_h)^\alpha dA_h,$$

are called *α-harmonic maps* (or *(α, h)-harmonic maps* if we want to emphasize the role of the metric. For the convenience of the reader, we begin this section with a few preliminary remarks about α-harmonic maps, referring to [17], [18] or further background results, and to [4] for the extension to the case where Σ has a boundary constrained to lie in a smooth submanifold of M. The case in which Σ has boundary fixed (Dirichlet boundary conditions) can be treated in much the same way. For simplicity of notation, we will often suppress the notation for the metric on Σ and write E_α for $E_{\alpha,h}$.

It is convenient to regard M as isometrically imbedded in some Euclidean space $\mathbb{R}^N$ of large dimension, always possible by the Nash embedding theorem. If Σ is a compact connected surface, possibly with boundary $\partial\Sigma$ consisting of several circles, and $p \geq 2$, we let $L^p_k(\Sigma, \mathbb{R}^N)$ denote the completion of the space of smooth maps from Σ to $\mathbb{R}^N$ with respect to the Sobolev L^p_k norm. The space $L^p_k(\Sigma, \mathbb{R}^N)$ is always a Banach space, and a Hilbert space if $p = 2$. If $k \geq 2$, or $p > 2$ and $k \geq 1$ and the boundary of Σ is empty, then

$$L^p_k(\Sigma, M) = \{f \in L^p_k(\Sigma, \mathbb{R}^N) : f(q) \in M \text{ for all } q \in \Sigma\},$$

is an infinite-dimensional smooth submanifold. More generally, if Σ has a boundary $\partial\Sigma$ and $f_0 : \partial\Sigma \to M$ is a fixed smooth map, we set

$$L^p_{k,0}(\Sigma, M) = \{f \in L^p_k(\Sigma, M : f|\partial\Sigma = f_0\}$$

which is also an infinite-dimensional smooth submanifold. Note that the map

$$E_{\alpha,h} : L^p_{k,0}(\Sigma, M) \longrightarrow \mathbb{R},$$

which is only C^2 when $p = 2\alpha$ and $k = 1$, is actually C^∞, when k is sufficiently large. To see this, regard $E_{\alpha,h}$ as a composition of several maps

$$f \mapsto df \mapsto (1 + |df|^2_h)^\alpha \mapsto \frac{1}{2}\int_\Sigma (1 + |df|^2_h)^\alpha dA_h.$$

When k is large, the first map is smooth into L^p_{k-1}, the second is smooth by the well-known ω-Lemma, while the third is always smooth since integration is continuous and linear.

For a fixed choice of metric h on Σ, we can differentiate to find the first variation of $E_{\alpha,h}$, obtaining

$$dE_{\alpha,h}(f)(Y) = \int_\Sigma \langle \mathcal{X}(f), Y\rangle dA, \quad \text{for} \quad Y \in T_f L^p_{k,0}(\Sigma, M),$$

where $f \mapsto \mathcal{X}(f)$ is the Euler-Lagrange operator, which depends smoothly on f, as well as on the metrics $\langle\cdot,\cdot\rangle$ on the ambient manifold M and h on Σ. We can think of $\mathcal{X}$ as an "L^2 gradient of $E_{\alpha,h}$." If (x, y) are isothermal parameters on Σ so that the Riemannian metric on Σ takes the form $\lambda^2(dx^2 + dy^2)$, a calculation shows that the Euler-Lagrange operator is given by the explicit formula

$$\mathcal{X}(f) = -\frac{\alpha}{\lambda^2}\left(\frac{\partial}{\partial x}\left(\mu^{2(\alpha-1)}\frac{\partial f}{\partial x}\right)\right)^\top - \frac{\alpha}{\lambda^2}\left(\frac{\partial}{\partial y}\left(\mu^{2(\alpha-1)}\frac{\partial f}{\partial y}\right)\right)^\top, \tag{20}$$

where $\mu^2 = 1 + |df|^2$ and $(\cdot)^\top$ denotes orthogonal projection into the tangent space to the submanifold M of $\mathbb{R}^N$. Alternatively, we can write

$$\mathcal{X}(f) = -\frac{\alpha}{\lambda^2}\frac{D^g}{\partial x}\left(\mu^{2(\alpha-1)}\frac{\partial f}{\partial x}\right) - \frac{\alpha}{\lambda^2}\frac{D^g}{\partial y}\left(\mu^{2(\alpha-1)}\frac{\partial f}{\partial y}\right). \tag{21}$$

where D^g is the covariant derivative for the Levi-Civita connection for the metric $g = \langle \cdot, \cdot \rangle$ on the ambient manifold M. Differentiating once again yields the second variation of E_α,

$$d^2 E_\alpha(f)(X,Y) = \int_\Sigma \langle L(X), Y\rangle dA, \quad \text{for} \quad X, Y \in T_f L^p_{k,0}(\Sigma, M), \tag{22}$$

where L is the Jacobi operator, a second-order formally self-adjoint elliptic operator.

Just as harmonic maps satisfy the unique continuation property (as proven for example in [16]), so do α-harmonic maps. To see this, we expand the right-hand side of (20) and write the Euler-Lagrange equation $F = 0$ as

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = A(f)(df, df) - (\alpha - 1)\left(\frac{\partial}{\partial x}(\log \mu^2)\frac{\partial f}{\partial x} + \frac{\partial}{\partial y}(\log \mu^2)\frac{\partial f}{\partial y}\right), \tag{23}$$

an equation in which $A(f)(df, df)$ stands for a certain expression in terms of the second fundamental form $A(f)$ of f. We can differentiate the logarithm to obtain

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = A(f)(df, df) - Q_f(f),$$

where Q_f is the nonlinear differential operator defined by

$$\begin{aligned} Q_f(u) = \frac{\alpha - 1}{1 + |df|^2} & \left(\left\langle \frac{\partial^2 u}{\partial x^2}, \frac{\partial f}{\partial x}\right\rangle + \left\langle \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial f}{\partial y}\right\rangle\right)\frac{\partial f}{\partial x} \\ & + \left(\left\langle \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial f}{\partial x}\right\rangle + \left\langle \frac{\partial^2 u}{\partial y^2}, \frac{\partial f}{\partial y}\right\rangle\right)\frac{\partial f}{\partial y}. \end{aligned}$$

We can write this more simply as $L_f(f) = A(f)(df, df)$, where

$$L_f(u) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + Q_f(u),$$

a uniformly elliptic operator when α is sufficiently close to one, whose coefficients depend on f and df. We observe that the values of Q_f are tangent to f, in the sense that at any point of Σ, they are linear combinations of $\partial f/\partial x$ and $\partial f/\partial y$, evaluated at that point. Given $\epsilon > 0$, we can choose an α_0 sufficiently close to one that for $\alpha \in (1, \alpha_0]$,

$$(1 - \epsilon)L_f(u) \leq \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \leq (1 + \epsilon)L_f(u). \tag{24}$$

When α is sufficiently close to one, (24) and complexification yields the estimate

$$\left|\frac{\partial^2 f}{\partial z \partial \bar{z}}\right| \leq K\left(\left|\frac{\partial f}{\partial z}\right| + |f|\right),$$

where $z = x + iy$ and K is a constant.

But this is just the estimate we need for the Lemma of Hartman and Wintner proven in [8], §2.6. The above estimate and the Hartman-Wintner Lemma imply that if the complex coordinate z is centered at a point $p \in \Sigma$,

$$|f(z)| = o(|z|^n) \quad \Rightarrow \quad \lim_{z \to 0} \frac{\partial f}{\partial z} z^{-n} \quad \text{exists}, \tag{25}$$

and if $|f(z)| = o(|z|^n)$ for all n, then f is constant. Thus an α-harmonic map which is an immersion on a dense set of points can drop rank at only isolated points, and the resulting singularity looks like a branch point to lowest order.

Using (24), we can modify the argument presented by Sampson [16] to prove unique continuation:

Unique Continuation Lemma. *If $\alpha_0 > 1$ is sufficiently close to one and $\alpha \in (1, \alpha_0]$, then any two α-harmonic maps from a connected surface Σ into M which agree on an open set must agree identically.*

This Lemma could also be obtained from the estimate (24) and the unique continuation theorem of Aronsjazn [1].

Remark. Following the proof of Theorem 3 in [16], we note that the Unique Continuation Lemma implies that if df has rank zero on a nonempty open set, the harmonic map f must be constant. We now ask what happens if df has rank one on a nonempty open set $U \subset \Sigma$. In this case, every point of U has an open neighborhood which is mapped by f onto a smooth arc C in M. We can suppose that coordinates (u, θ) have been constructed on U so that $\partial f / \partial \theta = 0$, and thus $f : U \to M$ reduces to a function of one variable, $f(u, \theta) = f_0(u)$, parametrizing C. The variational equation for f_0 is an ordinary differential equation

$$\frac{D}{du} \left(\mu^{2(\alpha-1)} \frac{df_0}{du} \right) = 0, \qquad \mu^2 = 1 + \left(\frac{df_0}{du} \right)^2,$$

where D is the covariant derivative defined by the Levi-Civita connection. If we let $(\cdot)^\perp$ denote orthogonal projection to the component normal to C, we find that

$$\left(\frac{D}{du} \left(\mu^{2(\alpha-1)} \frac{df_0}{du} \right) \right)^\perp = \mu^{2(\alpha-1)} \left(\frac{D}{du} \left(\frac{df_0}{du} \right) \right)^\perp = 0,$$

which implies that C must be a geodesic arc. Thus we can regard f_0 as a composition $f_0 = \gamma \circ \phi$ where $\gamma : \mathbb{R} \to M$ is a unit-speed geodesic and $\phi : (a, b) \to \mathbb{R}$. It remains only to determine the parametrization ϕ. The variational equation for f_0 now implies that ϕ must satisfy the ordinary differential equation

$$\mu^{2(\alpha-1)} \frac{d\phi}{du} = a, \tag{26}$$

where a is a constant of integration. The solutions to this equation depend on the metric h on Σ. Such α-harmonic parametrizations of geodesics play an important role n approximating necks between bubbles.

Using the estimate (24), we can generalize several other theorems from the theory of harmonic maps to the case of α-harmonic maps. For example:

Bochner Lemma. *For each $\alpha > 1$, there is a constant c_α depending continuously on α and a second-order elliptic operator L_α whose coefficients depend continuously on df and α such that*

1. *$c_\alpha \to 1$ and $L_\alpha \to \Delta$ as $\alpha \to 1$, where Δ is the usual Laplace operator, and*
2. *if $f : \Sigma \to M$ is a nonconstant α-harmonic map, then*

$$\frac{1}{2} L_\alpha(|df|^2) \geq c_\alpha |\nabla df|^2 + K|df|^2 - |R_{1212}||df|^4,$$

where K is the Gaussian curvature of the Riemannian metric on Σ and R_{1212} is the sectional curvature of the two-plane in M spanned by $f_*(T\Sigma)$.

We can use this lemma just as in the case of harmonic maps (see [16]) to prove the following result essentially due to Sacks and Uhlenbeck [17]:

ϵ-Regularity Theorem. *Let M be a compact Riemannian manifold whose sectional curvatures $K(\sigma)$ satisfy the estimate $K(\sigma) \leq 1$. Then there exists an $\alpha_0 > 1$ with the following property: If $f : D_r \to M$ is an α-harmonic map, where $1 \leq \alpha \leq \alpha_0$ and D_r is the disk of radius r in the complex plane with the standard Euclidean metric ds^2, then there exists $\epsilon > 0$ such that*

$$\int_{D_r} e(f) dA < \epsilon \qquad \Rightarrow \qquad \max_{\sigma \in (0,r]} \sigma^2 \sup_{D_{r-\sigma}} e(f) < 4, \tag{27}$$

where $e(f) = (1/2)|df|^2$ is the energy density, which is calculated with respect to the metric ds^2 on D_r and the Riemannian metric on M.

2.2 The Replacement Theorem

To study stability of α-harmonic maps, we need an explicit formula for the second variation of the α-energy E_α at a critical point f:

$$\begin{aligned} d^2 E_\alpha(f)(X,Y) = \alpha \int_\Sigma (1 + |df|^2)^{\alpha-1} [\langle \nabla X, \nabla Y \rangle - \langle \mathcal{K}(X), Y \rangle] dA \\ + 2\alpha(\alpha - 1) \int_\Sigma (1 + |df|^2)^{\alpha-2} \langle df, \nabla X \rangle \langle df, \nabla Y \rangle dA, \end{aligned} \tag{28}$$

for $X, Y \in T_f L^p_k(\Sigma, M)$. Here $\mathcal{K}(X)$ is defined by

$$\langle \mathcal{K}(X), Y \rangle = \frac{1}{\lambda^2} \left[\left\langle R\left(X, \frac{\partial f}{\partial x}\right) \frac{\partial f}{\partial x}, Y \right\rangle + \left\langle R\left(X, \frac{\partial f}{\partial y}\right) \frac{\partial f}{\partial y}, Y \right\rangle \right],$$

R being the Riemann-Christoffel curvature tensor of M. The factor of $\alpha - 1$ in the second term of (28) implies that the first of the two terms in the index formula dominates when α is close to one. Note also that the second term is positive semidefinite. As $\alpha \to 1$, the second variation of the α-energy approaches the familiar second variation for the usual energy. An integration by parts in (28) yields the Jacobi operator L which appears in (22).

Recall that if the metric on Σ is $h = \lambda^2(dx^2+dy^2)$, we can write the standard Laplace operator as

$$\Delta = \frac{1}{\lambda^2}\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right].$$

Moreover, if $h : M \to \mathbb{R}$ is a smooth function, we define its second covariant derivative by the formula

$$(\nabla^2 h)(X,Y) = X(Yh)) - (\nabla_X Y)(h),$$

where ∇ is the Levi-Civita connection. This induces a symmetric bilinear form

$$\nabla^2 h : T_pM \times T_pM \longrightarrow \mathbb{R},$$

for each $p \in M$. It follows from the chain rule that if $f : T^2 \to \mathbb{R}$ is a smooth map,

$$\Delta(h \circ f) = \frac{1}{\lambda^2}\left[\nabla^2 h\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial x}\right) + \nabla^2 h\left(\frac{\partial f}{\partial y}, \frac{\partial f}{\partial y}\right)\right] + dh(\tau(f)),$$

where $\tau(f)$ is the tension of the map f, defined by

$$\tau(f) = \frac{1}{\lambda^2}\left[\frac{D}{\partial x}\frac{\partial f}{\partial x} + \frac{D}{\partial y}\frac{\partial f}{\partial y}\right].$$

The tension vanishes for harmonic maps, while in the case of an α-harmonic map f, it follows from the Euler-Lagrange equations that the tension is given by

$$\tau(f) = -\frac{\alpha-1}{\lambda^2}\left[\frac{\partial}{\partial x}(\log \mu^2)\frac{\partial f}{\partial x} + \frac{\partial}{\partial y}(\log \mu^2)\frac{\partial f}{\partial y}\right],$$

where $\mu^2 = (1 + |df|^2)$. For α-harmonic maps, it is convenient to replace the Laplace operator by the operator

$$L = \frac{1}{\lambda^2}\frac{1}{\mu^{2\alpha-2}}\left[\frac{\partial}{\partial x}\left(\mu^{2\alpha-2}\frac{\partial}{\partial x}\right) + \frac{\partial}{\partial y}\left(\mu^{2\alpha-2}\frac{\partial}{\partial y}\right)\right].$$

This introduces a new term which exactly cancels the tension field when f is α-harmonic, leaving

$$L(h \circ f) = \frac{1}{\lambda^2}\left[\nabla^2 h\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial x}\right) + \nabla^2 h\left(\frac{\partial f}{\partial y}, \frac{\partial f}{\partial y}\right)\right]. \qquad (29)$$

Given a point $p \in M$, let $D_\rho(p)$ denote the open geodesic ball of radius ρ about p, where ρ is chosen so that $\rho < \pi/2\kappa$, where κ^2 is an upper bound on the setional curvature of M and $D_\rho(p)$ is disjoint from the cut locus of p. We can then define a smooth map

$$h : D_\rho(p) \longrightarrow \mathbb{R} \qquad \text{by} \qquad h(q) = \frac{1 - \cos(\kappa d(p,q))}{\kappa^2},$$

where d is the distance function on M. By comparison with the metric of constant curvature κ, Jäger and Kaul [7] show that

$$\nabla^2 h(q)(v,v) \geq \cos(\kappa d(p,q))|v|^2.$$

It follows that if U is an open subset of Σ, $f : U \to D_\rho(p)$ is an α-harmonic map, and

$$\phi(x) = \cos(\kappa d(p, f(x))), \qquad \text{for } x \in U, \tag{30}$$

then

$$L(\phi) \geq -\kappa^2 |df|^2 \phi. \tag{31}$$

Stability Lemma. *Let U be a domain in the Riemann surface Σ, and let $f : U \to M$ be an α-harmonic map such that $f(U) \subset D_\rho(p)$, where $D_\rho(p)$ is a geodesic ball of radius ρ about $p \in M$ disjoint from the cut locus of p and $\rho < \pi/(2\kappa)$, where κ^2 is an upper bound for the sectional curvature on M. Then f is stable.*

Proof: We modify the proof of Theorem B from [7] following the presentation in §2.2 of [8]. Suppose that X is a section of f^*TM which satisfies the elliptic equation

$$L(X) + \frac{1}{\lambda^2}\left[R\left(X, \frac{\partial f}{\partial x}\right)\frac{\partial f}{\partial x} + R\left(X, \frac{\partial f}{\partial y}\right)\frac{\partial f}{\partial y}\right] = 0, \tag{32}$$

where

$$L = \frac{1}{\lambda^2}\frac{1}{\mu^{2\alpha-2}}\left[\frac{D}{\partial x} \circ \left(\mu^{2\alpha-2}\frac{D}{\partial x}\right) + \frac{D}{\partial y} \circ \left(\mu^{2\alpha-2}\frac{D}{\partial y}\right)\right]. \tag{33}$$

Assuming without loss of generality that κ is positive, we define a map $\theta : U \to \mathbb{R}$ by

$$\theta(x) = \frac{|X(x)|^2}{\phi(x)^2},$$

where ϕ is defined by (30). Our first objective will be to show that this function satisfies the maximum principle.

Indeed, setting $\psi(x) = |X(x)|^2$, we find that

$$\theta = \frac{\psi}{\phi^2} \quad \Rightarrow \quad \nabla(\log\theta) = \frac{\nabla\psi}{\psi} - 2\frac{\nabla\phi}{\phi},$$

and hence

$$L(\log\theta) = \frac{L\psi}{\psi} - 2\frac{L\phi}{\phi} - \left[\left(\frac{|\nabla\psi|}{\psi}\right)^2 - 2\left(\frac{|\nabla\phi|}{\phi}\right)^2\right]. \tag{34}$$

It follows from the chain rule that

$$L(\psi) = \lambda^{-2}\mu^{2-2\alpha}\nabla\left(\mu^{2\alpha-2}\nabla(|X|^2)\right) = 2|\nabla X|^2 + 2\langle L(X), X\rangle,$$

and hence by (33),

$$L(\psi) = 2|\nabla X|^2 - \frac{2}{\lambda^2}\left\langle R\left(X, \frac{\partial f}{\partial x}\right)\frac{\partial f}{\partial x} + R\left(X, \frac{\partial f}{\partial y}\right)\frac{\partial f}{\partial y}, X\right\rangle.$$

Since

$$|\nabla\psi| = 2|\langle\nabla X, X\rangle| \leq 2|\nabla X||X| = 2|\nabla X|\sqrt{\psi} \quad \Rightarrow \quad \frac{|\nabla\psi|^2}{\psi} \leq 4|\nabla X|^2$$

and the sectional curvatures are bounded by κ^2, we obtain

$$\frac{L(\psi)}{\psi} \geq \frac{|\nabla\psi|^2}{2\psi^2} - 2\kappa^2|df|^2.$$

On the other hand, it follows from equation (31) that

$$-2\frac{L(\phi)}{\phi} \geq 2\kappa^2|df|^2.$$

Substituting into (34), we obtain

$$L(\log\theta) \geq -\left(\frac{|\nabla\psi|^2}{2\psi^2}\right) + 2\left(\frac{|\nabla\phi|^2}{\phi}\right).$$

Setting $\xi(x) = (1/2)[(\nabla\psi/\psi) + 2(\nabla\phi/\phi)]$, we obtain

$$L(\log\theta) + \xi(x)\cdot\nabla(\log\theta) \geq 0,$$

and by Hopf's maximum principle, we conclude that θ cannot assume a positive local maximum in the interior. Thus if X vanishes on the boundary of U, then θ and hence X must be identically zero.

Let $\Gamma_0(f^*TM)$ denote the space of smooth sections of f^*TM which vanish on the boundary of U. The above argument shows that the bilinear form

$$I : \Gamma_0(f^*TM) \times \Gamma_0(f^*TM) \longrightarrow \mathbb{R}$$

defined by

$$I(X, X) = \int_{T^2}(1 + |df|^2)^{\alpha-1}[|\nabla X|^2 - \langle\mathcal{K}(X), X\rangle]dA$$

is positive definite. Indeed, if it were not positive definite, there would be a solution to the elliptic equation $L(X) = 0$ which vanished on the boundary of some proper subdomain $U_1 \subset U$ by Smale's version of the Morse index theorem [22], contradicting the maximum principle for the function θ considered in the preceding paragraph. Thus I is positive definite and since it follows from (28) that the Hessian $d^2E_\alpha(f)$ for the α-energy satisfies $d^2E_\alpha(f) \geq I$, we see that it is also positive definite. Hence f is stable, establishing the Stability Lemma.

The following Replacement Theorem guarantees that we can replace any α-harmonic disk bounded by a curve Γ lying in a small normal coordinate neighborhood by a disk which minimizes α-energy:

Replacement Theorem. *Let U be a domain in the compact Riemann surface Σ with a smooth boundary ∂U consisting of a finite number of circles, and let $f_1, f_2 : U \cup \partial U \to M$ be two α-harmonic maps such that $f_i(U) \subset D_\rho(p)$, where $D_\rho(p)$ is a geodesic ball of radius ρ about $p \in M$ disjoint from the cut locus of p and $\rho < \pi/(2\kappa)$, where κ^2 is an upper bound for the sectional curvature on M. Then $f_1|\partial U = f_2|\partial U \Rightarrow f_1 = f_2$. Moreover, the unique α-harmonic f with given boundary values depends continuously on the boundary values and on the metric g on M.*

Proof: We can assume that $f_1(U)$ and $f_2(U)$ are contained in $D_{\rho-\epsilon}(p)$, for some $\epsilon > 0$. We can replace the Riemannian metric ds^2 on $D_\rho(p)$ by $\sigma(r)ds^2$, where r is the radial coordinate on $D_\rho(p)$ and $\sigma : [0, \rho) \to \mathbb{R}$ is a smooth function which is identically one for $r \leq \rho - \epsilon$, and goes off to infinity as $r \to \rho$ so fast that $(D_\rho(p), \sigma ds^2)$ is complete. Of course, f_1 and f_2 are still α-harmonic maps into the Riemannian manifold $(D_\rho(p), \sigma ds^2)$.

We claim that if $\sigma : [0, \rho) \to \mathbb{R}$ is chosen appropriately, no α-harmonic map into $(D_\rho(p), \sigma ds^2)$ with the same boundary as f_1 and f_2 can actually penetrate the region A defined by the inequalities $\rho - \epsilon < r < \rho$. Indeed, we can arrange that

$$\nabla^2\sigma(X, X) > \|X\|^2, \quad \text{for } X \in T_qM,\ q \in A.$$

If f is such an α-harmonic map into $(D_\rho(p), \sigma ds^2)$ which is tangent to one of the hyperspheres $r =$ constant within A at $f(q)$, it follows from (29) that $L(\sigma \circ f) \geq 0$ on a neighborhood of p. Noting that f is an immersion at almost all points, we quickly obtain a contradiction with the maximum principle for the elliptic operator L.

Since the manifold $(D_\rho(p), \sigma ds^2)$ is complete and hence the α-energy functional on the space of $L_1^{2\alpha}$-maps from U into $D_\rho(p)$ which take on given values on ∂U satisfies condition C, we can apply Lusternik-Schnirelman theory. It follows from Lemma 1 that all critical points of E_α are strict local minima. But applying the "mountain pass lemma" (via Lusternik-Schnirelman theory on Banach manifolds as in Palais [12], [13]) to the space of $L_1^{2\alpha}$ paths joining two distinct critical points would yield a critical point which would not be a strict local minimum, a contradiction. Hence there is only one critical point, proving the uniqueness statement of the Replacement Theorem.

Note that although Lusternik-Schnirelman theory produces critical points in $L_1^{2\alpha}$, regularity theory show that the α-harmonic maps are C^∞. For the sake of the following arguments, we let $(M, g) = (D_\rho(p), \phi ds^2)$, the complete Riemannian manifold described above. If k is sufficiently large, we can define a map

$$H : L^2_k(U, M) \to L^2_{k-2}(f^*TM) \times L^2_{k-1/2}(\partial U, M) \quad \text{by} \quad H(f) = (F(f), \mathrm{ev}(f)), \tag{35}$$

where F is the Euler-Lagrange operator and

$$\mathrm{ev} : L^2_k(U, M) \longrightarrow L^2_{k-1/2}(\partial U, M)$$

is evaluation on the boundary. The linearization of this map is

$$X \in L^2_k(f^*TM) \mapsto (L(X), X|\partial U) \in L^2_{k-2}(f^*TM) \times L^2_{k-1/2}(f^*TM|\partial U),$$

which is surjective by standard existence and regularity theory. An application of the inverse function theorem now implies that f depends continuously on the boundary values.

The proof that the the unique harmonic f depends continuously on the ambient metric is similar. Let $\mathrm{Met}^2_k(M)$ denote the space of L^2_k metrics on M, noting that the Euler-Lagrange operator (21) depends on the metric. Suppose that $h : \partial U \to M$ is given and let

$$L^2_{k,0}(U, M) = \{f \in L^2_k(U, M) : f|\partial U = h\}.$$

For this fixed choice of h, we can then define a map

$$F : L^2_{k,0}(U, M) \times \mathrm{Met}^2_{k-1}(M) \longrightarrow L^2_{k-2}(U, TM).$$

If (f, g) is a critical point for this map F, we can define

$$\pi_V \circ DF(f, g) : T_f L^2_{k,0}(U, M) \oplus T_g \mathrm{Met}^2_{k-1}(M) \longrightarrow T_f L^2_{k-2}(U, M),$$

π_V being the vertical projection. We can divide into components,

$$\pi_V \circ DF(f, g) = (\pi_V \circ D_1 F(f, g), \pi_V \circ D_2 F(f, g)),$$

the first component being the Jacobi operator L. The fact that there are no Jacobi fields implies that this first component of the derivative is an isomorphism. Smooth dependence on the metric therefore follows from the implicit function theorem.

The usual replacement procedure for harmonic maps on a disk utilized by Schoen and Yau [19] can now be obtained by taking the limit as $\alpha \to 1$, and noting that the α-harmonic maps must converge without bubbling, since the bubbling two-spheres cannot exist within the normal coordinate ball.

3 Length and α-energy of necks

The replacement procedure suggests using a combination of the two approaches, infinite-dimensional manifolds and finite-dimensional approximations, for studying gradient-like flows for the α-energy. On appropriate open subsets of the space $\mathrm{Map}(\Sigma, M)$, one can imagine dividing Σ into "thick" and "thin" subsets, the thin subsets approximating trees whose edges are cylindrical parametrizations of curves. as a fist step towards understanding this approximation, we now consider parametrizations of curves as α-energy critical points.

3.1 Curves parametrized as critical points for α-energy

If h is a Riemannian metric on the domain Σ, we can imbed the function $E_{\alpha,h}$ in a larger family of functions that is invariant under rescaling. Thus we define

$$E^{\beta}_{\alpha,h} : \mathrm{Map}(\Sigma, M) \to \mathbb{R} \quad \text{by} \quad E^{\beta}_{\alpha,h}(f) = \frac{1}{2}\int_{\Sigma} (\beta^2 + |df|^2)^{\alpha} dA, \tag{36}$$

where $\alpha > 1$ and $\beta^2 > 0$, so that $E^1_{\alpha,h} = E_{\alpha,h}$. Since we can write

$$E^{\beta}_{\alpha,h}(f) = \frac{\beta^{2(\alpha-1)}}{2}\int_{\Sigma}\left(1 + \frac{|df|^2}{\beta^2}\right)^{\alpha}\beta^2 dA, \tag{37}$$

we see that $E^{\beta}_{\alpha,h}$ can be obtained from $E_{\alpha,h}$ up to a constant multiple by simply rescaling the metric h on Σ. Thus all of the results mentioned before for $E_{\alpha,h}$ hold just as well as for $E^{\beta}_{\alpha,h}$, including the fact that critical points of $E^{\beta}_{\alpha,h}$ are automatically C^{∞}. One advantage of the larger family of functions is that we can take the limit as $\beta \to 0$, the resulting limit having very nice scaling properties that allow more precise estimates.

We want to consider critical points for (36) of the form $\gamma \circ f_0$, where $\gamma : [0, L] \to M$ is a smooth unit-speed geodesic and

$$f_0 \in \mathrm{Map}_0([0, b] \times S^1, [0, L]) = \{\text{smooth maps } f_0 : [0, b] \times S^1 \to [0, L] \\ \text{such that } f_0(0, \theta) = 0, f_0(b, \theta) = L\},$$

θ being the angular coordinate on S^1. In accordance with the discussion at the end of §2.1, f_0 is a critical point for the map

$$E^{\beta}_{\alpha,h} : \mathrm{Map}_0([0, b] \times S^1, \mathbb{R}) \longrightarrow \mathbb{R}. \tag{38}$$

Here the cylinder $[0, \infty) \times S^1$ is diffeomorphic to the punctured unit disk $D_1(0)$ via the map $(u, \theta) \mapsto (r, \theta)$, where $r = e^{-u}$ and the metric h is

$$ds^2 = e^{-2u}(du^2 + d\theta^2) = dr^2 + r^2 d\theta^2. \tag{39}$$

Since the curvature of $\mathbb{R}$ vanishes, we can apply the Replacement Theorem for arbitrarily large choice of ρ to conclude that there is a unique critical point for (38) in $\mathrm{Map}_0([0, b] \times S^1, [0, L])$, and an elementary argument using Fourier analysis shows that it must be of the form $f_0(u, \theta) = \phi(u)$ where ϕ is the unique critical point for

$$F^{\beta}_{\alpha,h} : \mathrm{Map}_0([0, b], [0, L]) \to \mathbb{R},$$

$$F^{\beta}_{\alpha,h}(\phi) = \pi \int_0^b (\beta^2 + e^{2u}|\phi'(u)|^2)^{\alpha} e^{-2u} du, \tag{40}$$

and

$$\mathrm{Map}_0([0, b], [0, L]) = \{\text{smooth maps } \phi : [0, b] \to [0, L] \\ \text{such that } \phi(0) = 0, \phi(b) = L\}.$$

Indeed, if we let $\mu^2 = \beta^2 + e^{2u}|\phi'(u)|^2$, the critical point for (40) is the solution to the Euler-Lagrange equation

$$\frac{d}{du}\left(\mu^{2(\alpha-1)}\frac{d\phi}{du}\right) = 0, \tag{41}$$

and, in agreement with (26),

$$\mu^{2(\alpha-1)}\phi'(u) = (\beta^2 + e^{2u}|\phi'(u)|^2)^{\alpha-1}\phi'(u) = a, \tag{42}$$

where a is a constant.

3.2 Estimates for length, conformal parameter and energy

Our next goal is to understand the relationships between the length L of the geodesic, the conformal parameter $b/2\pi$ and the α-energy of the map $\gamma \circ f_0$ we have just described.

This is easiest to do when we set $\beta = 0$. Then (42) has the explicit decaying exponential solutions

$$\phi'(u) = ce^{-2u(\alpha-1)/(2\alpha-1)u} = ce^{-k(\alpha)u}, \quad \textit{where} \quad k(\alpha) = 2\frac{\alpha-1}{2\alpha-1}. \tag{43}$$

Here the constant of integration $c = \phi'(0)$ is related to the constant a of (42) by $c^{2\alpha-1} = a$. We can calculate the length of the curve,

$$L = L(\phi) = \int_0^b \phi'(u)du = \frac{c}{k}[1 - e^{-kb}], \tag{44}$$

as well as the value of the function $F^0_{\alpha,h}$,

$$\begin{aligned} F^0_{\alpha,h}(\phi) &= \pi\int_0^b e^{2u(\alpha-1)}c^{2\alpha}\left(e^{-2u(\alpha-1)/(2\alpha-1)}\right)^{2\alpha} du \\ &= \pi c^{2\alpha}\int_0^b \exp\left[2u\left(\frac{(\alpha-1)(2\alpha-1)}{2\alpha-1} - \frac{2\alpha(\alpha-1)}{2\alpha-1}\right)\right] du \\ &= \pi c^{2\alpha}\int_0^b e^{-2u(\alpha-1)/(2\alpha-1)}du = \frac{\pi c^{2\alpha}}{k}[1 - e^{-kb}]. \end{aligned} \tag{45}$$

Eliminating c yields the relationship between L, b and the energy of the corresponding critical point for $E^0_{\alpha,h}$,

$$E^0_{\alpha,h}(f_0) = F^0_{\alpha,h}(\phi) = \frac{\pi L^{2\alpha}k^{2\alpha-1}}{[1 - e^{-kb}]^{2\alpha-1}} = \pi L^{2\alpha}\left(\frac{k}{1 - e^{-kb}}\right)^{2\alpha-1}. \tag{46}$$

Since $(d/du)(1 - e^{-ku}) = ke^{-ku}$,

$$be^{-kb} \leq \frac{1 - e^{-kb}}{k} = \int_0^b e^{-ku}du \leq b, \quad \text{when } a \leq b. \tag{47}$$

Moreover, if for some constant c_0,

$$b(\alpha) \le \frac{c_0}{(\alpha-1)^\sigma}, \quad \text{where } 0<\sigma<1, \tag{48}$$

then $e^{-kb} \to 1$ as $\alpha \to 1$. Thus given any $\epsilon > 0$, there is an $\alpha_0 > 1$ such that if $\alpha \in (1, \alpha_0]$, then it follows from (46) that the unique critical point ϕ_α satisfies

$$\frac{\pi L^{2\alpha}}{b^{2\alpha-1}} \le F^0_{\alpha,h}(\phi_\alpha) \le (1+\epsilon)\frac{\pi L^{2\alpha}}{b^{2\alpha-1}}.$$

As $\alpha \to 1$, ϕ_α approaches an affine function ϕ_1 such that

$$F^0_{1,h}(\phi_1) = \frac{\pi L^2}{b}. \tag{49}$$

We see that we can parametrize a curve of given length L with arbitrarily small energy if we let the conformal parameter $b/2\pi$ go to infinity.

We expect similar phenomena when β^2 is small but nonzero. To verify this, we set $\psi(u) = |\phi'(u)|^2$ and rewrite (42) as

$$(\beta^2 + e^{2u}\psi(u))^{2(\alpha-1)}\psi(u) = a^2.$$

Differentiation then gives

$$(\beta^2 + e^{2u}\psi)^{2(\alpha-1)}\psi' + 2(\alpha-1)(\beta^2+e^{2u}\psi)^{2\alpha-3}(2\psi(u)+\psi')e^{2u}\psi = 0.$$

This can be simplified to yield

$$\psi'(u) = -\frac{4(\alpha-1)\psi(u)^2}{\beta^2 e^{-2u} + (2\alpha-1)\psi(u)},$$

which is equivalent to equation (3.12) from [3] when $\beta^2 = 1$. In particular,

$$\psi'(u) \le -\frac{4(\alpha-1)}{2\alpha-1}\psi(u) \quad \text{and} \quad \psi(u) \le \psi(u_0)e^{-\frac{4(\alpha-1)}{2\alpha-1}(u-u_0)},$$

when $u > u_0$, which implies that

$$\phi'(u) \le \phi'(u_0)e^{-\frac{2(\alpha-1)}{2\alpha-1}(u-u_0)}, \quad \text{when } u \ge u_0. \tag{50}$$

On the other hand, we can follow (3.16) of [3] and set

$$v(u) = \psi(u) + \frac{\beta^2}{2\alpha-1}e^{-2u}.$$

A calculation yields the inequality,

$$v'(u) \ge -\frac{4(\alpha-1)}{2\alpha-1}v(u) \quad \text{which implies that}$$
$$\psi(u) + \frac{\beta^2}{2\alpha-1}e^{-2u} \ge \psi(u_0)e^{-\frac{4(\alpha-1)}{2\alpha-1}(u-u_0)}, \quad \text{when } u \ge u_0. \tag{51}$$

Thus

$$\left(\phi'(u) + \frac{\beta}{\sqrt{2\alpha - 1}} e^{-u}\right)^2 \geq (\phi'(u_0))^2 e^{-\frac{4(\alpha-1)}{2\alpha-1}(u-u_0)},$$

and using Taylor's theorem we conclude that

$$\phi'(u) \geq \phi'(u_0) e^{-\frac{2(\alpha-1)}{2\alpha-1}(u-u_0)} - \frac{\beta}{\sqrt{2\alpha-1}} e^{-u}, \quad \text{when } u \geq u_0. \tag{52}$$

If we set $u_0 = 0$, we can rewrite (50) and (52) as

$$\phi'(0)e^{-ku} - \frac{\beta}{\sqrt{2\alpha-1}} e^{-u} \leq \phi'(u) \leq \phi'(0) e^{-ku}. \tag{53}$$

It follows from (42) that

$$a = (\beta^2 + (\phi'(0))^2)^{\alpha-1}\phi'(0) = \phi'(0)^{2\alpha-1}\left(1 + \frac{\beta^2}{\phi'(0)^2}\right)^{\alpha-1}.$$

If we set $x = \phi'(0)$, then

$$c = a^{1/(2\alpha-1)} = x\left(1 + \frac{\beta^2}{x^2}\right)^{(\alpha-1)/(2\alpha-1)} \quad \Rightarrow \quad c \leq x\left(1 + \frac{\alpha-1}{2\alpha-1}\frac{\beta^2}{x^2}\right),$$

and since $x > c$,

$$\frac{\beta^2}{x^2} < \frac{\beta^2}{c^2} \quad \Rightarrow \quad c < x\left(1 + \frac{\alpha-1}{2\alpha-1}\frac{\beta^2}{c^2}\right) \quad \Rightarrow \quad c\left(1 - \frac{\alpha-1}{2\alpha-1}\frac{\beta^2}{c^2}\right) < x,$$

the last implication following from the inequality $1 - y < 1/(1+y)$ when $y > 0$. Thus we find that

$$c\left(1 - \frac{\alpha-1}{2\alpha-1}\frac{\beta^2}{c^2}\right) \leq \phi'(0) \leq c. \tag{54}$$

The estimate (53) integrates to yield an estimate for L in terms of c,

$$\frac{c}{k}[1 - e^{-kb}] - \frac{\beta}{\sqrt{2\alpha-1}}[1 - e^{-b}] \leq L \leq \frac{c}{k}[1 - e^{-kb}], \tag{55}$$

which extends (44) to the case where $\beta \neq 0$. When L is held fixed, we can estimate c by

$$L\frac{k}{1 - e^{-kb}} \leq c \leq \left(L + \frac{\beta}{\sqrt{2\alpha-1}}\right)\frac{k}{1 - e^{-kb}}. \tag{56}$$

From this we conclude that

$$\frac{1}{c} \leq \frac{b}{L} \quad \text{or} \quad \frac{\beta}{c} \leq \frac{b\beta}{L}. \tag{57}$$

To estimate $E_\alpha^\beta(f)$, where $f(u,\theta) = \gamma(\phi(u))$, we use (50) and Hölder's inequality to conclude that

$$
\begin{aligned}
F_\alpha^\beta(\phi) &\leq \pi \int_0^b \left(\beta^2 e^{-2u} + c^2 e^{-4u(\alpha-1)/(2\alpha-1)}\right)^\alpha e^{2(\alpha-1)u} du \\
&\leq \left\{ \left[\pi \int_0^b \left(c^2 e^{-4u(\alpha-1)/(2\alpha-1)}\right)^\alpha e^{2(\alpha-1)u} du \right]^{1/\alpha} + \text{Error} \right\}^\alpha \\
&\leq \left\{ \left[F_\alpha^0(\phi)\right]^{1/\alpha} + \text{Error} \right\}^\alpha ,
\end{aligned}
$$

where

$$
\text{Error} = \left[\pi \int_0^b \beta^{2\alpha} e^{-2\alpha u} e^{2(\alpha-1)u} du \right]^{1/\alpha} .
$$

In view of (56) and the fact that

$$
\frac{k}{1 - e^{-kb}} \leq \frac{1}{be^{-kb}},
$$

we obtain the following:

Lemma. *Suppose that a length L and a small constant $\epsilon > 0$ are given. There is an $\alpha_0 \in (1, \infty)$ such that when $b > 0$ is sufficiently large, whenever $\alpha \in [1, \alpha_0)$ and $\beta > 0$ satisfies the inequality $\pi\beta^{2\alpha} < \epsilon/4$, the E_α^β-minimizing parametrization of any curve of length L parametrized on $[0,b] \times S^1$ has (α, β)-energy $< \epsilon$.*

In other words, we can parametrize a curve of any given length L so that it has arbitrarily small α-energy when α is close to one.

Moreover, if L is smaller than the distance from any point in M to its cut locus and $\gamma_0, \gamma_1 : S^1 \to M$ are curves lying in sufficiently small normal coordinate neighborhoods of points p and q in M such that $d(p,q) = L$, then by continuous dependence on boundary values in the Replacement Theorem, there is a unique E_α^β-minimizing map

$$
h' : [0,b] \times S^1 \longrightarrow M \quad \text{such that} \quad h'(0,t) = \gamma_0(t), \quad h'(b,t) = \gamma_1(t),
$$

with energy satisfying the estimate similar to that of the above Lemma.

3.3 Energy loss in necks and the thick-thin decomposition

We now continue the discussion of the Parker-Wolfson bubble tree started in §1.2, and consider implications of the preceding estimates.

Without making assumptions on the compact manifold M, one could construct sequences of α-energy critical points which lose energy in the necks in the limit as $\alpha \to 1$. However, under the assumption $\pi_1(M)$ be finite, the Replacement Lemma and the Lemma from the previous section imply no loss of energy in necks for minimax sequences corresponding to a given homology or cohomology constraint, as we next explain.

Given a choice of multi-index $(i_1, \ldots, i_k)$, we consider the restrictions of f to the corresponding annulus $A_m = A_{i_1,\ldots,i_k;m}$. On the disk $D_m = A_m \cup B_m$, where $B_m = B_{i_1,\ldots,i_k;m}$ is the corresponding bubble region, we use the polar coordinates (r, θ) centered at the bubble point, and the related coordinates (u, θ), where $u = -\log r + c$, so that the boundary ∂D_m corresponds to $r = 0$. Moreover, we use the flat metric,

$$ds^2 = dr^2 + r^2 d\theta^2 = e^{-2u}(du^2 + d\theta^2).$$

(If the genus of Σ is at least two, this is only an approximation to the restriction of the metric of Σ to the disk, but the approximation becomes better and better as $m \to \infty$.) The annulus A_m is described by the inequalities

$$0 \leq u \leq b_m, \qquad \text{where} \qquad b_m \to \infty,$$

and the restrictions of f_m to A_m can be reparametrized as cylinder maps

$$h'_m = h'_{i_1,\ldots,i_k;m} : [0, b_m] \times S^1 \longrightarrow M.$$

Estimate (14) implies that (after possibly contracting $[0, b_m]$ slightly) the curve $h'_m(\{t\} \times S^1)$ is contained in an $\pi\epsilon_1$-neighborhood of some point in M, for each $t \in [0, b_m]$. Moreover, (14) also implies that cylindrical regions $[a, b] \times S^1$ of given length $c = b - a$ lying within the neck, $h'_m([a, b] \times S^1)$ is contained in an $c\epsilon_1/2$-neighborhood of some point $p \in M$. Thus for a fixed choice of m, we can consider subintervals

$$[a, b] \subset [0, b_m] \qquad \text{such that} \qquad h'_m([a, b] \times S^1) \subset B(p, \rho),$$

for some point $p \in M$ and some radius ρ satisfying the inequality

$$\rho < \min(\pi/(2\kappa), (\text{distance from } p \text{ to its cut locus})),$$

where κ^2 is an upper bound for the sectional curvature on M. This allows us to use the Replacement Theorem from §2.2.

Choose N so that $b_m/N < c$ and divide the interval $[0, b_m]$ into N subintervals

$$[0, b_m/N], [b_m/N, 2b_m/N], \ldots, [(N-1)b_m/N, b_m].$$

It follows from continuous dependence upon boundary conditions, that for any $\epsilon > 0$ there exists $\delta > 0$ such that

$$\begin{aligned} &h'_m(\{(k-1)b_m/N\} \times S^1) \subset B_\delta(p) \quad \text{and} \quad h'_m(\{kb_m/N\} \times S^1) \subset B_\delta(q) \\ &\quad \Rightarrow \quad h'_m([(k-1)b_m/N, kb_m/N] \times S^1) \subset B(C, \epsilon) = \{p \in M : d(p, C) \leq \epsilon\}, \end{aligned}$$

where C is a geodesic arc in M, and d is the distance function defined by the Riemannian metric on M. We can apply the same argument to the intervals

$$\left[\left(k - \frac{1}{2}\right)\frac{b_m}{N}, \left(k + \frac{1}{2}\right)\frac{b_m}{N}\right],$$

to conclude that h'_m approximates a broken geodesic path C_m.

The approximation becomes better and better the further one stays from the ends of the neck because of exponential decay of $\partial h'_m/\partial\theta$ on the interior of the neck. Indeed, after translation, a portion of given length, say $[-10,10]\times S^1$, will get taken to smaller and smaller normal coordinate neighborhoods of a point $p\in M$ as $\alpha\to 1$. In terms of Fermi coordinates along C_m, the equation for α-harmonic maps on $[-10,10]\times S^1$,

$$\frac{D}{\partial u}\left(\frac{\partial h'_m}{\partial u}\right)+\frac{D}{\partial\theta}\left(\frac{\partial h'_m}{\partial\theta}\right)=-(\alpha-1)\left[\frac{\partial}{\partial u}(\log\mu^2)\frac{\partial h'_m}{\partial u}+\frac{\partial}{\partial\theta}(\log\mu^2)\frac{\partial h'_m}{\partial\theta}\right], \tag{58}$$

where $\mu^2=(1+|dh'_m|^2)$, more and more closely approximates the standard equation for harmonic maps in Euclidean space

$$\frac{\partial^2\hat{h}_m}{\partial u^2}+\frac{\partial^2\hat{h}_m}{\partial\theta^2}=0.$$

The latter equation can be be solved by separation of variables and Fourier series, the solutions being

$$\begin{aligned}\hat{h}_m(u,\theta)=\mathbf{a}_0+\mathbf{b}_0u&\\ +\sum_n\,[\mathbf{a}_n\cosh nu\cos n\theta+\mathbf{b}_n\sinh nu\cos n\theta&\\ +\mathbf{c}_n\cosh nu\sin n\theta+\mathbf{d}_n\sinh nu\sin n\theta]&,\end{aligned}$$

the vectors $\mathbf{a}_n$, $\mathbf{b}_n$, $\mathbf{c}_n$ and $\mathbf{d}_n$ being determined by the boundary conditions $\hat{h}|\{-10\}\times S^1$ and $\hat{h}|\{10\}\times S^1$. Except for the linear terms all the terms in the sum exhibit exponential decay on the interior of the interval $[-10,10]$. Since the solution to (58) with given boundary conditions depends smoothly on α as $\alpha\to 1$, the same must be true for h'_m.

The angles between successive geodesic segments must go to zero as $b_m\to\infty$, because otherwise α-energy could be decreased by making the angles smaller, so C_m approaches a geodesic which extends the full length of the neck. (An alternative approach to this convergence to a geodesic is presented in [3].) Inductive application of the Replacement Theorem shows that there is only one α_m-harmonic map which takes on the given boundary values and lies in a given ϵ-tube about C_m, and it must have less α_m-energy than any other map in the given ϵ-tube. We can get an upper bound on the α-energy by comparing with a map of the annulus that maps small annular bands near the boundary to disks and the remainder of the annulus to a parametrization of part of the curve C_m. Thus the Lemma at the end of §3.2 shows that all we need is a bound on the length of C_m to show that the total energy within the neck goes to zero as $\alpha_m\to 1$.

To get the needed estimate on the length, we utilize the assumption that M has finite fundamental group. This allows us to apply Gromov's estimate (12) which states that length of an unstable geodesic is bounded by a constant times its Morse index. Since the Morse index of a minimax sequence corresponding to a homology or cohomology constraint is bounded, this estimate gives a bound on length L of any neck, finishing the proof that no energy is lost in necks in the limit. Thus we conclude that

$$\lim_{m\to\infty} E(f_m, \omega_m) = E(f_\infty, \omega_\infty) + \sum E(g_{i_1,\ldots,i_k}), \tag{59}$$

where $f_\infty : \Sigma \to M$ is the base map and the sum is taken over all the harmonic two-spheres $g_{i_1,\ldots,i_k}$ in the bubble tree.

Finally, we reparametrize the neck region once again, and define

$$h_{i_1,\ldots,i_k;m} : [0,1] \times S^1 \longrightarrow M \qquad \text{by} \qquad h_{i_1,\ldots,i_k;m}(t,\theta) = h'_{i_1,\ldots,i_k;m}(tb_m, \theta). \tag{60}$$

When estimate (48) holds, and hence $e^{-kb} \to 1$, the preceding argument shows that each sequence $m \mapsto h_{i_1,\ldots,i_k;m}$ converges uniformly in C^k on compact subsets to a geodesic $\gamma_{i_1,\ldots,i_k} : [0,1] \to M$, which may be constant. Moreover, $\gamma_{i_1,\ldots,i_k}(1)$ lies in the image of the minimal two-sphere $g_{i_1,\ldots,i_k} : S^2 \to M$, while $\gamma_{i_1,\ldots,i_k}(0)$ lies in the image of $g_{i_1,\ldots,i_{k-1}}$ if $k > 1$, or in the image of $f_\infty : \Sigma \to M$ if $k = 1$.

Definition. A sequence of bubble disks $\{B_{i_1,\ldots,i_k;m}\}$ is *essential* if the corresponding sequence of rescaled maps $g_{i_1,\ldots,i_k:m}$ described in the previous section converge to a nonconstant harmonic two-sphere $g_{i_1,\ldots,i_k}$; otherwise, it is *inessential*. Thus the inessential bubble disks converge to ghost bubbles.

Remark. Given any disk $D_{i_1,\ldots,i_k;m}$ in the above construction, we can we can consider all essential bubble disks $B'_{1;m}$, ..., $B'_{l;m}$ contained within it. We then have a family

$$m \mapsto E_{i_1,\ldots,i_k;m} = D_{i_1,\ldots,i_k;m} - \bigcup_{i=1}^{l} B'_{i;m}.$$

of planar domains possesses a reparametrization which converges to a tree of geodesics which connect nonconstant harmonic two-spheres or nonconstant harmonic two-spheres to the base. This tree is very nicely described at the end of §1 of [14]. As $m \to \infty$ the disks $B'_{i;m}$ contract to points.

A bound on the α-energy yields a bound on the number of bubbles as well as the level of any bubble.

Recall that the sequence of bubble disks $\{B_{i_1,\ldots,i_k;m}\}$ is *essential* if $g_{i_1,\ldots,i_k}$ is a nonconstant harmonic two-sphere. Suppose that hypothesis (10) is satisfied for some union $\mathcal{N}$ of components of $\mathcal{M}(\Sigma, M)$, and that the sequence $\{(f_m, \omega_m)\}$ is chosen to lie in $\mathcal{N}$, so that when energy is bounded by E_0, there is a lower bound L_0 between minimal distance L_0 from base minimal surfaces of genus at least one to minimal two-spheres. A bound E_1 on the α-energy gives a bound on the number of edges in the bubble tree and a bound k on the level of an essential sequence of bubble disks. Moreover, any essential bubble disk $B_{i_1,\ldots i_k;m}$ of level k is contained in a bubble disk $B_{i_1,\ldots i_j;m}$ of level j, with $1 \le j \le k$, such that the corresponding annulus $A_{i_1,\ldots i_j;m}$ parametrizes a curve of length at least L_1, where $L_1 = L_0/k$. The corresponding $h_{i_1,\ldots i_j;m}$'s will then approach a parametrization of a curve C of length $\geq L_1$.

3.4 Conformality of the base

Finally we show that the base map $f_\infty : \Sigma \to M$ obtained in the limit is indeed a parametrized minimal surface:

Conformality Lemma. *The map $f_\infty : \Sigma \to M$ is conformal with respect to the limit conformal structure ω_∞.*

Proof: Since (f_m, ω_m) is a critical points for

$$E_{\alpha_m} : \mathrm{Map}(\Sigma, M) \times \mathrm{Met}_0(\Sigma) \longrightarrow \mathbb{R},$$

it must be the case that

$$\frac{d}{dt} E_\alpha(f, (h_{ab}(t)))\bigg|_{t=0} = 0, \tag{61}$$

whenever $t \mapsto (h_{ab}(t))$ is a variation through constant curvature metrics of total area one on Σ such that $(h_{ab}(0)) = (h_{ab})$ represents the conformal class ω_m. In fact the variation needs only to be tangent to the space of constant curvature metrics of total area one. Thus, if we choose isothermal parameter $z = x_1 + ix_2$ for the initial metric, we can consider a metric variation of the form

$$h_{ab}(t) = \lambda^2 \delta_{ab} + t\dot{h}_{ab}, \qquad a, b = 1, 2, \tag{62}$$

λ^2 being a positive smooth function, where as shown in [20] (see also §5 of [**?**]),

$$\dot{h}_{11} + \dot{h}_{22} = 0, \quad \text{and} \quad (\dot{h}_{11} - i\dot{h}_{12})dz^2$$

is a holomorphic quadratic differential on Σ.

Carrying out the differentiation on the left-hand side of (61) yields

$$\begin{aligned} &\frac{d}{dt} E_\alpha(f, (h_{ab}(t)))\bigg|_{t=0} \\ &\qquad = \frac{\alpha}{2} \int_\Sigma (1 + |df|^2)^{\alpha-1} \sum_{a,b} \frac{d}{dt} \sqrt{h(t)} h^{ab}(t)\bigg|_{t=0} \left\langle \frac{\partial f}{\partial x_a}, \frac{\partial f}{\partial x_b} \right\rangle dx_1 dx_2, \end{aligned} \tag{63}$$

where (h^{ab}) is the matrix inverse to (h_{ab}) and $h = \det(h_{ab})$. Since $dh/dt(0) = 0$,

$$\frac{d}{dt} \begin{pmatrix} \sqrt{h}h^{11} & \sqrt{h}h^{12} \\ \sqrt{h}h^{21} & \sqrt{h}h^{22} \end{pmatrix}\bigg|_{t=0} = \lambda^{-2} \begin{pmatrix} \dot{h}_{22} & -\dot{h}_{12} \\ -\dot{h}_{21} & \dot{h}_{11} \end{pmatrix}.$$

Thus we find that

$$\sum_{a,b} \frac{d}{dt}\sqrt{\eta}\eta^{ab}\bigg|_{t=0} \left\langle \frac{\partial f}{\partial x_a}, \frac{\partial f}{\partial x_b} \right\rangle dx_1 dx_2$$
$$= -\left[\frac{\dot h_{11}}{\lambda^2} \left(\left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_1} \right\rangle - \left\langle \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_2} \right\rangle \right) + \frac{2\dot h_{12}}{\lambda^2} \left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right\rangle \right] dx_1 dx_2$$
$$= -\frac{4}{\lambda^2} \mathrm{Re}\left[\left(\dot h_{11} + i\dot h_{12} \right) \left\langle \frac{\partial f}{\partial z}, \frac{\partial f}{\partial z} \right\rangle \right] dx_1 dx_2.$$

Substitution into (63) yields

$$\frac{d}{dt} E_\alpha(f, (h_{ab}(t)))\bigg|_{t=0} = \frac{\alpha}{2} \int_\Sigma (1 + |df|^2)^{\alpha - 1} \phi dA,$$
$$\text{where} \qquad \phi = -\frac{4}{\lambda^4} \mathrm{Re}\left[\left(\dot h_{11} + i\dot h_{12} \right) \left\langle \frac{\partial f}{\partial z}, \frac{\partial f}{\partial z} \right\rangle \right]. \quad (64)$$

Note that the measure ϕdA is absolutely continuous with respect to $(1/2)|df|^2 dA$, and hence

$$\lim_{m\to\infty} \frac{\alpha}{2} \int_\Sigma (1 + |df|^2)^{\alpha-1} \phi_m dA = \lim_{m\to\infty} \frac{\alpha}{2} \int_{\Sigma_{0;m}} (1 + |df|^2)^{\alpha-1} \phi_m dA$$
$$+ \sum \frac{\alpha}{2} \int_{B_{i_1,\ldots,i_k;m}} (1 + |df|^2)^{\alpha-1} \phi_m dA,$$

the last sum being taken over all bubble regions in the bubble tree. Taking the limits on the right we obtain

$$\lim_{m\to\infty} \frac{\alpha}{2} \int_\Sigma (1 + |df|^2)^{\alpha-1} \phi_m dA$$
$$= \frac{\alpha}{2} \int_\Sigma \mathrm{Re}\left[\left(\dot h_{11} + i\dot h_{12} \right) \left\langle \frac{\partial f_\infty}{\partial z}, \frac{\partial f_\infty}{\partial z} \right\rangle \right] dA$$
$$+ \sum \frac{\alpha}{2} \int_{S^2} \mathrm{Re}\left[\left(\dot h_{11} + i\dot h_{12} \right) \left\langle \frac{\partial g_{i_1,\ldots,i_k}}{\partial z}, \frac{\partial g_{i_1,\ldots,i_k}}{\partial z} \right\rangle \right] dA.$$

All the limits in the sum vanish because the bubble harmonic two-spheres are all conformal. Thus (61) implies that

$$\frac{\alpha}{2} \int_\Sigma \mathrm{Re}\left[\left(\dot h_{11} + i\dot h_{12} \right) \left\langle \frac{\partial f_\infty}{\partial z}, \frac{\partial f_\infty}{\partial z} \right\rangle \right] dA = 0$$

Thus the Hopf differential

$$\left\langle \frac{\partial f_\infty}{\partial z}, \frac{\partial f_\infty}{\partial z} \right\rangle dz^2,$$

a holomorphic quadratic differential, is perpendicular with respect to the natural inner product to all holomorphic quadratic differentials, and this implies that it vanishes. It is well-known that vanishing of the Hopf differential of f_∞ is equivalent to conformality, so the lemma is proven.

4 Proof of the Main Theorem

4.1 Strategy of the proof

Recall that according to the Main Theorem stated at the end of §1.2 the ratio between the inner and outer radii of a neck annulus $A_{i_1,\ldots,i_k;m}$ which approaches a parametrization of ageodesic of length L satisfies

$$\frac{\eta_{i_1,\ldots,i_k;m}}{\epsilon_{i_1,\ldots,i_k;m}} \sim \exp\left(\frac{-c_0}{\sqrt{\alpha_m - 1}}\right) \tag{65}$$

as $\alpha_m \to 1$, where $\sim$ denotes asymptotic behavior and c_0 denotes a constant depending upon L and the energy within the bubble. It suffices to prove this with the expression $\sqrt{\alpha - 1}$ in (65) replaced by $(\alpha - 1)^\sigma$, where $\sigma = 1/2\alpha$, because of the limits

$$\lim_{\alpha\to 1}(\alpha - 1)\log(\alpha - 1) = 0, \qquad \lim_{\alpha\to 1}\frac{(\alpha-1)^\alpha}{\alpha - 1} = 1, \qquad \lim_{\alpha\to 1}\frac{\alpha - 1}{(\alpha-1)^{1/\alpha}} = 1.$$

Thus to establish (17) it will suffice show that there is an $\alpha_0 \in (1,\infty)$ such that for $1 < \alpha_m \le \alpha_0$,

$$\frac{\eta_{i_1,\ldots,i_k;m}}{\epsilon_{i_1,\ldots,i_k;m}} \le \exp\left(\frac{-c_1}{(\alpha_m - 1)^{1/2\alpha_m}}\right), \tag{66}$$

and

$$\frac{\eta_{i_1,\ldots,i_k;m}}{\epsilon_{i_1,\ldots,i_k;m}} \ge \exp\left(\frac{-c_2}{(\alpha_m - 1)^{1/2\alpha_m}}\right), \tag{67}$$

where c_1 and c_2 are positive constants which approach

$$\sqrt{\frac{\pi}{2}}\frac{L_{i_1,\ldots,i_k}}{\sqrt{E_b}} \quad \text{as } \alpha_0 \to 1,$$

E_b being the limit of the usual energy of f_m restricted to the ball $B_{i_1,\ldots,i_k;m}$ as $m \to \infty$. Indeed, if

$$g_m : D_1 \to M \qquad \text{denotes the rescaling of} \qquad f_m : B_{i_1,\ldots,i_k;m} \to M$$

to the unit disk D_1 and

$$E_b = \lim_{m\to\infty}\frac{1}{2}\int |dg_m|^2 dxdy, \quad E_{\alpha_0,b} = \lim_{m\to\infty}\frac{1}{2}\int |dg_m|^{2\alpha_0} dxdy,$$

we will prove these estimates for

$$c_1 = (1 - \delta_1(\alpha_0))L_{i_1,\ldots,i_k}\left(\frac{\pi}{2E_{\alpha_0,b}}\right)^{1/2}$$

$$\text{and} \quad c_2 = (1 + \delta_2(\alpha_0))L_{i_1,\ldots,i_k}\left(\frac{\pi}{2E_b}\right)^{1/2\alpha_0},$$

where and $\delta_1(\alpha_0)$ and $\delta_2(\alpha_0)$ approach zero as $\alpha_0 \to 1$.

4.2 A model space for bubbling

To better understand the proof, we introduce a model for bubbling. It is helpful to consider first the case of a single bubble. Thus we imagine that a single bubble two-sphere forms, and consider a space of maps f from Σ to M that can approximate a sequence of critical points for E_α as $\alpha \to 1$. Let p be a point in M and divide Σ into three pieces:

1. a *base* $\Sigma_0 = \Sigma - D_\epsilon(p)$, where $D_\epsilon(p)$ is a disk of radius ϵ about p,
2. a smaller concentric disk $D_\eta(p)$ called the *bubble*, and
3. an annulus $A = D_\epsilon(p) - D_\eta(p)$ called the *neck*.

As before, we imagine that the disk $D_\epsilon(p)$ is given the standard flat metric

$$ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\theta^2 = e^{-2u}(du^2 + d\theta^2),$$

where (r, θ) are the usual polar coordinates and $r = e^{-u}$.

Definition. We let $\mathrm{Map}_{p,\epsilon,\eta}(\Sigma, M)$ denote the subspace of $\mathrm{Map}(\Sigma, M)$ consisting of the piecewise smooth maps $f : \Sigma \to M$, smooth on each of the three regions Σ_0, $D_\eta(p)$ and A, which satisfy the following conditions:

1. f takes the circles $\partial D_\epsilon(p)$ and $\partial D_\eta(p)$ to points, and
2. $f|A$ is independent of θ, of the form $f(u, \theta) = \gamma \circ \phi(u)$, where $\gamma : [0, L] \to M$ is a unit-speed curve of length L, and

   $$\phi : [-\log \epsilon, -\log \eta] \to [0, L]$$

   is the unique critical point for the function

   $$F_\alpha(\phi) = \pi \int_{-\log \epsilon}^{-\log \eta} (1 + e^{2u}|\phi'(u)|^2)^\alpha e^{-2u} du, \tag{68}$$

   which we considered in § 3.2.

We can regard $\mathrm{Map}_{p,\epsilon,\eta}(\Sigma, M)$ as a smooth infinite-dimensional manifold when completed with respect to a suitable Sobolev norm. Closely related is the smooth manifold

$$\mathrm{Map}^\star_\epsilon(\Sigma, M) = \{(f, p, \eta) \in \mathrm{Map}(\Sigma, M) \times M \times (\epsilon, \infty) \\ : f \in \mathrm{Map}_{p,\epsilon,\eta}(\Sigma, M)\}. \tag{69}$$

in which we allow the bubble point p and the parameter η to vary. The previous space $\mathrm{Map}_{p,\epsilon,\eta}(\Sigma, M)$ can be regarded as the fiber over (p, η) of a continuous map

$$\pi : \mathrm{Map}^\star_\epsilon(\Sigma, M) \longrightarrow M \times (\epsilon, \infty),$$

the projection on the last two factors.

The model is flexible enough to apply to several cases:

1. If the single bubble forms at the point $p \in \Sigma$, we have a single sequence of bubble disks $B_{1;m}$ centered at p, a single sequence of neck regions $A_{1;m}$, and a single sequence of rescaled maps $g_{1;m} : D_m(0) \to M$ which converges uniformly on compact subsets of $\mathbb{C}$ to a harmonic two-sphere. In this case, we can set

$$\epsilon_m = \epsilon_{1;m} = (\text{radius of } A_{1;m} \cup B_{1;m}), \quad \eta_m = \eta_{1;m} = (\text{radius of } B_{1;m}).$$

When m is sufficiently large, f_m will closely approximate an element of the space $\mathrm{Map}_{p,\epsilon_m,\eta_m}(\Sigma, M)$.

2. The same model can be adapted to apply to the case in which several bubbles are forming, by setting

$$\epsilon = \epsilon_{i_1,\ldots,i_k;m}, \qquad \eta = \eta_{i_1,\ldots,i_k;m}.$$

In this case, the restriction of f_m to D_η may approach a tree consisting of minimal two spheres connected by geodesics, while the restriction of f_m to $\Sigma - D_\epsilon$ may approach a base minimal surface connected to minimal two-spheres by geodesics.

3. We can contract the neck to have a given length within M. In either of the previous cases, we can make η larger than needed, so that the restriction of f_m to the annular region, when properly rescaled, approaches a curve of a given fixed length, the length being less than the distance from any point in M to its conjugate locus.

In each of the three cases, we fix $\epsilon > 0$, and consider a family of maps depending continuously on the parameter b,

$$b \in (0, \infty) \quad \mapsto \quad \zeta(b) = (f_b, p, \eta(b)) \in \mathrm{Map}_{\epsilon,\eta_0}(\Sigma, M), \qquad \text{where} \quad \eta(b) = e^{-b}\epsilon, \tag{70}$$

satisfying the additional conditions that

1. the restriction of each f_b to the base Σ_0 is fixed,
2. the restriction of each f_b to the neck A is of the form $f(u, \theta) = \gamma \circ \phi_b(u)$, where γ is a fixed curve,
3. the restriction of each f_b to the bubble $B_\eta(p)$ is obtained from a fixed map $g : D_1 \to M$ by rescaling, where D_1 is the unit disk.

The third condition means that there is a map

$$g : D_1 \to M, \quad \text{such that} \quad f_b(x, y) = g(\bar{x}, \bar{y}), \quad \text{where} \quad \begin{cases} x = \eta\bar{x}, \\ y = \eta\bar{y}, \end{cases}, \tag{71}$$

$(\bar{x}, \bar{y})$ being the standard coordinates on the unit disk D_1.

We consider the effect on the energy of reparametrizations in which b, and hence $\eta = e^{-b}\epsilon$, are varied, thereby changing the size of the bubble region. Let

$$G_\alpha(b) = E_{\alpha,h}(f_b|D_{e^{-b}\epsilon}(p)), \qquad H_\alpha(b) = E_{\alpha,h}(f_b|(D_\epsilon(p) - D_{e^{-b}\epsilon}(p))).$$

It is clear that as b increases, $G_\alpha(b)$ increases while $H_\alpha(b)$ decreases. Our goal is to show that when b is sufficiently small, the derivative of H_α dominates and we can decrease energy by increasing b. Conversely, when b is sufficiently large, the derivative of G_α dominates, and we can decrease energy by making b smaller.

4.3 Estimates when $\beta = 0$

The needed estimates are easier to understand if we introduce an additional parameter β and consider the family of functions

$$E^\beta_{\alpha,h}(f) = \frac{1}{2}\int_\Sigma (\beta^2 + |df|^2_h)^\alpha dA,$$

just as we did in §3.2, where $\beta \to 0$. The parameter β appears when making a change of scale and expanding the disk $D_\epsilon(p) = A \cup B_\eta(p)$ to a disk of unit radius. Indeed, we replace f_b by $\tilde{f}_{\tilde{b}}$, where

$$f_b(x,y) = \tilde{f}_{\tilde{b}}(\tilde{x},\tilde{y}), \quad \text{where} \quad \begin{cases} x = \epsilon\tilde{x}, \\ y = \epsilon\tilde{y}. \end{cases}$$

Although the ordinary energy would be invariant under such a rescaling, the α-energy is not as we have already noted, and in fact

$$\begin{aligned} \frac{1}{2}\int_{D_\epsilon}(1+|df_b|^2)^\alpha dxdy &= \frac{1}{2}\int_{D_1}(1+\frac{|d\tilde{f}_{\tilde{b}}|}{\epsilon}^2)^\alpha \epsilon^2 d\tilde{x}d\tilde{y} \\ &= \frac{1}{2\epsilon^{2(\alpha-1)}}\int_{D_1}(\epsilon^2 + |d\tilde{f}_{\tilde{b}}|^2)^\alpha d\tilde{x}d\tilde{y} = \frac{1}{\epsilon^{2(\alpha-1)}}E^\epsilon_{\alpha,\tilde{h}}(\tilde{f}_{\tilde{b}}), \end{aligned}$$

where $\tilde{h} = d\tilde{x}^2 + d\tilde{y}^2$.

To simplify notation, we drop the tilde, and obtain a new family of maps

$$b \in (1,\infty) \quad \mapsto \quad f_b \in \mathrm{Map}(D_1(0), M).$$

Our goal is to minimize $G^\beta_\alpha(b) + H^\beta_\alpha(b)$, where

$$G^\beta_\alpha(b) = E^\beta_{\alpha,h}(f_b|D_{e^{-b}}(0)), \qquad H^\beta_\alpha(b) = E^\beta_{\alpha,h}(f_b|A), \tag{72}$$

where h is the standard flat metric on D_1 and we have replaced the old A with

$$A = D_1(0) - D_{e^{-b}}(0).$$

In this section, we describe the needed estimates in the limiting case in which $\beta \to 0$. We assume that the restriction of f_b to A is a critical point for the α-energy with Dirichlet boundary conditions for each choice of b.

Recall from §3.2 that since $f(u,\theta) = \gamma \circ \phi(u)$ on A,

$$E^0_{\alpha,h}(f|A) = F^0_{\alpha,h}(\phi) = \pi \int_0^b |\phi'(u)|^{2\alpha} e^{2(\alpha-1)u} du,$$

and there is a unique critical point ϕ to the functional $F^0_{\alpha,h}$ which satisfies the end point conditions $\phi(0) = 0$ and $\phi(b) = L$. This critical point is given by the formula

$$\phi'(u) = ce^{-ku} = ce^{-(2(\alpha-1))u/(2\alpha-1)}, \qquad \text{where } c \text{ is a constant,}$$

the constant being determined by the end point conditions. It follows from (44) that the relationship between b and c (for given choices of α and L) is given by

$$c = \mu(b) = L\frac{k}{1-e^{-kb}}, \quad \text{where } \mu'(b) < 0 \text{ and } \quad \lim_{b\to\infty} \mu(b) = kL. \tag{73}$$

From (46), we conclude that

$$H^0_\alpha(b) = E^0_{\alpha,h}(f_b|A) = \pi L^{2\alpha} \left(\frac{k}{1-e^{-kb}}\right)^{2\alpha-1}. \tag{74}$$

On the other hand, we can consider the energy within $D_\eta(0)$, where for the rescaled map, we take $\eta = e^{-b}$. Note that it follows from (71) that $dg = \eta df_b$, and hence

$$\begin{aligned} G^0_\alpha(b) &= E^0_{\alpha,h}(f_b|D_{e^{-b}}(0)) \\ &= \frac{1}{2}\left(\frac{1}{\eta}\right)^{2(\alpha-1)} \int_{D_1} |dg|^{2\alpha} d\bar{x} d\bar{y} = \frac{1}{2} e^{2(\alpha-1)b} \int_{D_1} |dg|^{2\alpha} d\bar{x} d\bar{y}. \end{aligned} \tag{75}$$

To see that there is a unique value of b which minimizes the sum

$$E^0_{\alpha,h}(f_b|D_\epsilon(p)) = G^0_\alpha(b) + H^0_\alpha(b),$$

we first note that

$$G^0_\alpha(b) \geq \frac{1}{2}\int_{D_1} |dg|^2 dx dy,$$

so $G^0_\alpha(b)$ is bounded below by a positive constant. Moreover,

$$\frac{dG^0_\alpha}{db}(b) = (\alpha-1)e^{2(\alpha-1)b} \int_{D_1} |dg|^{2\alpha} d\bar{x} d\bar{y} \geq 0,$$

$$\frac{d^2G^0_\alpha}{db^2}(b) = 2(\alpha-1)^2 e^{2(\alpha-1)b} \int_{D_1} |dg|^{2\alpha} d\bar{x} d\bar{y} \geq 0,$$

so $G^0_\alpha(b)$ is strictly increasing and concave up. On the other hand,

$$H^0_\alpha(b) = \pi L^{2\alpha} \left(\frac{\frac{2(\alpha-1)}{2\alpha-1}}{1-e^{-kb}}\right)^{2\alpha-1} \quad \to \quad \pi L^{2\alpha}\left(\frac{2(\alpha-1)}{2\alpha-1}\right)^{2\alpha-1} \tag{76}$$

as $b \to \infty$, and $H^0_\alpha(b) \to \infty$ as $b \to 0$. Moreover, since

$$\frac{d}{db}\left(\frac{1-e^{-kb}}{k}\right) = e^{-kb},$$

we see that

$$\frac{dH^0_\alpha}{db}(b) = (1-2\alpha)e^{-kb}\pi L^{2\alpha}\left(\frac{k}{1-e^{-kb}}\right)^{2\alpha} \leq 0,$$

$$\frac{d^2H^0_\alpha}{db^2}(b) = (-k-2\alpha)(1-2\alpha)e^{-2kb}\pi L^{2\alpha}\left(\frac{k}{1-e^{-kb}}\right)^{2\alpha+1} \geq 0.$$

so $H^0_\alpha(b)$ is strictly decreasing and concave up. Thus

$$E^0_\alpha(b) = G^0_\alpha(b) + H^0_\alpha(b) \quad \text{satisfies} \quad \frac{d^2E^0_\alpha}{db^2}(b) \geq 0,$$

and hence E^0_α can have at most one local minimum, and any such minimum must be a global minimum. Such a minimum can occur only if

$$\frac{dG^0_\alpha}{db}(b) = -\frac{dH^0_\alpha}{db}(b),$$

or equivalently,

$$(\alpha-1)e^{2(\alpha-1)b}\int_{D_1}|dg|^{2\alpha}d\bar{x}d\bar{y} = (2\alpha-1)e^{-kb}\pi L^{2\alpha}\left(\frac{k}{1-e^{-kb}}\right)^{2\alpha}. \tag{77}$$

If we choose α small enough, there will in fact be a unique solution to (77) for a given choice of L, in which the "energy in the bubble" balances the "energy in the neck."

We want to give estimates on the choice of b which achieves this minimum. To do this, we first note that

$$e^{-kb}b \leq \frac{1-e^{-kb}}{k} = \int_0^b e^{-kx}dx \leq b. \tag{78}$$

It therefore follows from (77) that

$$(\alpha-1)e^{2(\alpha-1)b}\int_{D_1}|dg|^{2\alpha}d\bar{x}d\bar{y} \geq (2\alpha-1)e^{-kb}\pi L^{2\alpha}\frac{1}{b^{2\alpha}}, \tag{79}$$

which in turn implies that when $\alpha \leq \alpha_0$,

$$(\alpha-1)b^{2\alpha} \geq \frac{\pi L^{2\alpha}}{2E_{\alpha_0,b}(g)}e^{-2\alpha kb}, \quad \text{where} \quad E_{\alpha_0,b}(g) = \frac{1}{2}\int_{D_1}|dg|^{2\alpha_0}d\bar{x}d\bar{y}.$$

Since $b \leq (\text{constant})(\alpha-1)^{-\sigma}$, for $\sigma \in (0,1)$ implies that $e^{-2\alpha kb} \to 1$ as $\alpha \to 1$, we conclude that there is a small $\delta > 0$ such that

$$b \geq \frac{c_1}{(\alpha-1)^{1/2\alpha}}, \qquad \text{where} \qquad c_1 = (1-\delta)L\left(\frac{\pi}{2E_{\alpha_0,b}}\right)^{1/2}, \tag{80}$$

for $1 < \alpha \leq \alpha_0$. (Here we use the fact that $0 < x < 1$ implies that $x^{1/2\alpha} > x^{1/2}$.) Thus we obtain an upper bound an radius,

$$e^{-b} \leq \exp\left(\frac{-c_1}{(\alpha-1)^{1/2\alpha}}\right), \tag{81}$$

the first estimate of the Main Theorem when $\beta = 0$.

A lower bound on the radius of the bubble region can be obtained in a similar fashion. It follows from (77) and (78) that

$$(\alpha-1)e^{2(\alpha-1)b} \int_{D_1} |dg|^{2\alpha} d\bar{x}d\bar{y} \leq (2\alpha-1)e^{-kb}\pi L^{2\alpha}\frac{e^{2\alpha kb}}{b^{2\alpha}}. \tag{82}$$

and since $e^{2(\alpha-1)b} = e^{(2\alpha-1)kb}$, we conclude that

$$b^{2\alpha} \leq (2\alpha-1)\pi L^{2\alpha}\frac{1}{2(\alpha-1)E_b(g)}, \quad \text{where} \quad E_b(g) = \frac{1}{2}\int_{D_1} |dg|^2 d\bar{x}d\bar{y},$$

or there is a small $\delta > 0$ such that

$$b \leq \frac{c_2}{(\alpha-1)^{1/2\alpha}}, \qquad \text{where} \qquad c_2 = (1+\delta)L\left(\frac{\pi}{2E_b}\right)^{1/2\alpha_0}, \tag{83}$$

for $1 < \alpha \leq \alpha_0$. (Here we use the fact that $0 < x < 1$ implies that $x^{1/2\alpha} < x^{1/2\alpha_0}$.) This gives the second estimate of the Main Theorem when $\beta = 0$.

4.4 The upper bound on radius in the general case

Once one has the estimate (80) for the case $\beta = 0$, a similar estimate can be obtained when $\beta > 0$ is sufficiently small by a relatively straightforward application of Taylor's theorem. As in the preceding section, we consider a family of rescaled maps,

$$b \in (1,\infty) \quad \mapsto \quad f_b \in \mathrm{Map}(D_1(0), M),$$

and we seek to establish estimates for the functions

$$G^\beta_\alpha(b) = E^\beta_{\alpha,h}(f_b|D_{e^{-b}}(0)) \quad \text{and} \quad H^\beta_\alpha(b) = E^\beta_{\alpha,h}(f_b|A),$$

where $A = D_1(0) - D_{e^{-b}}(0)$, and now β is nonzero. Our strategy is to show that unless the first estimate (66) of the Scaling Theorem holds,

$$\frac{dG^\beta_\alpha}{db}(b) < -\frac{dH^\beta_\alpha}{db}(b).$$

In other words, we show that the sum can be decreased by increasing b, unless b satisfies (81) when α is sufficiently close to one.

As far as the restriction to the bubble region goes, we now need to estimate the derivative of

$$G^\beta_\alpha(b) = \frac{1}{2}e^{2(\alpha-1)b}\int_{D_1} (e^{-2b}\beta^2 + |dg|^2)^\alpha d\bar{x}d\bar{y}. \tag{84}$$

We quickly verify that

$$\frac{dG_\alpha^\beta}{db}(b) = (\alpha-1)G_\alpha^\beta(b) \\ - \alpha e^{2(\alpha-1)b}(\beta^2 e^{-2b}) \int_{D_1} (e^{-2b}\beta^2 + |dg|^2)^{\alpha-1} d\bar{x} d\bar{y}, \quad (85)$$

and hence

$$\frac{dG_\alpha^\beta}{db}(b) \leq 2(\alpha-1)G_\alpha^\beta(b), \quad (86)$$

where $G_\alpha^\beta(b)$ is the (α,β)-energy of the bubble.

We also need to consider the effect of varying b on

$$H_\alpha^\beta(b) = \pi \int_0^b (\beta^2 e^{-2u} + \phi_a'(u)^2)^\alpha e^{2(\alpha-1)u} du, \quad (87)$$

where $\phi_a'(u)$ is the function determined implicitly by

$$(e^{-2u}\beta^2 + \phi_a'(u)^2)^{\alpha-1}\phi_a'(u) = ae^{-2(\alpha-1)u}, \quad (88)$$

a being the constant chosen so that

$$\int_0^b \phi_a'(u) du = L, \quad \text{where } L \text{ is a given constant.} \quad (89)$$

Thus

$$\frac{dH_\alpha^\beta(b)}{db} = \frac{d}{db} h(a(b), b), \\ \text{where} \quad h(a,b) = \pi \int_0^b (\beta^2 e^{-2u} + \phi_a'(u)^2)^\alpha e^{2(\alpha-1)u} du,$$

and $a(b)$ is determined implicitly by the constraint

$$\int_0^b \phi_a'(u) du = L. \quad (90)$$

The condition $dL = 0$ implies that

$$\frac{da}{db} = \frac{-\phi_a'(b)}{\int_0^b \varepsilon'(u) du}, \quad \text{where} \quad \varepsilon(u) = \frac{\partial \phi_a}{\partial a}(u). \quad (91)$$

We calculate

$$\frac{dH_\alpha^\beta}{db} = \frac{d}{db}\left(\pi \int_0^b (\phi_a'(u)^2 + e^{-2u}\beta^2)^\alpha e^{2(\alpha-1)u} du\right) \\ = \pi \left[(\phi_a'(b)^2 + e^{-2b}\beta^2)^\alpha e^{2(\alpha-1)b}\right] \\ + \pi \left[\int_0^b 2\alpha(\phi_a'(u)^2 + e^{-2u}\beta^2)^{\alpha-1}\phi_a'(u)\varepsilon'(u) e^{2(\alpha-1)u} du\right]\left[\frac{da}{db}\right], \quad (92)$$

where da/db is given by (91). It follows from (88) that

$$\int_0^b 2\alpha(\phi_a'(u)^2 + e^{-2u}\beta^2)^{\alpha-1}\phi_a'(u)\varepsilon'(u)e^{2(\alpha-1)u}du = \int_0^b 2a\alpha\varepsilon'(u)du, \tag{93}$$

and hence using (91), we can simplify (92) to

$$\frac{dH_\alpha^\beta}{db} = \left[(\phi_a'(b)^2 + e^{-2b}\beta^2)^\alpha e^{2(\alpha-1)b}\right] - 2a\alpha\phi_a'(b). \tag{94}$$

The idea is to regard each of the terms in (94) as a perturbation of the corresponding expression for the case where $\beta = 0$.

For estimating the first of these expressions, we can regard $\phi_a'(u)^2 + e^{-2u}\beta^2$ as a small perturbation of $\phi_a'(u)^2$. It follows from Taylor's theorem that if $1 < \alpha < 2$, $x > 0$ and $\zeta \geq 0$,

$$x^\alpha \leq (x+\zeta)^\alpha \leq x^\alpha\left(1 + \alpha\frac{\zeta}{x} + \frac{1}{2}\alpha(\alpha-1)\frac{\zeta^2}{x^2}\right).$$

Thus

$$\begin{aligned}\phi_a'(b)^{2\alpha}e^{2(\alpha-1)b} &\leq \left[(\phi_a'(b)^2 + e^{-2b}\beta^2)^\alpha e^{2(\alpha-1)b}\right] \\ &\leq \phi_a'(b)^{2\alpha}e^{2(\alpha-1)b}\left(1 + \alpha\frac{e^{-2b}\beta^2}{\phi_a'(b)} + \frac{1}{2}\alpha(\alpha-1)\frac{e^{-4b}\beta^4}{\phi_a'(b)^2}\right).\end{aligned}$$

Applying (50) and (54) then yields

$$\begin{aligned}&\left[(\phi_a'(b)^2 + e^{-2b}\beta^2)^\alpha e^{2(\alpha-1)b}\right] \\ &\qquad \leq a^{2\alpha/(2\alpha-1)}e^{-kb}\left(1 + \alpha\frac{e^{-2b}\beta^2}{\phi_a'(b)} + \frac{1}{2}\alpha(\alpha-1)\frac{e^{-4b}\beta^4}{\phi_a'(b)^2}\right),\end{aligned} \tag{95}$$

giving an estimate for the first term in (94).

On the other hand, the second term can be estimated directly by (54):

$$\begin{aligned}&2a\alpha\left(a^{1/(2\alpha-1)}(1-\epsilon)e^{-kb} - \frac{\beta}{\sqrt{2\alpha-1}}e^{-b}\right) \leq 2a\alpha\phi_a'(b), \\ &\qquad\qquad \text{where} \quad \epsilon = \frac{\alpha-1}{2\alpha-1}\frac{\beta^2}{a^{2/(2\alpha-1)}}.\end{aligned} \tag{96}$$

We can substitute (95) and (96) into $H_{\alpha,\beta}$ to obtain the result:

$$\begin{aligned}\frac{dH_{\alpha,\beta}}{db}(b) \leq c^{2\alpha}&\left[e^{-kb}\left(1 + \alpha\frac{e^{-2b}\beta^2}{\phi_a'(b)} + \frac{1}{2}\alpha(\alpha-1)\frac{e^{-4b}\beta^4}{\phi_a'(b)^2}\right)\right. \\ &\qquad \left. -2\alpha\left((1-\epsilon)e^{-kb} - \frac{\beta}{c\sqrt{2\alpha-1}}e^{-b}\right)\right].\end{aligned} \tag{97}$$

where

$$c^{2\alpha-1} = a \qquad \text{and} \qquad \epsilon = \frac{\alpha-1}{2\alpha-1}\frac{\beta^2}{c^2}.$$

Finally, it follows from (89), (53) and (54) that

$$L = \int_0^b \phi_a'(u)du \le \int_0^b \phi_a'(0)e^{-ku}du \le \phi'(0)\frac{1-e^{-kb}}{k} \le c\frac{1-e^{-kb}}{k},$$

which implies that

$$c \ge \frac{Lk}{1-e^{-kb}}.$$

Thus it follows from (97) that

$$\begin{aligned} -\frac{dH_\alpha^\beta}{db}(b) &\ge (2\alpha-1)e^{-kb}\left(\frac{Lk}{1-e^{-kb}}\right)^{2\alpha}(1-\text{Error}) \\ &\ge (2\alpha-1)e^{-kb}\left(\frac{L}{b}\right)^{2\alpha}(1-\text{Error})\,, \end{aligned} \tag{98}$$

where the error term is small as long as β/c is sufficiently small.

It follows from (86) and (98) that

$$\frac{dG_\alpha^\beta}{db}(b) + \frac{dH_\alpha^\beta}{db}(b) \le 2(\alpha-1)G_\alpha^\beta(b) - (2\alpha-1)e^{-kb}\left(\frac{L}{b}\right)^{2\alpha}(1-\text{Error})\,,$$

and hence the sum on the left is negative unless

$$2(\alpha-1)G_\alpha^\beta(b) \ge (2\alpha-1)e^{-kb}\left(\frac{L}{b}\right)^{2\alpha}(1-\text{Error})\,,$$

or by (84),

$$\begin{aligned} (\alpha-1)e^{2(\alpha-1)b}\int_{D_1}(e^{-2b}\beta^2 + |dg|^2)^\alpha d\bar{x}d\bar{y} & \\ \ge (2\alpha-1)e^{-kb}\left(\frac{L}{b}\right)^{2\alpha}(1-\text{Error})\,. & \end{aligned} \tag{99}$$

We now proceed as in the case $\beta = 0$, using (99) instead of (79). We conclude as in the preceding section that if (66) does not hold, then

$$\frac{dG_\alpha^\beta}{db} + \frac{dH_\alpha^\beta}{db} < 0$$

whenever α is sufficiently close to one, under the assumption that we can choose β/c to be arbitrarily small.

4.5 The lower bound on radius in the general case

Our goal in this case is to show that

$$\frac{dG_\alpha^\beta}{db}(b) > -\frac{dH_\alpha^\beta}{db}(b).$$

unless b satisfies the second estimate (67) of the Scaling Theorem.

To get an estimate for the derivative of G_α^β, we let $h = \sup(|df|^2, 1)$ and let A be the subset of D_1 on which $h \geq 1$. Then

$$\int_{D_1} (e^{-2b}\beta^2 + |dg|^2)^{\alpha-1} d\bar{x}d\bar{y} \\ \leq B = \int_{D_1 - A} (e^{-2b}\beta^2 + 1)^{\alpha-1} d\bar{x}d\bar{y} + \int_A (e^{-2b}\beta^2 + |dg|^2) d\bar{x}d\bar{y}.$$

Thus it follows from (85) that

$$\frac{dG_\alpha^\beta}{db}(b) \geq 2(\alpha-1)G_\alpha^\beta(b) - (2\alpha)e^{2(\alpha-1)b}(\beta^2 e^{-2b})B, \tag{100}$$

the last term going to zero as $\beta^2 \to 0$, since boundedness of energy implies that $e^{2(\alpha-1)b}$ is bounded, for reasons described following the statement of the Main Theorem in §1.2.

To estimate the derivative of H_α^β, we use (94),

$$\frac{dH_\alpha^\beta}{db} = \left[(\phi_a'(b)^2 + e^{-2b}\beta^2)^\alpha e^{2(\alpha-1)b}\right] - 2a\alpha\phi_a'(b), \tag{101}$$

and once again regard each term as a perturbation of the corresponding expression for the case where $\beta = 0$.

We note first that

$$(\phi_a'(b)^2 + e^{-2b}\beta^2)^\alpha e^{2(\alpha-1)b} \geq \phi_a'(b)^{2\alpha} e^{2(\alpha-1)b}.$$

Applying (52) and (54) then yields

$$\left[(\phi_a'(b)^2 + e^{-2b}\beta^2)^\alpha e^{2(\alpha-1)b}\right] \geq \phi_a'(0)e^{-kb} - \frac{\beta}{\sqrt{2\alpha-1}}e^{-b} \\ \geq c^{2\alpha}\left(1 - \frac{\alpha-1}{2\alpha-1}\frac{\beta^2}{c^2}\right)\left(\phi_a'(0)e^{-kb} - \frac{\beta}{\sqrt{2\alpha-1}}e^{-u}\right), \tag{102}$$

giving an estimate for the first term in (101). Note that the error term is small if both β/c and βe^{-b} are small.

On the other hand, the second term can be estimated directly by (54):

$$2\alpha\phi_a'(b) \leq 2\alpha c. \tag{103}$$

We can substitute (102) and (103) into $H_{\alpha,\beta}$ to obtain the result:

$$\frac{dH_{\alpha,\beta}}{db}(b) \geq c^{2\alpha}\left[\left(1-\frac{\alpha-1}{2\alpha-1}\frac{\beta^2}{c^2}\right)\left(\phi_a'(0)e^{-kb}-\frac{\beta}{\sqrt{2\alpha-1}}\right)-2\alpha e^{-kb}\right]. \tag{104}$$

Finally, it follows from (89), (53) and (54) that

$$\begin{aligned} L = \int_0^b \phi_a'(u)du \geq \int_0^b \left(ce^{-ku} - \frac{\beta}{\sqrt{2\alpha-1}}e^{-u}\right)du \\ \geq c\frac{1-e^{-kb}}{k} - \frac{\beta}{\sqrt{2\alpha-1}}(1-e^{-b}), \end{aligned}$$

which implies that

$$c \leq \left(L + \frac{\beta}{\sqrt{2\alpha-1}}(1-e^{-b})\right)\frac{k}{1-e^{-kb}}.$$

Thus it follows from (104) that

$$\begin{aligned} -\frac{dH_\alpha^\beta}{db}(b) &\leq (2\alpha-1)e^{-kb}\left(\frac{Lk(1+\mathrm{Error})}{1-e^{-kb}}\right)^{2\alpha} \\ &\leq (2\alpha-1)e^{-kb}\left(\frac{L\,(1+\mathrm{Error})}{be^{-kb}}\right)^{2\alpha}, \end{aligned} \tag{105}$$

where the error term is small as long as β and β/c are sufficiently small.

It follows from (100) and (105) that

$$\begin{aligned} \frac{dG_\alpha^\beta}{db}(b) + \frac{dH_\alpha^\beta}{db}(b) \geq 2(\alpha-1)G_\alpha^\beta(b) - (2\alpha)e^{2(\alpha-1)b}(\beta^2 e^{-2b})B \\ - (2\alpha-1)e^{-kb}\left(\frac{L\,(1+\mathrm{Error})}{be^{-kb}}\right)^{2\alpha}, \end{aligned}$$

and hence the sum on the left is positive unless

$$\begin{aligned} 2(\alpha-1)G_\alpha^\beta(b) - (2\alpha)e^{2(\alpha-1)b}(\beta^2 e^{-2b})B \\ \leq (2\alpha-1)e^{-kb}\left(\frac{L\,(1+\mathrm{Error})}{be^{-kb}}\right)^{2\alpha}, \end{aligned}$$

or by (84),

$$\begin{aligned} (\alpha-1)e^{2(\alpha-1)b}\int_{D_1}(e^{-2b}\beta^2 + |dg|^2)^\alpha d\bar{x}d\bar{y} - \mathrm{Error} \\ \leq (2\alpha-1)e^{-kb}\left(\frac{L\,(1+\mathrm{Error})}{be^{-kb}}\right)^{2\alpha}, \end{aligned} \tag{106}$$

both error terms being small if β and β/c are sufficiently small.

We now proceed as in the case $\beta = 0$, using (106) instead of (82). We conclude as in the preceding section that if (67) does not hold, then

$$\frac{dG_\alpha^\beta}{db} + \frac{dH_\alpha^\beta}{db} > 0$$

whenever α is sufficiently close to one, under the assumption that we can choose β and β/c to be arbitrarily small.

4.6 Proof of the Main Theorem

Case I: only one bubble forms. We now compare the models of the previous sections with a sequence $\{(f_m, \omega_m)\}$ of critical points for E_{α_m} as described in §1.2, under the assumption that only one bubble forms. In this case, we have a single sequence of bubble disks $\{B_m\}$ and a single sequence of necks $\{A_m\}$. The rescalings g_m and h_m of the restrictions of f_m to B_m and A_m converge to a limit harmonic two sphere $g : S^2 \to M$ and a cylinder parametrization

$$h = \gamma \circ \pi : [0,1] \times S^1 \to [0,1] \to M$$

of a geodesic γ. Our assumption is that the geodesic γ has length bounded below by $L = L_0$ in accordance with (10).

To apply the model of §4.2, we take η to be the radius η_m of of the bubble region B_m, but we need to describe what corresponds to the disk of radius ϵ. We have some freedom of choice of ϵ because of the nonzero length of the curves connecting bubbles to base. We will set

$$\epsilon = \epsilon'_m = \exp\left(\zeta \log \eta_m + (1-\zeta)\log \epsilon_m\right), \tag{107}$$

where $\zeta > 0$ is very close to zero. The length of the curve parametrized by the corresponding annular region approaches $(1-\zeta)L$ as $\alpha \to 1$.

We recall that after rescaling to a unit disk, we replace E_α by E_α^β, where $\beta = \epsilon$, the radius of the outer disk in the model. The choice (107) allows us to take β and β/c small, as needed in the estimates in §4.4 and §4.5. Recalling that

$$b_m = -[\log \eta_m - \log \epsilon'_m], \quad \text{we see that} \quad \beta = \epsilon = \epsilon_m e^{-\zeta b_m},$$

which becomes arbitrarily small as $b_m \to \infty$. Moreover, if follows from (57) that

$$\frac{\beta}{c} \le \frac{b\beta}{(1-\zeta)L} \le \frac{b\epsilon_m e^{-\zeta b}}{(1-\zeta)L} \longrightarrow 0 \quad \text{as} \quad b \to \infty.$$

Thus we can indeed ensure that

$$\frac{dG_\alpha^\beta}{db} + \frac{dH_\alpha^\beta}{db} < 0$$

unless the estimate (66) of the Main Theorem is satisfied. In particular, we can arrange that f_m is close to an element f'_m of the model space $\mathrm{Map}_\epsilon^\star(\Sigma, M)$ where ϵ is given by (107) and $\eta = \eta_m$, in the following sense:

1. the restriction of f_m $D_\eta(p)$ is $L_1^{2\alpha}$ close to f'_m, and
2. the restriction of f_m to $D_\epsilon(p)-D_\eta(p)$ is C^0 and $L_1^{2\alpha}$ close to a parametrization (as discussed in §3.2) of a smooth curve C of length at least $L/2$.

Suppose that $\{\psi_s : s \in \mathbb{R}\}$ is a smooth family of piecewise smooth diffeomorphisms of Σ with ψ_0 the identity, each ψ_s being continuous and smooth on each piece,

$$\Sigma - D_\epsilon(p), \quad D_\epsilon(p) - D_\eta(p) \quad \text{and} \quad D_\eta(p).$$

Since f_m is a critical point for E_{α_m}, $s = 0$ must be a critical point for the map $s \mapsto E(f_m \circ \psi_s)$. Assuming that η does not satisfy the estimate of the Scaling Theorem, we claim that when m is sufficiently large, we will construct such a family of diffeomorphisms such that

$$\frac{d}{ds}(E_\alpha^\beta(f_m \circ \psi_s))\bigg|_{s=0} \neq 0. \tag{108}$$

Recall that since $f'_m \circ \psi_s$ is an element of the model space $\mathrm{Map}^\star_\epsilon(\Sigma, M)$, the restriction of f'_m to $D_\epsilon(p) - D_\eta(p)$ is of the form $f'_m(u,\theta) = \gamma \circ \phi(u)$ for some diffeomorphism $\phi : [a,b] \to [0,L]$. We construct the family of diffeomorphisms of Σ by stipulating that the restriction of each ψ_s to $\Sigma - D_\epsilon(p)$ is the identity, that $\psi_s|D_\eta(p)$ is a rescaling which sends the disk of radius η about p to the disk of radius $e^{-s}\eta$, and that in terms of the coordinates (u,θ) on $D_\epsilon(p)$, the restriction of ψ_s to $D_\epsilon(p) - D_{e^{-s}\eta}(p)$ is of the form $\psi_s(u,\theta) = \phi_s(u)$, where $\gamma \circ \phi_s(u)$ is the parametrization which gives a critical point for the α-energy with Dirichlet boundary conditions. With this choice of ψ_s, it follows from § 4.4 that the model f'_m satisfies

$$\frac{d}{ds}(E_\alpha^\beta(f'_m \circ \psi_s))\bigg|_{s=0} \neq 0.$$

The nature of the convergence of f_m to the model f'_m now implies that (108) holds, finishing the proof of the Main Theorem in the case of one bubble.

Case II: the general case. Given a sequence $\{(f_m, \omega_m)\}$ with convergence properties described in §1.2, we choose an outermost essential sequence of bubble disks $m \mapsto B_{i_1,\ldots,i_k;m}$, and choose a sequence $m \mapsto B_{i_1,\ldots,i_j;m}$ with $j \leq k$ such that the corresponding $h_{i_1,\ldots i_j;m}$'s approach a curve of length $\geq L_1$, where $L_1 = L/k$. We let p be the center of $B_{i_1,\ldots,i_j;m}$, let $\eta_{i_1,\ldots,i_j;m}$ be the radius of $B_{i_1,\ldots,i_j;m}$ and let $\epsilon_{i_1,\ldots,i_j;m}$ be the radius of $A_{i_1,\ldots,i_j;m} \cup B_{i_1,\ldots,i_j;m}$. We apply the model this time with $\eta = \eta_{i_1,\ldots,i_j;m}$ and

$$\epsilon = \epsilon'_{i_1,\ldots,i_j;m} = \exp\left(\zeta \log \eta_{i_1,\ldots,i_j;m} + (1-\zeta)\log \epsilon_{i_1,\ldots,i_j;m}\right),$$

with ζ very close to zero. Of course, the restriction of f_m to D_η may bubble into several minimal two-spheres, some being ghosts, connected by geodesics. Similarly, the restriction of f_m to $\Sigma - D_\epsilon$ may converge to a base connected by geodesics to a collection of minimal two-spheres. Our use of the model focuses on one neck at a time.

We can now repeat the argument of Case I with virtually no change.

4.7 Stretching part of one neck

Finally, we note that a slightly different version of the Main Theorem holds, in which we focus on the stretching of part of one of the necks, ignoring the structure of the remaining portion of the maps. Recall that the Main Theorem implies that necks between bubbles stretch to infinite conformal length as $\alpha \to 1$.

The idea is to apply the argument for the Scaling Theorem to the third of the three cases described in §4.2. Our goal is to study the restriction of a collection of ω_m-harmonic maps $f_m : \Sigma \to M$ to annuli $A_m \subset \Sigma$ of the form

$$A_m = D_{\epsilon_m} - D_{\eta_m},$$

where

1. D_{ϵ_m} is one of the disks $D_{i_1,\ldots,i_k;m}$ appearing in the bubble tree construction, and
2. D_{η_m} contains $B_{i_1,\ldots,i_k;m}$, and η_m is chosen so that the corresponding rescaled maps of the annulus, $h_m : [0,1] \times S^1 \to M$, converge to a parametrization of a curve of length L.

Thus the inner disk D_η may parametrize a collection of bubbles connected by annuli converging to geodesics, and the complement of the outer disk $\Sigma - D_\epsilon$ may parametrize not only the base in the Parker-Wolfson bubble tree, but also several of the bubbles connected by annuli converging to geodesics. Our goal is to understand the behavior of the function

$$\zeta : (0,\infty) \longrightarrow \mathrm{Map}^\star_\epsilon(\Sigma, M), \quad \zeta(b) = (f_b, p, \eta(b)), \quad \eta(b) = e^{-b}\epsilon,$$

described in in the paragraph containing (70). We let

$$b_m = -(\log \epsilon_m - \log \eta_m).$$

By modifying the proof given in §4.6, we can then establish:

Neck Stretching Theorem. *Suppose that M be a compact Riemannian manifold, that Σ is a closed surfaces of genus $g \geq 1$ and that $\{(f_m, \omega_m)\}$ is a sequence of (α_m, ω_m)-harmonic maps from Σ into M with ω_m converging to a limit conformal structure ω_∞. If $\sigma = 1/2\alpha$ and b_m is defined as above, then*

$$b_m(\alpha_m - 1)^\sigma > c_1 L \quad \Rightarrow \quad \frac{d}{db}(E_{\alpha_m,h}(\zeta(b))\Big|_{b_m} < 0,$$

and

$$b_m(\alpha_m - 1)^\sigma < c_0 L \quad \Rightarrow \quad \frac{d}{db}(E_{\alpha_m,h}(\zeta(b))\Big|_{b_m} > 0,$$

when α_m is sufficiently close to one.

5 Appendix

We now prove that the α-energy $E_\alpha : \mathcal{M}^{(2)}(T^2, M) \longrightarrow \mathbb{R}$ satisfies Condition C.

To prove this, we recall that in the case where Σ is a torus, the Teichmüller space $\mathcal{T}$ is the upper half plane, and after a change of basis we can arrange that an element $\omega \in \mathcal{T}$ lies in the fundamental domain

$$D = \{u + iv \in \mathbb{C} : -(1/2) \leq u \leq (1/2), u^2 + v^2 \geq 1\} \tag{109}$$

for the action of the mapping class group $\Gamma = SL(2, \mathbb{Z})$. The moduli space $\mathcal{R}$ is obtained from D by identifying points on the boundary. The complex torus corresponding to $\omega \in \mathcal{T}$ can be regarded as the quotient of $\mathbb{C}$ by the abelian subgroup generated by d and ωd, where d is any positive real number, or alternatively, this torus is obtained from a fundamental parallelogram spanned by d and ωd by identifying opposite sides. The fundamental parallelogram of area one can be regarded as the image of the unit square $\{(t_1, t_2) \in \mathbb{R}^2 : 0 \leq t_i \leq 1\}$ under the linear transformation

$$\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{v}} \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix},$$

where $z = x + iy$ is the usual complex coordinate on $\mathbb{C}$. A straightforward calculation gives a formula for the usual energy

$$\begin{aligned} E(f, \omega) &= \frac{1}{2} \int_P \left(\left| \frac{\partial f}{\partial x} \right|^2 + \left| \frac{\partial f}{\partial y} \right|^2 \right) dx dy \\ &= \frac{1}{2} \int_P \left(v \left| \frac{\partial f}{\partial t_1} \right|^2 + \frac{1}{v} \left| \frac{\partial f}{\partial t_2} - u \frac{\partial f}{\partial t_1} \right|^2 \right) dt_1 dt_2, \end{aligned}$$

P denoting the image of the unit square. The only way that ω can approach the boundary of Teichmüller space while remaining in the fundamental domain D is for $v \to \infty$. The rank two condition implies that the maps $t_1 \mapsto f(t_1, b)$ must be homotopically nontrivial, and hence the length in M of $t_1 \mapsto f(t_1, b)$ is bounded below by a positive constant c. This implies that

$$\begin{aligned} E(f, \omega) &\geq \frac{1}{2} \int_0^1 \int_0^1 v \left| \frac{\partial f}{\partial t_1} \right|^2 dt_1 dt_2 \\ &\geq \frac{v}{2} (\text{average length of } t_1 \mapsto f(t_1, b))^2 \geq \frac{c^2 v}{2} \end{aligned} \tag{110}$$

by the Cauchy-Schwarz inequality, and hence $E_\alpha(f, \omega)$ (which is $\geq E(f, \omega)$) must approach infinity.

Suppose now that $[f_i, \omega_i]$ is a sequence of points in $\mathcal{M}^{(2)}(\Sigma, M)$ on which E_α is bounded and for which $\|dE_\alpha([f_i, \omega_i])\| \to 0$, and for each i, $(f_i, \omega_i) \in \text{Map}(T^2, M) \times \mathcal{T}$ is a representative for $[f_i, \omega_i]$. Then the projection $[\omega_i] \in \mathcal{R}$ of

$\omega_i \in \mathcal{T}$ is bounded, and must therefore have a subsequence which converges to an element $[\omega_\infty] \in \mathcal{R}$. Hence there are elements $\phi_i \in \Gamma$ such that a subsequence of $\phi_i^* \omega_i$ converges to an element $\omega_\infty \in \mathcal{T}$. Then $E_{\alpha,\omega_\infty}(f_i \circ \phi_i)$ is bounded and $\|dE_{\alpha,\omega_\infty}(f_i \circ \phi_i)\| \to 0$, so by Condition C for E_{α,ω_∞}, a subsequence of $\{(f_i \circ \phi_i, \phi_i^* \omega_i)\}$ converges to a critical point for E_α on $\mathrm{Map}(T^2, M) \times \mathcal{T}$. This establishes Condition C for the function E_α.

References

[1] N. Aronszajn, *A unique continuation theorem for solutions of elliptic partial differential equations or inequalities of second order*, J. Math. Pure Appl. **36** (1957), 235-249.

[2] R. Bott, *Lectures on Morse theory, old and new*, Bull. Amer. Math. Soc. **7** (1982), 331-358.

[3] J. Chen and G. Tian, *Compactification of moduli space of harmonic mapping*, Comm. Math. Helv. **74** (1999), 201-237.

[4] A. Fraser, *On the free boundary variational problem for minimal disks*, Comm. Pure Appl. Math. **53** (2000), 931-971.

[5] D. Gromoll and W. Meyer, *Periodic geodesics on compact Riemannian manifolds*, J. Differential Geometry **3** (1969), 493-510.

[6] M. Gromov, *Homotopical effects of dilatation*, J. Differential Geometry **13** (1978), 303-310.

[7] W. Jäger and H. Kaul, *Uniqueness and stability of harmonic maps and their Jacobi fields*, Manuscripta Math. **28** (1979), 269-291.

[8] J. Jost, *Two-dimensional geometric variational problems*, John Wiley and Sons, New York, 1991.

[9] D. McDuff and D. Salamon, *J-holomorphic curves and symplectic topology*, Amer. Math. Soc., Providence Rhode Island, 2004.

[10] J. Milnor, *Morse theory*, Annals of Math. Studies **51**, Princeton Univ. Press, Princeton, NJ, 1963.

[11] J. Milnor, *Lectures on the h-cobordism theorem*, Princeton Univ. Press, Princeton, NJ, 1965.

[12] R. Palais, *Lusternik-Schnirelmann theory on Banach manifolds*, Topology **5** (1966), 115-132.

[13] R. Palais, *Critical point theory and the minimax principle*, Proc. Symposia in Pure Math. **15** (1970), 185-212.

[14] T. Parker, *Bubble tree convergence for harmonic maps curvature flow*, J. Differential Geometry **44** (1996), 595-633.

[15] T. Parker and J. Wolfson, *Pseudoholomorphic maps and bubble trees*, J. Geometric Analysis **3** (1993), 63-98.

[16] J. H. Sampson, *Some properties and applications of harmonic mappings*, Annales Scientifiques de l'École normale Supérieure **11** (1978), 211-228.

[17] J. Sacks and K. Uhlenbeck, *The existence of minimal immersions of* 2-*spheres*, Annals of Math. **113** (1981), 1-24.

[18] J. Sacks and K. Uhlenbeck, *Minimal immersions of closed Riemann surfaces*, Trans. Amer. Math. Soc. **271** (1982), 639-652.

[19] R. Schoen and S. T. Yau, *Existence of incompressible minimal surfaces and the topology of three dimensional manifolds with non-negative scalar curvature*, Annals of Math. **110** (1979), 127-142.

[20] R. Schoen and S. T. Yau, *Lectures on harmonic maps*, International Press, Boston, 1997.

[21] M. Schwarz, *Morse homology*, Birkhäuser Verlag, Basel, 1993.

[22] S. Smale, *On the Morse index theorem*, J. Math. Mech. **14** (1965), 1049-1055; Corrigendum **16** (1967), 1069-1070.

[23] K. Uhlenbeck, *Morse theory on Banach manifolds*, J. Functional Analysis **10** (1972), 430-445.