COMMUNICATIONS IN
ANALYSIS AND GEOMETRY
Volume 21, Number 5, 917–955, 2013

On the Sacks–Uhlenbeck flow of Riemannian surfaces

Min-Chun Hong and Hao Yin

In this paper, we study an α-flow for the Sacks–Uhlenbeck functional on Riemannian surfaces and prove that the limiting map of the α-flow is a weak solution to the harmonic map flow. By an application of the α-flow, we present a simple proof of an energy identity of a minimizing sequence in each homotopy class.

1. Introduction

Suppose that M is a Riemannian manifold and N is a closed manifold embedded in $\mathbb{R}^k$. A critical point u of the Dirichlet energy

$$E(u) = \int_M |\nabla u|^2 \, dv$$

is called a harmonic map.

Harmonic maps between Riemannian manifolds can be thought of as a natural generalization of geodesics, minimal surfaces and harmonic functions. A fundamental question on harmonic maps is: given a smooth map u_0 from M to N, does there exists a smooth harmonic map representative in the homotopy class of u_0? In a pioneering work [7], Eells and Sampson answered this question in the case that the sectional curvature of N is nonpositive by introducing the harmonic map flow to deform an initial map to a harmonic map in the same homotopy class. In general, one cannot expect to have the existence of smooth harmonic maps into general target manifolds due to the fact that singularities do occur. From now on, we assume that M is a closed Riemannian surface; i.e., $\dim M = 2$. Under certain topological conditions on N, the existence of minimizing harmonic maps in a homotopy class was proved by Lemaire [12] and Schoen and Yau [22]. In a well-known paper [19], Sacks and Uhlenbeck established many existence results of minimizing harmonic maps in their homotopy classes by introducing a family of

functionals

$$E_\alpha(u) = \int_M (1 + |\nabla u|^2)^\alpha dv$$

for $\alpha > 1$. The α-functional E_α is now called the 'Sacks–Uhlenbeck functional' and satisfies the Palais–Smale condition, so one can prove existence of critical points of E_α by either minimizing the energy functional or using Morse theory. Moreover, a bubbling phenomenon occurs by studying the limits of the critical points of E_α as $\alpha \to 1$ (see Section 4 in [19]). The method of the α-functional has been very useful in studying many other problems. For example, Struwe in [23] used the α-functional to show that for any closed surface $S \subset \mathbf{R}^3$, which is diffeomorphic to S^2, there is a non-trivial minimal surface intersecting S orthogonally along its boundary.

On the other hand, Struwe [24] proved the global existence of the weak solution to the harmonic map flow and that the solution to the flow converges to a harmonic map as $t \to \infty$. Chang et al. [2] constructed an example where the harmonic map flow blows up at finite time, so in general the limiting harmonic map under the flow may not be in the same homotopy class as the initial map.

In this paper, we study an α-flow for the perturbed energy E_α in the same homotopy class of the initial map. More precisely, we consider the following evolution problem:

$$\partial_t u = \triangle_M u + (\alpha - 1)\frac{\nabla |\nabla u|^2 \cdot \nabla u}{1 + |\nabla u|^2} + A(u)(\nabla u, \nabla u), \tag{1.1}$$

with $u(x, 0) = u_0$, where $\triangle_M$ is the Laplacian operator with respect to the Riemannian metric of M and A is the second fundamental form of N. We call (1.1) the Sacks–Uhlenbeck flow (or α-flow). We would like to point out that the α-flow is not the standard gradient flow, but the flow has some analytic advantages.

Firstly, we have

Theorem 1.1. *Let $u_0 : M \to N$ be a smooth map. For sufficiently small $\alpha - 1$, there exists a unique global smooth solution $u_\alpha(x, t)$ to the evolution problem (1.1) in $M \times [0, \infty)$ with initial value u_0. Moreover, for any $t_i \to \infty$, $u_\alpha(\cdot, t_i)$ converges smoothly to a limit map u_α, which is a critical point of the Sacks–Uhlenbeck functional.*

For each $\alpha > 1$, let $u_\alpha(x, t)$ be the global smooth solution to (1.1) in the same homotopy class with the initial map u_0 from M to N. It is a nature

problem to study the limiting behaviour of the solution $u_\alpha(x,t)$ as $\alpha \to 1$ as in [19]. Thus, we prove

Theorem 1.2. *(i) As $\alpha \to 1$, the solution u_α converges smoothly to u on $M \times [0,\infty) \setminus \Sigma$, where the concentration set Σ is a closed set defined by*

$$\Sigma = \bigcap_{0<R<R_M} \left\{ z \in M \times [0,\infty) | \quad \liminf_{\alpha \to 1} \Psi_R^\alpha(u_\alpha, z) \geq \varepsilon_0 \right\}$$

for some ε_0 to be determined. For the precise definition of Ψ_R^α, see Section 3.

(ii) For any two positive t_1 and t_2, $\mathcal{P}^2(\Sigma \cap (M \times [t_1,t_2]))$ is finite, where $\mathcal{P}^2$ denotes the 2-dimensional parabolic Hausdorff measure. Moreover, for any $t \in (0,+\infty)$, $\Sigma_t = \Sigma \cap (M \times \{t\})$ consists of at most finitely many points.

(iii) u is a weak solution to the harmonic map flow.

In fact, this result is similar to the one using the Ginzburg–Landau flow approximation. In [4], Chen and Struwe used the Ginzburg–Landau flow approximation to construct a global weak solution of the harmonic map flow for any dimension larger than two. Further results obtained by the Ginzburg–Landau flow were discussed by Lin and Wang in [16]. The method of the Ginzburg–Landau flow is very powerful for showing the global existence and partial regularity of a weak solution to the harmonic map flow, but it seems that the flow approximation loses control on the topological quantity of maps. The advantage of the Sacks–Uhlenbeck flow approximation is that the solution u_α remains in the same homotopy class of u_0, therefore it seems that the Sacks–Uhlenbeck flow provides a nice geometric picture and can be used to produce some geometric applications (see Theorem 1.4 below).

With Theorem 1.2 in hand, it is very interesting to ask whether the limiting solution u of the Sacks–Uhlenbeck flow is the global weak solution $\hat{u}$ constructed by Struwe in [24]. Although we do not have a complete answer to the question, we can compare some properties of u with the solution $\hat{u}$ in the following:

Theorem 1.3. *(i) The first concentration time for u_α as $\alpha \to 1$, i.e.,*

$$T = \inf_{(x,t)\in\Sigma} t$$

is the same as the first singular time of the weak solution $\hat{u}$ constructed by Struwe. Hence, u_α converges smoothly to the solution $\hat{u}$ in $M \times [0,T)$.

(ii) Consider the following two limits of measures:

$$|\nabla \hat{u}(\cdot,t)|^2 dv \to |\nabla \hat{u}(\cdot,T)|^2 dv + \sum_{i=1}^{\hat{k}} \hat{m}_i \delta_{\hat{p}_i} \quad \text{as } t \to T,$$

and

$$(1+|\nabla u_\alpha(\cdot,T)|^2)^\alpha dv \to (1+|\nabla u(\cdot,T)|^2)\, dv + \sum_{i=1}^{k} m_i \delta_{p_i} \quad \text{as } \alpha \to 1.$$

Then $k = \hat{k}$, $p_i = \hat{p}_i$ and $\hat{m}_i \geq m_i$.

Without the assumption of the energy inequality, weak solutions to the heat flow for harmonic maps may not always be unique (see [1] and [26]), so we cannot exclude the possibility that the weak solution u is different from the solution $\hat{u}$ constructed in [24] after the first singular time T. On this issue, we would like to make the following remark: Consider maps from S^2 to itself. According to [2], there exists an initial value map u_0 of degree three such that Struwe's solution of the harmonic map flow blows up at some finite time T and then the homotopy class of the solution is changed for $t > T$. Hence, the solutions u_α of the α-flow cannot converge strongly to Struwe's solution after T. A natural question is what we can say about Σ (See [13]).

In the final part of this paper, we apply the α-flow to study a minimizing sequence in a given homotopy class. Let u_i be a sequence of smooth maps minimizing $E(u) = \int_M |\nabla u|^2\, dv$ in a fixed homotopy class of maps. Since u_i is bounded in $W^{1,2}$, there is a weak limit u in $W^{1,2}(M,N)$. In general, u may not be in the same homotopy class, but we can show:

Theorem 1.4. *Let u be the weak limit of the above minimizing sequence $\{u_i\}$. Then it is a harmonic map from M to N and there exist harmonic maps $\omega_k : S^2 \to N$ with $k = 1, \ldots, l$ such that*

$$\lim_{i\to\infty} E(u_i) = E(u) + \sum_{k=1}^{l} E(\omega_k). \tag{1.2}$$

Moreover, if $\pi_2(N)$ is trivial, then u_i converges strongly to u in $W^{1,2}(M,N)$ and u is a minimizer in the homotopy class of u_i.

The last part of this theorem can be compared with Theorem 5.1 of [19], where under the same assumption that $\pi_2(N)$ is trivial, the existence of a

minimizer in each homotopy class is proved (see also [12, 22]). We improve the result a little by obtaining that every smooth minimizing sequence converges to such a minimizer. In the proof, we use the α-flow to modify the original minimizing sequence and study the blow-up of the new sequence as in Theorem 5.1 of [19]. The energy identity of minimizers of the Sacks–Uhlenbeck functional was implicitly established by Chen and Tian in [3]. Our proof of the energy identity is different from the one in [3]. Although the result of Theorem 1.4 may be regarded as a consequence of the theory developed by Duzaar and Kuwert [6], the advantage is that we can avoid the use of the concept of "weak homotopy class" in the Sobolev space $W^{1,2}(M,N)$, which is formulated using big machinery. Furthermore, we would like to mention that the related energy identity of critical points of the Sacks–Uhlenbeck functionals was recently discussed by Li and Wang [14], and by Lamm [11].

Remark 1.5. After we finished the first version of this paper, Yuxiang Li informed us that in [15] they used a similar idea to show the energy identify for a sequence of minimizers of E_α as $\alpha \to 1$. In fact, while Li and Wang in [15] used a reduction procedure of Ding and Tian [5], we use the bubble tree construction of Parker in [18] to find the connecting geodesics, so our approach is different. Our result is also stronger than theirs.

Finally, we would like to mention some results about the approximation method of harmonic maps in a homotopy class for higher dimensional manifolds (e.g. see [9, 28]).

The paper is organized as follows. In Section 2, we prove Theorem 1.1. In Section 3, we give a proof of Theorem 1.2. In Section 4, we complete a proof of Theorem 1.3 except for the energy identity, which is proved in Section 5.

2. Global existence and convergence of α-flow

This section is devoted to the proof of Theorem 1.1. The system of the α-flow (1.1) is a non-linear parabolic system. For a smooth initial value u_0, there is a local smooth solution to the system (see below appendix for details in Appendix A); i.e., there exists $T > 0$ and a smooth solution $u(x,t)$ defined on $[0,T)$. In fact, the α-flow (1.1) can be written as follows:

$$(1+|\nabla u|^2)^{\alpha-1}\partial_t u = \operatorname{div}\Big((1+|\nabla u|^2)^{\alpha-1}\nabla u\Big) + (1+|\nabla u|^2)^{\alpha-1}A(u)(\nabla u,\nabla u). \tag{2.1}$$

The first global estimate is:

Lemma 2.1. *Let $u(x,t)$ be a solution to the α-flow in $M \times [0,T)$. Then we have*

$$E_\alpha(u(\cdot,0)) = E_\alpha(u(\cdot,t)) + \int_0^t \int_M (1+|\nabla u|^2)^{\alpha-1} |\partial_t u|^2 \, dvdt$$

for any $t \in [0,T)$.

Proof. The result follows from multiplying (2.1) by $\partial_t u$ and integrating by parts. □

The second estimate is a local version of the energy inequality:

Lemma 2.2. *Let u be a solution to the α-flow. For any constant $R > 0$ and any point $x \in M$, we have*

(2.2)
$$\int_{B_R(x)} (1+|\nabla u(\cdot,t_2)|^2)^\alpha dv \le \int_{B_{2R}(x)} (1+|\nabla u(\cdot,t_1)|^2)^\alpha dv + C\frac{t_2-t_1}{R^2}E_0$$

for all $t_1, t_2 \in [0,T)$ with $t_1 \le t_2$, where E_0 is an upper-bound of the overall energy.

Proof. Let φ be a cut-off function supported in $B_R(x)$ with $\varphi \equiv 1$ on $B_{R/2}(x)$. Multiplying Equation (2.1) by $\partial_t u \varphi^2$, we obtain

$$\int_{B_R(x)} (1+|\nabla u(\cdot,t_2)|^2)^\alpha dv \le \int_{B_{2R}(x)} (1+|\nabla u(\cdot,t_1)|^2)^\alpha dv + C\frac{t_2-t_1}{R^2}E_0.$$

□

One of key steps in proving Theorem 1.1 is to derive a Bochner-type formula. Let (g_{ij}) be a Riemannian metric on M. Then

Lemma 2.3. *(Bochner's type formula) Let $u(x,t)$ be a smooth solution to the α-flow (1.1) and set $e(u) := |\nabla u|^2$. Then, for $\alpha - 1$ sufficiently small, we have*

$$\begin{aligned}(2.3)\quad \partial_t e(u) - \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x_i}\left(\sqrt{|g|}\left(g^{ij} + \frac{(\alpha-1)}{1+|\nabla u|^2} g^{ik}\frac{\partial u^\beta}{\partial x_k} g^{jl}\frac{\partial u^\beta}{\partial x_l}\right)\frac{\partial e(u)}{\partial x_j}\right) \\ \le Ce(u)(e(u)+1),\end{aligned}$$

where the constant C does not depend on α and u.

Proof. In a neighbourhood of each point $x \in M$, we can choose an orthonormal frame $\{e_1, e_2\}$. We denote by ∇_i the first covariant derivative with respect to e_i and by u_{ji} the second covariant derivatives of u and so on. In a local frame, we have

$$\nabla_j e(u) = 2u_k^\gamma u_{kj}^\gamma, \quad |\nabla^2 u|^2 = \sum_{k,i,\gamma} \left|u_{ki}^\gamma\right|^2.$$

Then we have

$$\begin{aligned}
&\nabla_i \left(\left(\delta_{ij} + 2(\alpha-1)\frac{u_i^\beta u_j^\beta}{1+|\nabla u|^2} \right) \nabla_j \, \mathrm{e}(u) \right) \\
&= 2\nabla_i \left(u_k^\gamma u_{ki}^\gamma + 2(\alpha-1)\frac{u_i^\beta u_j^\beta u_k^\gamma u_{kj}^\gamma}{1+|\nabla u|^2} \right) \\
&\geq 2|\nabla^2 u|^2 + 2u_k^\gamma u_{iik}^\gamma + 4(\alpha-1)\nabla_i \left(\frac{u_i^\gamma u_k^\gamma u_j^\beta u_{kj}^\beta}{1+|\nabla u|^2} \right) - Ce(u) \\
&\geq \frac{3}{2}|\nabla^2 u|^2 + 2u_k^\gamma u_{iik}^\gamma + 4(\alpha-1)\frac{u_i^\gamma u_k^\gamma u_j^\beta u_{kji}^\beta}{1+|\nabla u|^2} - Ce(u) \\
&\geq |\nabla^2 u|^2 + 2u_k^\gamma \nabla_k \left(u_{ii}^\gamma + 2(\alpha-1)\frac{u_i^\gamma u_j^\beta u_{ji}^\beta}{1+|\nabla u|^2} \right) - Ce(u).
\end{aligned}$$

Here we have twice used the Ricci identity for switching third-order derivatives. Using (1.1), we have

$$\begin{aligned}
&\frac{\partial}{\partial t} e(u) - \nabla_i \left(\left(\delta_{ij} + 2(\alpha-1)\frac{u_i^\beta u_j^\beta}{1+|\nabla u|^2} \right) \nabla_j e(u) \right) \\
&\quad \leq -|\nabla^2 u|^2 - 2u_k^\gamma \nabla_k \left(A^\gamma(u)(\nabla u, \nabla u) \right) + Ce(u) \\
&\quad \leq Ce(u)(e(u)+1).
\end{aligned}$$

This proves our claim. □

Although the Bochner formula is motivated by the result of Uhlenbeck [27] on p-harmonic functions, it seems that this type of Bochner's formula is new. Using the formula, we can show a small ϵ-regularity estimate by following a method of Schoen [21] and Struwe [25]. More precisely, we have

Lemma 2.4. *Let $u(x,t)$ be a solution of the α-flow in $M \times [0,T)$ for a $T > 0$. There is a small constant $\varepsilon_0 > 0$ such that if*

$$\sup_{0\le t<T} \int_{B_{r_0}(x_0)} |\nabla u(x,t)|^2 \, dx \le \varepsilon_0$$

for a point $(x_0, t_0) \in M \times (0,T)$ with $r_0^2 \le T$, then

$$|\nabla u(x,t)| \le \frac{C}{r_0^2}, \quad \forall (x,t) \in B_{r_0/2}(x_0) \times \left[t_0 - \frac{r_0^2}{4}, t_0\right],$$

where the constant C depends not on (x,t).

Proof. Set

$$P_r = \{(x,t) | x \in B_r(x_0) \quad \text{and } t_0 - r^2 \le t \le t_0\}.$$

We can find $\rho \in [0, r_0]$ such that

$$(r_0 - \rho)^2 \sup_{P_\rho} e(u) = \max_{\sigma \in [0, r_0]} \left\{ (r_0 - \sigma)^2 \sup_{P_\sigma} e(u) \right\}.$$

Let (x_1, t_1) be the point in P_ρ such that

$$e \stackrel{\triangle}{=} e(u)(x_1, t_1) = \sup_{P_\rho} e(u).$$

If $e(r_0 - \rho)^2 \le 4$, then

$$\left(\frac{r_0}{2}\right)^2 \sup_{P_{r_0/2}} e(u) \le e(r_0 - \rho)^2 \le 4,$$

which means the lemma is proved. Hence, we may assume that $e(r_0 - \rho)^2 > 4$. Set

$$v(x,t) = u\left(x_1 + \frac{x}{e^{1/2}}, t_1 + \frac{t}{e}\right)$$

for $(x,t) \in B_1(0) \times [-1, 0]$, so v satisfies the scaled α-flow equation

$$\partial_t v = \frac{\operatorname{div}\left(\left(e^{-1} + |\nabla v|^2\right)^{\alpha_k - 1} \nabla v\right)}{(e^{-1} + |\nabla v|^2)^{\alpha_k - 1}} + A(v)(\nabla v, \nabla v).$$

By the scaling invariance of the Dirichlet energy, we have

$$\sup_{t\in[-1,0]}\int_{B_1(0)} e(v)dv \le \varepsilon_0. \tag{2.4}$$

Moreover, we have $e(v)(0,0)=1$ and

$$\sup_{B_1(0)\times[-1,0]} e(v) \le e^{-1}\sup_{P_{\frac{r_0+\rho}{2}}} e(u) \le e^{-1}\frac{e(r_0-\rho)^2}{(r_0-\frac{r_0+\rho}{2})^2} \le 4. \tag{2.5}$$

By Lemma 2.3 and (2.5), we have

$$\frac{\partial}{\partial t}e(v) - \nabla_i(a_{ij}(v)\nabla_j e(v)) \le Ce(v),$$

where

$$a_{ij}(v) = \delta_{ij} + 2(\alpha-1)\frac{v_i^\beta v_j^\beta}{e^{-1}+|\nabla v|^2}.$$

The symmetric matrix $(a_{ij}(v))$ has eigenvalues satisfying the uniform elliptic condition. By a well-known Moser iteration and (2.4), we have

$$1 = e(v)(0,0) \le C\left(\int_{B_1(0)\times[-1,0]} e(v)^2\right)^{1/2}.$$

By (2.5), we obtain

$$1 \le C\int_{B_1(0)\times[-1,0]} e(v) \le C\varepsilon_0,$$

which is impossible if we choose ε_0 small. This proves our claim. □

For each fixed $\alpha > 1$, the small energy assumption in Lemma 2.4 is automatically true because of the Hölder inequality and the uniform boundness of the energy $E_\alpha(u)$. More precisely, For any $t \in [0,T)$, Lemma 2.1 implies that

$$\int_M (1+|\nabla u|^2)^\alpha(t)dv \le C.$$

For each $\varepsilon_0 > 0$, there exists a constant $r_0 > 0$, depending on ε_0 and α, such that for all $x \in M$ and $t < T$, we have

$$\int_{B_{r_0}(x)} |\nabla u(\cdot,t)|^2\, dv \le \varepsilon_0.$$

We now complete the proof of Theorem 1.1.

Proof. By the local existence, there is the unique smooth solution $u(x,t)$ of the alpha-flow in $M \times [0,T)$ for some $T > 0$. we will show that the solution u can be extended smoothly to the time T. By Lemma 2.4, we have the C^0-estimate for $\nabla u(x,t)$ for all $t \in [0,T)$. Next, we show higher-order derivative estimates. For $x_0 \in M$ and $T/2 < t_0 < T$, set $P_r = B_r(x_0) \times [t_0 - r^2, t_0]$. Suppose φ is a cut-off function supported in $P_{T/2}$ with $\varphi \equiv 1$ in $P_{T/4}$. Multiplying (1.1) by φ, we obtain

$$\partial_t(u\varphi) - \triangle(u\varphi) = (\alpha - 1)\frac{(\nabla^2(u\varphi), \nabla u)\nabla u}{1 + |\nabla u|^2} + \mathcal{R}[\varphi, u, \nabla u], \tag{2.6}$$

where $\mathcal{R}$ is a term involving $u, \nabla u$ and derivatives of φ. By Lemma 2.4 and the L^p estimate of linear parabolic equations, we have

$$\|\varphi u\|_{W^{2,p}(P_{T/2})} \leq C(\alpha - 1)\left\|\nabla^2(\varphi u)\right\|_{L^p(P_{T/2})} + C,$$

for any $p > 0$. When $\alpha - 1$ is small,

$$\|\varphi u\|_{W^{2,p}(P_{T/2})} \leq C.$$

We then take one more space derivative of (2.6) to get

$$\partial_t(\varphi \nabla u) - \triangle(\varphi \nabla u) = (\alpha - 1)\frac{(\nabla^3(\varphi u), \nabla u)\nabla u}{1 + |\nabla u|^2} + \mathcal{Q}[\varphi, u, \nabla u, \nabla^2 u].$$

Here $\mathcal{Q}$ involves $u, \nabla u, \nabla^2 u$ and derivatives of φ. Using the L^p estimate again, we have (when $\alpha - 1$ is sufficiently small)

$$\|\varphi u\|_{W^{3,p}(P_{T/2})} \leq C,$$

for any p. Therefore,

$$\|\partial_t \nabla(\varphi u)\|_{L^p(P_{T/2})}, \quad \left\|\nabla^2(\varphi u)\right\|_{L^p(P_{T/2})} \leq C.$$

For $p > 3$, the Sobolev embedding theorem yields that

$$\|\nabla(\varphi u)\|_{C^\beta(P_{T/2})} \leq C$$

for some $\beta > 0$. We can now apply the Schauder theory for parabolic equations to obtain higher-order estimates, which means that we can extend the solution $u(x,t)$ smoothly to $t = T$. Then the local existence result for smooth initial data implies $T = \infty$. $\square$

3. The limiting behaviour of the Sacks–Uhlenbeck flow

For a fixed initial map u_0, it follows from Theorem 1.1 that there is a unique global solution u_α to (1.1) in $M \times [0,\infty)$ for each $\alpha > 1$. In this section, we study the limit of the solutions u_α as $\alpha \to 1$. We start with two lemmas. The first one is another Bochner-type formula.

Lemma 3.1. *Let $u(x,t)$ be a classical solution to the α-flow (1.1) and set $e_\alpha(u) := (1+|\nabla u|^2)^\alpha$. If $\alpha - 1$ is sufficiently small, then there is a constant $C > 0$ such that*

(3.1)
$$\frac{\partial}{\partial t} e_\alpha(u) - \nabla_i \left(\left(\delta_{ij} + 2(\alpha-1) \frac{u_i^\beta u_j^\beta}{1+|\nabla u|^2} \right) \nabla_j e_\alpha(u) \right) \le C e_\alpha(u)(1+|\nabla u|^2).$$

Proof. By direct computation and switching indices as in Lemma 2.3, we have

$$\begin{aligned}
&\nabla_i \left(\left(\delta_{ij} + 2(\alpha-1) \frac{u_i^\beta u_j^\beta}{1+|\nabla u|^2} \right) \nabla_j e_\alpha(u) \right) \\
&\quad = 2\alpha \nabla_i \left((1+|\nabla u|^2)^{\alpha-1} u_k^\gamma u_{ki}^\gamma + 2(\alpha-1)(1+|\nabla u|^2)^{\alpha-2} u_i^\gamma u_k^\gamma u_j^\beta u_{kj}^\beta \right) \\
&\quad = 2\alpha \nabla_i \left[u_k^\gamma \nabla_k (u_i^\gamma (1+|\nabla u|^2)^{\alpha-1}) \right] \\
&\quad = 2\alpha \left| u_{ki}^\gamma \right|^2 (1+|\nabla u|^2)^{\alpha-1} + 4\alpha(\alpha-1)(1+|\nabla u|^2)^{\alpha-2} u_{ki}^\gamma u_{jk}^\beta u_i^\gamma u_j^\beta \\
&\qquad + 2\alpha u_k^\gamma \nabla_i \nabla_k (u_i^\gamma (1+|\nabla u|^2)^{\alpha-1}).
\end{aligned}$$

Again, we assume $\alpha - 1$ is small whenever necessary. Using the Ricci identity to switch the order of covariant derivatives, we have

$$\begin{aligned}
&\nabla_i \left(\left(\delta_{ij} + 2(\alpha-1) \frac{u_i^\beta u_j^\beta}{1+|\nabla u|^2} \right) \nabla_j e_\alpha(u) \right) \\
&\quad \ge 2\alpha \left| u_{ki}^\gamma \right|^2 (1+|\nabla u|^2)^{\alpha-1} + 2\alpha u_k^\gamma \nabla_k \nabla_i (u_i^\gamma (1+|\nabla u|^2)^{\alpha-1}) \\
&\qquad - C(1+|\nabla u|^2)^\alpha \\
&\quad \ge \alpha \left| u_{ki}^\gamma \right|^2 (1+|\nabla u|^2)^{\alpha-1} + 2\alpha (1+|\nabla u|^2)^{\alpha-1} u_k^\gamma \nabla_k \frac{\partial u^\gamma}{\partial t} \\
&\qquad + 2\alpha (1+|\nabla u|^2)^{\alpha-1} u_k^\gamma \nabla_k (A^\gamma(u)(\nabla u, \nabla u)) - C(1+|\nabla u|^2)^\alpha \\
&\quad \ge \frac{1}{2} \left| u_{ki}^\gamma \right|^2 (1+|\nabla u|^2)^{\alpha-1} + \frac{\partial}{\partial t} e_\alpha(u) - C e_\alpha(u)(1+|\nabla u|^2).
\end{aligned}$$

Here in the last two inequalities, we have used the equation of u and the fact that $\alpha - 1$ is small. $\square$

The second lemma is a parabolic monotonicity formula. Such a formula for the harmonic map flow was first established by Struwe in [25] for the Euclidean case and by Chen–Struwe in [4] for general case. Let u be a solution to (1.1). Suppose i_M is the injectivity radius of M. For a fixed point x_0 in M, choose the normal coordinates $\{x_i\}$ and a cut-off function φ supported in $B_{i_M}(x_0)$ such that $0 \le \varphi \le 1$ and $\varphi \equiv 1$ in a neighbourhood of p. For $z_0 = (x_0, t_0)$ with some $t_0 \in [0, \infty)$, set

$$G_{z_0}(x,t) = \frac{1}{|t-t_0|} e^{\frac{-|x-x_0|^2}{4|t-t_0|}}$$

and

$$\Psi_\rho^\alpha(u, z_0) = \rho^{2\alpha-2} \int_{T_\rho(t_0)} (1 + |\nabla u|^2)^\alpha G_{z_0} \varphi^2 \sqrt{|g|}\, dxdt,$$

where $\sqrt{|g|}dx$ is the volume element of M and $T_\rho(t_0) = M \times [t_0 - 4\rho^2, t_0 - \rho^2]$.

Lemma 3.2. *Let u be a smooth solution to (1.1) defined on $M \times [0, \infty)$ and $E_\alpha(u(t)) \le E_0$. Then for each $z_0 \in M \times (0, +\infty)$ and for any $0 < r < \rho \le \min\{i_M, 1\}$, we have*

$$\Psi_r^\alpha(u, z_0) \le e^{c(\rho - r)} \Psi_\rho^\alpha(u, z_0) + CE_0(\rho - r),$$

with a uniform constant C depending only on M and N.

Proof. After a translation, we can assume $z_0 = (0, 0)$. For simplicity, we write $\Psi_\rho(u) = \Psi_\rho^\alpha(u, z)$ and $G = G_{(0,0)}$. Letting $u_\rho(y, s) = u(\rho y, \rho^2 s)$, we obtain

$$\Psi_\rho(u) = \int_{T_1} (\rho^2 + g^{ij}(\rho y)\partial_{y_i} u_\rho \partial_{y_j} u_\rho)^\alpha G \varphi^2(\rho y) \sqrt{|g|}(\rho y) dyds.$$

For simplicity, we write $g^{ij} := g^{ij}(\rho y)$ and the same convention applies to φ, $\bar\nabla g^{ij}$ and so on. Here we use $\bar\nabla$ to denote the gradient in local coordinates;

i.e. $(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2})$, and $\tilde{\nabla}$ to denote $(\frac{\partial}{\partial y_1}, \frac{\partial}{\partial y_2})$. Denoting $\frac{\partial}{\partial y_i} u_\rho$ by $u_{\rho,i}$, we have

$$\begin{aligned}
\frac{d}{d\rho}\Psi_\rho(u) &= \int_{T_1} \alpha(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\left(2\rho + y\cdot\bar{\nabla}g^{ij}u_{\rho,i}u_{\rho,j} + 2g^{ij}u_{\rho,i}\frac{du_{\rho,j}}{d\rho}\right) \\
&\quad \cdot G\varphi^2\sqrt{|g|}\,dy\,ds + \int_{T_1}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha}G\frac{d}{d\rho}(\varphi^2\sqrt{|g|})\,dy\,ds \\
&\geq -\int_{T_1} 2\alpha\frac{\partial}{\partial y_j}\left(\sqrt{|g|}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}g^{ij}u_{\rho,i}G\right)\frac{du_\rho}{d\rho}\varphi^2\,dy\,ds \\
&\quad - \int_{T_1} 4\alpha(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}g^{ij}u_{\rho,i}\partial_{y_j}\varphi G\frac{du_\rho}{d\rho}\varphi\sqrt{|g|}dy\,ds \\
&\quad + \int_{T_1} \alpha(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}y\cdot\bar{\nabla}g^{ij}u_{\rho,i}u_{\rho,j}G\varphi^2\sqrt{|g|}\,dy\,ds \\
&\quad + \int_{T_1} (\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha}G2\varphi y\cdot\bar{\nabla}\varphi\sqrt{|g|}\,dy\,ds \\
&\quad + \int_{T_1} (\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha}G\varphi^2\frac{y\cdot\bar{\nabla}\,|g|}{2\,|g|}\sqrt{|g|}\,dy\,ds \\
&= I + II + III + IV + V.
\end{aligned}$$

By the definition of u_ρ, we have

$$\frac{du_\rho}{d\rho} = y\cdot\bar{\nabla}u + 2\rho s\frac{\partial u}{\partial t} = \frac{2s}{\rho}\left(\frac{1}{2s}y\cdot\tilde{\nabla}u_\rho + \frac{\partial u_\rho}{\partial s}\right).$$

Since u_ρ satisfies

$$\begin{aligned}
(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\frac{\partial u_\rho}{\partial s} &= \frac{\partial}{\partial y_j}\left(\sqrt{|g|}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}g^{ij}u_{\rho,i}\right) \\
&\quad + (\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}A(u_\rho)(\tilde{\nabla}u_\rho, \tilde{\nabla}u_\rho),
\end{aligned}$$

we have

$$\begin{aligned}
I &\geq -2\alpha\int_{T_1}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\left(\frac{\partial u_\rho}{\partial s} - \frac{g^{ij}u_{\rho,i}y_j}{2|s|}\right)\frac{du_\rho}{d\rho}G\varphi^2\sqrt{|g|}dy\,ds \\
&\geq \int_{T_1}\frac{4\alpha|s|}{\rho}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\left|\frac{\partial u_\rho}{\partial s} - \frac{y\cdot\tilde{\nabla}u_\rho}{2\,|s|}\right|^2 G\varphi^2\sqrt{|g|}dy\,ds
\end{aligned}$$

$$
\begin{aligned}
&- \int_{T_1} \frac{4\alpha\,|s|}{\rho}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\left|\frac{g^{ij}u_{\rho,i}y_j}{2\,|s|} - \frac{y\cdot\tilde{\nabla}u_\rho}{2\,|s|}\right| \\
&\times \left|\frac{\partial u_\rho}{\partial s} - \frac{y\cdot\tilde{\nabla}u_\rho}{2\,|s|}\right| G\varphi^2\sqrt{|g|}dy\,ds \\
\geq &\int_{T_1} \frac{2\alpha|s|}{\rho}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\left|\frac{\partial u_\rho}{\partial s} - \frac{y\cdot\tilde{\nabla}u_\rho}{2\,|s|}\right|^2 G\varphi^2\sqrt{|g|}dy\,ds \\
&- C\int_{T_1} \frac{1}{\rho}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\left|\frac{g^{ij}u_{\rho,i}y_j}{2} - \frac{y\cdot\tilde{\nabla}u_\rho}{2}\right|^2 G\varphi^2\sqrt{|g|}dy\,ds.
\end{aligned}
$$

Since $\{x_i\}$ are normal coordinates, we have $\left|g^{ij}(\rho y) - \delta^{ij}\right| \leq C\rho\,|y|$. The absolute value of the last term is not bigger than

$$
\begin{aligned}
&C\int_{T_1} \rho^{-1}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\left|g^{ij}(\rho y) - \delta^{ij}\right|^2 |y|^2\,|u_{\rho,i}|^2\,G\varphi^2\sqrt{|g|}dy\,ds \\
&\leq C\int_{T_1} \rho^{-1}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\rho^2\,|y|^4\,|u_{\rho,i}|^2\,G\varphi^2\sqrt{|g|}dy\,ds \\
&\leq C\int_{T_1} (\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha}(G+C)\varphi^2\sqrt{|g|}dy\,ds \\
&\leq C\Psi_\rho(u) + C\max_{-4\leq s\leq -1}\int_M \rho^{2\alpha-2}(1+|\nabla u|^2)^\alpha\varphi^2\sqrt{|g|}dx \\
&\leq C\Psi_\rho(u) + CE_0,
\end{aligned}
$$

if $\rho < 1$. Here we use the fact that for $s \in [-4,-1]$, $|y|^4\,G \leq G + C$. In summary, we have

$$
\begin{aligned}
I \geq &\int_{T_1} \frac{2\alpha|s|}{\rho}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\left|\frac{\partial u_\rho}{\partial s} - \frac{y\cdot\tilde{\nabla}u_\rho}{2\,|s|}\right|^2 G\varphi^2\sqrt{|g|}dy\,ds \\
&- C\Psi_\rho(u) - CE_0.
\end{aligned}
$$

For the remaining terms, we have for a sufficiently small ϵ

$$
\begin{aligned}
|II| \leq &\epsilon\int_{T_1} (\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\frac{4s^2}{\rho}\left|\frac{\partial u_\rho}{\partial s} - \frac{y\cdot\tilde{\nabla}u_\rho}{2\,|s|}\right|^2 G\varphi^2\sqrt{|g|}\,dy\,ds \\
&+ C(\epsilon)\int_{T_1} \rho^{-1}(\rho^2 + g^{ij}u_{\rho,i}u_{\rho,j})^{\alpha-1}\left(g^{ij}u_{\rho,i}\partial_{y_j}\varphi\right)^2\sqrt{|g|}dy\,ds
\end{aligned}
$$

$$
\begin{aligned}
&\leq \epsilon \int_{T_1} (\rho^2 + g^{ij} u_{\rho,i} u_{\rho,j})^{\alpha-1} \frac{4s^2}{\rho} \left| \frac{\partial u_\rho}{\partial s} - \frac{y \cdot \tilde{\nabla} u_\rho}{2\,|s|} \right|^2 G\varphi^2 \sqrt{|g|}\, dy\, ds \\
&\quad + C(\epsilon) \int_{T_1} \rho \sum_k |\partial_{x_k} \varphi|^2 \, (\rho^2 + g^{ij} u_{\rho,i} u_{\rho,j})^\alpha \sqrt{|g|}\, dy\, ds \\
&\leq \frac{1}{2} I + C\Psi_\rho(u) + CE_0 + C \max_{-4 \leq s \leq -1} \int_M \rho^{1+2\alpha-2} (1 + |\nabla u|^2)^\alpha \sqrt{|g|}\, dx \\
&\leq \frac{1}{2} I + C\Psi_\rho(u) + CE_0.
\end{aligned}
$$

Here we use again that $\rho < 1$, and that G is bounded on T_1. There is some constant C depending only on the geometry of M such that $\bar{\nabla} g^{ij}$ and $\bar{\nabla}\,|g|$ are bounded by C. Therefore, we have

$$
|III| + |V| \leq C\Psi_\rho(u) + CE_0
$$

and

$$
\begin{aligned}
|IV| &\leq \frac{1}{2}\Psi_\rho(u) + C \int_{T_1} (\rho^2 + g^{ij} u_{\rho,i} u_{\rho,j})^\alpha \left| y \cdot \bar{\nabla}\varphi \right|^2 G \sqrt{|g|}\, dy\, ds \\
&\leq \frac{1}{2}\Psi_\rho(u) + CE_0,
\end{aligned}
$$

where we used the fact that $|y|^2 G$ is bounded on T_1. In conclusion, we show that for $\rho < 1$,

$$
\frac{d}{d\rho}\Psi_\rho(u) \geq C\Psi_\rho(u) + CE_0.
$$

The lemma follows from integrating this differential inequality. □

With Lemmas 3.1 to 3.2 in hand, we now prove Theorem 1.2.

Proof. (i) Let z_i be a sequence of points in Σ which converges to $\tilde{z} = (\tilde{x}, \tilde{t})$. For any $0 < R < R_M$, by the definition of Σ, we have

$$
\liminf_{\alpha \to 1} R^{2\alpha-2} \int_{T_R(z_i)} \varphi^2 e_\alpha(u_\alpha) G_{z_i} \sqrt{|g|}\, dx\, dt \geq \varepsilon_0.
$$

Since $\int_M e_\alpha(u_\alpha)\sqrt{|g|}dx$ is uniformly bounded by E_0, and G_{z_i} converges to $G_{\tilde{z}}$ uniformly away from $\tilde{z}$, we can take $i \to \infty$ in the above inequality and

switch the order of the limits to get

$$\liminf_{\alpha\to 1} R^{2\alpha-2}\int_{T_R(\tilde{z})}\varphi^2 e_\alpha(u_\alpha)G_{\tilde{z}}\sqrt{|g|}dxdt \geq \varepsilon_0.$$

Since R is arbitrary, we know $\tilde{z}\in\Sigma$, which shows that Σ is closed.

The proof on the convergence of u_α outside Σ is similar to the proof of Theorem 1.1. When α is fixed, we know that there is no concentration for the Dirichlet energy, so we can use the small energy condition with (2.3) to obtain a C^0-bound on the gradient as in Lemma 2.4. However, as α goes to 1, the difference is that there may be a concentration point of the Dirichlet energy. But for $z\notin\Sigma$, we can find $1/2>R>0$ such that

$$\liminf_{\alpha\to 1}\Psi_R^\alpha(u_\alpha,z)<\varepsilon_0.$$

We claim: There is an $\varepsilon_0>0$ which we use in the definition of Σ, a sequence $\{\alpha_k\}$, a constant $\delta>0$ and some constant C depending on N,R,E_0 and ε_0 such that

$$\|\nabla u_{\alpha_k}\|_{C^0(P_{\delta R}(z))}\leq C.$$

Assume $z=(0,0)$ and set $r_1=\delta R$ for some $\delta<1/2$ to be determined. Suppose for all $r,\sigma\in(0,r_1)$, $r+\sigma<r_1$ and $z_0=(x_0,t_0)\in P_r$. We will show that there are sufficiently small R and δ small such that

$$\sigma^{2\alpha_k-4}\int_{P_\sigma(z_0)} e_{\alpha_k}(u_{\alpha_k})\,dv\,dt\leq C\varepsilon_0$$

for all $r>0$, $\sigma>0$ with $r+\sigma<r_1$. To see this, Lemma 3.2 implies

$$\begin{aligned}
(3.2)\qquad &\sigma^{2\alpha_k-4}\int_{P_\sigma(z_0)} e_{\alpha_k}(u_{\alpha_k})\,dvdt\\
&\leq C\sigma^{2\alpha_k-2}\int_{P_\sigma(z_0)} e_{\alpha_k}(u_{\alpha_k})G_{(x_0,t_0+2\sigma^2)}dvdt\\
&\leq C\sigma^{2\alpha_k-2}\int_{T_\sigma(t_0+2\sigma^2)}\varphi^2 e_{\alpha_k}(u_{\alpha_k})G_{(x_0,t_0+2\sigma^2)}dvdt\\
&\leq Ce^{cR-\sigma}\Psi_R^{\alpha_k}(u_{\alpha_k},(x_0,t_0+2\sigma^2))+CE_0(R-\sigma).
\end{aligned}$$

We can choose R to be small so that the last term is no larger than ε_0.

$$
\begin{aligned}
(3.3)\quad & \Psi_R^{\alpha_k}(u_{\alpha_k},(x_0,t_0+2\sigma^2)) \\
&\le CR^{2\alpha_k-2}\int_{t_0+2\sigma^2-4R^2}^{t_0+2\sigma^2-R^2}\int_M \varphi^2 e_{\alpha_k}(u_{\alpha_k})G_{(x_0,t_0+2\sigma^2)}dvdt \\
&\le CR^{2\alpha_k-2}\int_{-4R^2}^{-R^2}\int_M \varphi^2 e_{\alpha_k}(u_{\alpha_k})G_{(x_0,t_0+2\sigma^2)}dvdt + (\delta R)^2 E_0 R^{-2} \\
&\le C\Psi_R^{\alpha_k}(u_{\alpha_k},(x_0,t_0)) + \varepsilon E_0 + \delta E_0.
\end{aligned}
$$

Here in the last line we used the fact that on T_R, for any $\varepsilon > 0$ we can find δ sufficiently small so that

$$G_{(x_0,t_0+2\sigma^2)} \le CG(x,t) + \varepsilon R^{-2}$$

(which is the exact estimate (2.18) of [4]). Hence, by choosing ε, δ properly, we have

$$\sigma^{2\alpha_k-4}\int_{P_\sigma(x_0,t_0)} e_{\alpha_k}(u_{\alpha_k})dvdt \le C\varepsilon_0 \tag{3.4}$$

for any $(x_0,t_0) \in P_r$ and any two r, σ with $r+\sigma < r_1 = \delta R$.

We can find $\rho \in [0,r_1]$ such that

$$(r_1-\rho)^{2\alpha_k}\sup_{P_\rho} e_{\alpha_k}(u_{\alpha_k}) = \max_{\sigma\in[0,r_1]}\left\{(r_1-\sigma)^{2\alpha_k}\sup_{P_\sigma} e_{\alpha_k}(u_{\alpha_k})\right\}.$$

Let (x_1,t_1) be the point in P_ρ such that

$$e_0 \triangleq e_{\alpha_k}(u_{\alpha_k})(x_1,t_1) = \sup_{P_\rho} e_{\alpha_k}(u_{\alpha_k}).$$

If $e_0(r_1-\rho)^{2\alpha_k} \le 4$, then

$$\left(\frac{r_1}{2}\right)^{2\alpha_k}\sup_{P_{r_1/2}} e_{\alpha_k}(u_{\alpha_k}) \le e_0(r_1-\rho)^{2\alpha_k} \le 4,$$

which means the claim is true. Hence, we may assume that $e_0(r_1-\rho)^{2\alpha_k} > 4$. Set $\lambda = e_0^{\frac{1}{2\alpha_k}}$ and

$$v(x,t) = u_{\alpha_k}\left(x_1+\frac{x}{\lambda}, t_1+\frac{t}{\lambda^2}\right)$$

for $(x,t) \in B_1(0) \times [-1,0]$. Then v satisfies a scaled $\alpha-$flow equation

$$(\lambda^{-2} + |\nabla v|^2)^{\alpha_k - 1}\partial_t v = \operatorname{div}\left((\lambda^{-2} + |\nabla v|^2)^{\alpha_k - 1}\nabla v\right) + (\lambda^{-2} + |\nabla v|^2)^{\alpha_k - 1}A(v)(\nabla v, \nabla v).$$

If we set $e(v) := (\lambda^{-2} + |\nabla v|^2)^{\alpha_k}$, then $e(v)(0,0) = 1$ and

$$\begin{aligned}\sup_{B_1(0)\times[-1,0]} e(v) &\le e_0^{-1} \sup_{P_{\frac{r_1+(2^{1/\alpha_k}-1)\rho}{2^{1/\alpha_k}}}} e_\alpha(u_\alpha)\\ &\le e_0^{-1}\frac{e_0(r_1-\rho)^{2\alpha_k}}{(r_1 - \frac{r_1+(2^{1/\alpha_k}-1)\rho}{2^{1/\alpha_k}})^{2\alpha_k}}\\ &= \left(\frac{2^{1/\alpha_k}}{2^{1/\alpha_k}-1}\right)^{2\alpha_k} \le 5\end{aligned}$$

for k sufficiently large. By Lemmas 3.1 and (3.5), we obtain

$$\frac{\partial}{\partial t}e(v) - \nabla_i(a_{ij}(v)\nabla_j e(v)) \le Ce(v),$$

where (a_{ij}) is a symmetric matrix whose eigenvalues satisfy a uniform elliptic condition. By the well-known Moser iteration again, we have

$$\begin{aligned}1 = e(v)(0,0) &\le C\int_{B_1(0)\times[-1,0]} e(v)\\ &\le C\lambda^{4-2\alpha_k}\int_{t_1-\lambda^{-2}}^{t_1}\int_{B_{1/\lambda}(x_1)}(1+|du_{\alpha_k}|)^{\alpha_k}dvdt\\ &\le C\varepsilon_0.\end{aligned}$$

Here in the last step, we used (3.4) with $\sigma = \frac{1}{\lambda}$ and $(x_0,t_0) = (x_1,t_0)$. Since $\lambda = e_0^{\frac{1}{2\alpha_k}}$ and $e_0(r_1-\rho)^{2\alpha_k} > 4$, it is easy to check that $P_{1/\lambda}(x_1,t_1) \subset P_{r_1}(z_0)$, which is necessary for (3.4) to be true. This is a contradiction if we choose ε_0 small. This proves the claim.

We can establish higher order estimates for u_{α_k} in a smaller neighbourhood as in the proof of Theorem 1.1 so that (i) of Theorem 1.2 follows.

(ii) For any $R > 0$, since $M \times [t^1, t^2]$ is compact, we can find a finite cover of Σ; i.e., $z_i = (x_i, t_i) \in M \times [t^1, t^2]$ such that

$$\Sigma \cap (M \times [t^1, t^2]) \subset \bigcup_{i=1}^{l} Q_{5R}(z_i)$$

and $\{Q_R(z_i)\}$ are disjoint, where $Q_r(z) := B_r(x) \times (t - r^2, t + r^2)$.

Let $\bar{z}_i = (x_i, t_i + R^2)$. For some constant $\delta \in (0, 1/4)$ to be determined later, we have (by definition of Σ and for $\alpha - 1$ sufficiently small; in fact, for each i, $\alpha - 1$ is required to be smaller)

$$\begin{aligned}
(3.5)\qquad \frac{1}{2}\varepsilon_0 &\le (\delta R)^{2\alpha-2} \int_{t_i-4\delta^2R^2}^{t_i-\delta^2R^2} \int_M \varphi^2 e_\alpha(u_\alpha) G_{z_i} \sqrt{|g|}\, dxdt \\
&\le C(\delta) R^{2\alpha-4} \int_{Q_R(z_i)} e_\alpha(u_\alpha)\sqrt{|g|}\, dxdt \\
&\quad + C\delta^{-2} e^{-\frac{1}{16\delta^2}} (\delta R)^{2\alpha-2} \int_{t_i-4\delta^2R^2}^{t_i-\delta^2R^2} \int_M \varphi^2 e_\alpha(u_\alpha) G_{\bar{z}_i} \sqrt{|g|}\, dxdt
\end{aligned}$$

for $1 \le i \le l$. Here in the last step, we use the fact that

$$G_{z_i} \le \delta^{-2} e^{-\frac{1}{16\delta^2}} G_{\bar{z}_i} \text{ in } B_{R_M}(x_i) \times [t_i - 4\delta^2R^2, t_i - \delta^2R^2] \setminus Q_R(z_i).$$

By Lemma 3.2, the second term in the right-hand side of (3.5) can be estimated as

$$\begin{aligned}
C\delta^{-2} e^{-\frac{1}{16\delta^2}} \Psi^\alpha_{\delta R}(u_\alpha, \bar{z}_i) &\le C\delta^{-2} e^{-\frac{1}{16\delta^2}} (\Psi^\alpha_{\sqrt{t^1}/4}(u_\alpha, \bar{z}_i) + E_0) \\
&\le C(E_0, t^1)\delta^{-2} e^{-\frac{1}{16\delta^2}}.
\end{aligned}$$

Hence, we can choose δ small enough so that it is smaller than $\varepsilon_0/4$. Therefore,

$$R^2 \le R^{4-2\alpha} \le C \int_{Q_R} e_\alpha(u_\alpha)\sqrt{|g|}\, dxdt.$$

Since $Q_R(z_i)$ are disjoint, we have

$$\sum_{i=1}^{l} (5R)^2 \le C \sum_{i=1}^{l} \int_{Q_R(z_i)} e_\alpha(u_\alpha)\sqrt{|g|}\, dxdt \le CE_0.$$

By sending R to zero, we see that $\mathcal{P}^2(\Sigma \cap (M \times [t^1, t^2]))$ is finite. One can prove that Σ_t is a finite set in a similar way. If it is an infinite set, then we can find l distinct points $x_1, \ldots, x_l$ in Σ_t and $R > 0$ such that balls $B_R(x_i)$ are disjoint. For the same $\delta > 0$ as before, we may repeat the above argument for (x_i, t) to see

$$l(5R^2) \le C \sum_{i=1}^{l} \int_{Q_R((x_i,t))} e_\alpha(u_\alpha)\sqrt{|g|} dxdt \le CE_0R^2.$$

This gives an upper bound on the number of points in Σ_t.

(iii) Applying the uniform bound of ∇u_α to (1.1), the same proof of Theorem 1.1 yields that u_α converges smoothly to u outside Σ. Then u satisfies the harmonic map flow equation on $M \times [0,\infty) \setminus \Sigma$. The rest of the proof is exactly the same as in Theorem 7.2.3 of [17]. $\square$

We now prove Theorem 1.3.

Proof. (i) Let T be the first concentration time of Σ. By the local energy inequality, there is a unform constant $a > 0$ such that $T \geq a > 0$. In fact, there is a constant $R_0 > 0$ such that $\int_{B_{2R_0}(x)} e_\alpha(u_0)dv \leq \frac{\varepsilon_0}{4}$ for all $x \in M$. By the local energy inequality, we have

$$\int_{B_R(x)} e_\alpha(u_\alpha)(t)dv \leq \int_{B_{2R}(x)} e_\alpha(u_0)dv + CE_0\frac{t}{R_0^2} < \varepsilon_0 \tag{3.6}$$

for $t \in [0,a]$, where the constant a is independent of α. This shows that $T \geq a > 0$. Let T_1 be the first singular time of the solution $\hat{u}(x,t)$ constructed by Struwe. By uniqueness of the smooth solution of the harmonic map flow, $u(x,t) = \hat{u}(x,t)$ for $t < T \leq T_1$. We claim $T = T_1$. Otherwise, assume that $T < T_1$. Let $z = (x_0, T)$ be a most-left point of Σ; i.e. $T = \inf_{(x,t)\in\Sigma} t$. Since $T < T_1$, there exist $r_1 > 0$ and $r_2 > 0$ such that

$$\Sigma \cap (B_{r_1}(x_0) \times (T - r_1^2, T)) = \emptyset \tag{3.7}$$

and

$$\|\nabla \hat{u}\|_{C^0(B_{r_2}(x_0)\times(T-r_2^2,T+r_2^2))} \leq C_1. \tag{3.8}$$

We may assume that $r_1 > r_2$. There exists an $R \in (0, r_2)$ depending only on the constant C in (3.8) and ε_0 such that

$$|B_R(x_0)|(1+4C^2)^2 \leq \varepsilon_0/8. \tag{3.9}$$

Given this R, we then find $\rho > 0$ depending only on R, ε_0 and the overall energy upper-bound E_0 such that

$$\int_{M\setminus B_R(x_0)} (1+|\nabla u_\alpha|^2)^\alpha(s)e^{-\frac{|x-x_0|^2}{4(T-s)}}\varphi^2 dv \leq \varepsilon_0/8 \tag{3.10}$$

for all $s \in [T-4\rho^2, T-\rho^2]$. Since

$$(B_R(x) \times [t-4\rho^2, t-\rho^2]) \cap \Sigma = \emptyset,$$

by (i) of Theorem 1.2, we have for k sufficiently large

$$|\nabla u_{\alpha_k}|(y,s) \le 2C \tag{3.11}$$

for $y \in B_R(x)$ and $s \in [t-4\rho^2, t-\rho^2]$. Now we can estimate for k sufficiently large

$$\begin{aligned}\Psi_\rho^{\alpha_k}(u,(x_0,T)) &= \rho^{2\alpha_k-2}\int_{M\times[T-4\rho^2,T-\rho^2]}(1+|\nabla u_\alpha|^2)^\alpha G_z\varphi^2\,dvdt\\ &\le 3\max_{s\in[T-4\rho^2,T-\rho^2]}\int_M(1+|\nabla u_{\alpha_k}|^2)^{\alpha_k}e^{-\frac{|x|^2}{4(t-s)}}\varphi^2\,dv\\ &\le \frac{7}{8}\varepsilon_0,\end{aligned}$$

where we used (3.9), (3.10) and (3.11). This contradicts the fact that $(x,t) \in \Sigma$, so this shows (i) of Theorem 1.3.

(ii) Let $\{\hat{p}_i\}_{i=1}^{\hat{k}}$ be the set of finite singularities of the solution $\hat{u}$ constructed by Struwe at the first singular time T. By Theorem 1.2, u_α converges smoothly to u away from Σ. Set $\Sigma_T := \Sigma \cap \{t=T\} = \{p_i\}_{i=1}^k$. Then, for each $x_0 \in M \setminus \Sigma_T$, it follows from Theorem 1.2 that u is smooth in $B_r(x_0)\times[T-r^2,T+r^2]$ for some small $r>0$. This implies that $\nabla\hat{u}$ is bounded in $B_r(x_0)\times[T-r^2,T)$, which means that x_0 does not belong to $\{\hat{p}_i\}_{i=1}^{\hat{k}}$. On the other hand, if x_0 is not a blow-up point of $\{\hat{p}_i\}_{i=1}^{\hat{k}}$ for $\hat{u}$ at time T, then there exists $r>0$ such that

$$|\nabla u|(x,t) \le C, \tag{3.12}$$

for (x,t) in $B_r(x_0)\times[T-r^2,T)$. The above proof in the part (i) gives that $x_0 \notin \Sigma_T$. Hence, $\{p_i\}$ and $\{\hat{p}_i\}$ are the same set of finite points; i.e., $k=\hat{k}$. Without loss of generality, assume $k=1$. It remains to show $\hat{m} \le m$. For each $R>0$ and $\delta>0$, the local energy inequality gives

$$\int_{B_R(p)} e_\alpha(u_\alpha)(T)dv \le \int_{B_{2R}} e_\alpha(u_\alpha)(T-\delta)dv + CE_0\frac{\delta}{R^2}. \tag{3.13}$$

For any $\varepsilon>0$, choose R so that $\int_{B_{2R}(p)} e(\hat{u})(T) \le \varepsilon/6$ and $\mathrm{Vol}(B_{2R}(p)) \le \varepsilon/6$. Let $\delta>0$ be a small number such that $CE_0\frac{\delta}{R^2} \le \varepsilon/6$. Taking $\alpha\to 1$ in (3.6), we have

$$m \le \int_{B_{2R}(p)} e(\hat{u})(T-\delta)dv + \varepsilon/3.$$

Finally, letting δ go to zero, we obtain

$$m \leq \int_{B_{2R}(p)} e(\hat{u})(T)dv + \hat{m} + \varepsilon/3 \leq \hat{m} + \varepsilon/2.$$

Theorem 1.3 follows by the arbitrariness of ε. □

4. An application

In this section, we apply our results about the α-flow to the study of minimizing sequence of the Dirichlet energy in a homotopy class. The following lemma is a variant of the main estimate in [19].

Lemma 4.1. *Let w be a smooth map from B_1 to N satisfying the following scaled equation for some $R > 0$:*

$$\triangle w + (\alpha - 1)\frac{\nabla |\nabla w|^2 \cdot \nabla w}{R^2 + |\nabla w|^2} + A(\nabla w, \nabla w) = h. \tag{4.1}$$

Then there exists $\varepsilon_0 > 0$ such that if $E(w) < \varepsilon_0$ and $\alpha - 1$ is sufficiently small we have

$$\|w - \bar{w}\|_{W^{2,2}(B_{1/2})} \leq C(\|\nabla w\|_{L^2(B_1)} + \|h\|_{L^2(B_1)}).$$

Here, $\bar{w}$ is the mean value of w on B_1.

Proof. The proof is similar to the main estimate in [19]. During the proof, we write $\|\cdot\|_{p,q}$ for $\|\cdot\|_{W^{p,q}(B_1)}$. Multiplying the equation by φ and taking the L^p norm, we have

$$\begin{aligned} \|\triangle(\varphi w)\|_{0,p} \leq\ & 2(\alpha - 1)\|\varphi w\|_{2,p} + C\||\nabla(\varphi w)|\,|\nabla w|\|_{0,p} \\ & + C\|w\|_{1,p} + C\|h\|_{0,p}. \end{aligned} \tag{4.2}$$

For $1 < p < 2$, the Hölder inequality implies that

$$\||\nabla(\varphi w)|\,|\nabla w|\|_{0,p} \leq \|\nabla(\varphi w)\|_{0,q}\|\nabla w\|_{0,2},$$

where $q = 2p/(2-p)$. By the L^p-estimate, for $\alpha - 1$ small, we have

$$\|\varphi w\|_{2,p} \leq C\varepsilon_0\|\nabla(\varphi w)\|_{0,q} + C\|w\|_{1,p} + C\|h\|_{0,p}.$$

If we further assume that ε_0 is small, then it follows from the Sobolev embedding theorem that

$$\|\varphi w\|_{2,p} \leq C \|w\|_{1,p} + C \|h\|_{0,p}.$$

Setting $p = 4/3$ and using Sobolev embedding again, we have

$$\|\varphi w\|_{1,4} \leq C(\|w\|_{1,4/3} + \|h\|_{0,4/3}).$$

With this, we can apply the interior L^2-estimate to (4.1) and multiply it by φ to get

$$\|w - \bar{w}\|_{W^{2,2}(B_{1/2})} \leq C(\|\nabla w\|_{L^2(B_1)} + \|h\|_{L^2(B_1)}).$$

□

We now prove Theorem 1.4.

Proof. Since the minimizing sequence u_i is smooth, there is a corresponding sequence $\alpha_i > 1$ with $\alpha_i \to 1$ such that

$$E_{\alpha_i}(u_i) \leq E(u_i) + \mathrm{Vol}(M) + \frac{1}{i},$$

which implies

$$\lim_{i\to\infty} E_{\alpha_i}(u_i) = \lim_{i\to\infty} E(u_i) + \mathrm{Vol}(M). \tag{4.3}$$

Then there is a unique solution $u_{\alpha_i}(x,t)$ to the α_i-flow (1.1) with initial value u_i. By Lemma 2.1, we have

$$E_{\alpha_i}(u_{\alpha_i}(\cdot,1)) + \int_0^1 \int_M (1 + |\nabla u_{\alpha_i}|^2)^{\alpha_i - 1} |\partial_t u_{\alpha_i}|^2 \, dvdt = E_{\alpha_i}(u_i).$$

Since u_i is a minimizing sequence of E, we know

$$\lim_{i\to\infty} E_{\alpha_i}(u_{\alpha_i}(\cdot,1)) \geq \lim_{i\to\infty} E(u_i) + \mathrm{Vol}(M).$$

Combining the above three identities yields

$$\lim_{i\to\infty} \int_0^1 \int_M (1 + |\nabla u_{\alpha_i}|^2)^{\alpha_i - 1} |\partial_t u_{\alpha_i}|^2 \, dvdt = 0. \tag{4.4}$$

Let u be the weak limit of u_{α_i} in $W^{1,2}(M \times [0,1])$ by passing to a subsequence if necessary. By Theorem 1.2, we know that u is a weak solution to

the harmonic map flow. It follows from (4.4) that u is a map independent of $t \in [0,1]$. Since $u_i(x,t)$ converges weakly to $u(x,t)$ in $W^{1,2}(M \times [0,1])$ and the trace operator $\mathbb{T} : W^{1,2}(M \times [0,1]) \to L^2(M \times \{0\})$ is a bounded linear operator, $u_{\alpha_i}(\cdot, 0) = u_i$ converges weakly to $u(\cdot, 0)$ in L^2. Hence, $u(\cdot, t) \equiv u(\cdot, 0)$ is a (weakly) harmonic map from M to N.

Instead of proving the energy identity for u_i directly, we will find another minimizing sequence of maps v_i in the same homotopy class, satisfying some perturbed harmonic map equation (see (4.7) below).

Due to (4.4), passing to a subsequence if necessary, we assume

$$\int_{1/2}^{1} \int_M |\partial_t u_{\alpha_i}|^2 \, dv dt \leq \frac{1}{16^i}. \tag{4.5}$$

Consider

$$I_i = \left\{ t \in [1/2, 1] : \quad \int_M |\partial_t u_{\alpha_i}(\cdot, t)|^2 \, dv \geq \frac{1}{2^i} \right\}.$$

It follows from (4.5) that

$$|I_i| \leq \frac{1}{8^i},$$

where $|I_i|$ is the Lebesgue measure of I_i. Since

$$\sum_{i=1}^{\infty} |I_i| \leq \sum_{i=1}^{\infty} \frac{1}{8^i} \leq \frac{1}{4},$$

there exists at least one $t_0 \in [1/2, 1]$ such that for all i

$$\int_M |\partial_t u_{\alpha_i}(\cdot, t_0)|^2 \, dv \leq \frac{1}{2^i}. \tag{4.6}$$

For simplicity, denote $u_{\alpha_i}(x, t_0)$ by v_i and write h_i for $\partial_t u_{\alpha_i}(x, t_0)$. Hence, v_i satisfies the equation

$$\triangle v_i + 2(\alpha_i - 1)(\nabla^2 v_i, \nabla v_i) \nabla v_i (1 + |\nabla v_i|^2)^{-1} + A(\nabla v_i, \nabla v_i) = h_i. \tag{4.7}$$

As shown above, v_i also converges weakly to u in $W^{1,2}$. Moreover, by (4.3) and Lemma 2.1, v_i is also a minimizing sequence of E in the same homotypy class. Using Lemma 4.1, there exist finitely many points $x_1, \ldots, x_k$ such that v_i converges to u smoothly in M away from these points. By the well-known removable singularity theorem on harmonic maps, u can be extended to a smooth map on M. The proof of the identity (1.2) will be given in Section 5.

Now, let us prove the second part of Theorem 1.3 assuming that $\pi_2(N)$ is trivial. For simplicity, we assume $k = 1$. Let $\eta(r)$ be a smooth cutoff function in $\mathbb{R}$ with the property that $\eta \equiv 1$ for $r \geq 1$ and $\eta \equiv 0$ for $r \leq 1/2$. For some $\rho > 0$, we define a new sequence of maps $\tilde{v}_i : M \to N$ such that $\tilde{v}_i$ is the same as v_i outside $B_\rho(x_1)$, and for $x \in B_\rho(x_1)$,

$$\tilde{v}_i(x) = \exp_{u(x)}\left(\eta\left(\frac{|x|}{\rho}\right)\exp^{-1}_{u(x)} \circ v_i(x)\right),$$

where exp is the exponential map on N. We claim that

$$\|\tilde{v}_i - u\|_{W^{1,2}(M)} \to 0 \tag{4.8}$$

as $i \to \infty$. To see this, it suffices to consider $B_\rho(x_1) \setminus B_{\rho/2}(x_1)$ because $\tilde{v}_i \equiv u$ on $B_{\rho/2}(x_1)$ and $\tilde{v}_i \equiv v_i$ outside $B_\rho(x_1)$. On the other hand, Lemma 4.1 implies that v_i converges to u on $B_\rho(x_1) \setminus B_{\rho/2}(x_1)$ strongly in $W^{1,2}$ and C^β for some $\beta > 0$. Hence for large i, $v_i(B_{\rho(x_1)} \setminus B_{\rho/2}(x_1))$ lies in a small neighbourhood of $u(x_1)$, where $\exp^{-1}_{u(x)}$ is a well defined smooth map (if ρ is small). Since $F(y) = \exp_{u(x)}\left(\eta(\frac{|x|}{\rho})\exp^{-1}_{u(x)} y\right)$ is a smooth map from a neighbourhood of $u(x_1)$ into itself, we have

$$\begin{aligned}\|\tilde{v}_i - u\|_{W^{1,2}(B_\rho \setminus B_{\rho/2}(x_1))} &= \|F \circ v_i - F \circ u\|_{W^{1,2}(B_\rho \setminus B_{\rho/2}(x_1))} \\ &\leq C\,\|v_i - u\|_{W^{1,2}(B_\rho \setminus B_{\rho/2}(x_1))} \to 0 \quad \text{as } i \to \infty.\end{aligned}$$

Thus the claim follows. Since $\pi_2(N)$ is trivial, $\tilde{v}_i$ is in the same homotopy class as v_i. Since v_i is a minimizing sequence of the Dirichlet energy and v_i converges weakly to u in $W^{1,2}$, we have

$$E(u) \leq \lim_{i\to\infty} E(v_i) \leq \lim_{i\to\infty} E(\tilde{v}_i), \tag{4.9}$$

which implies

$$E(u) = \lim_{i\to\infty} E(v_i).$$

Now, v_i converges to u strongly in $W^{1,2}(M, N)$, which means that there is no energy concentration and Lemma 4.1 in turn shows that the convergence is in C^β for some $\beta > 0$. □

5. An energy identity

In the previous section, by using the $\alpha-$flow, we have constructed a sequence of v_i satisfying

(1) an $\varepsilon-$regularity lemma such as Lemma 4.1 holds for each v_i;

(2) v_i minimizes E within some fixed homotopy class.

The purpose of this section is to prove that for such a sequence v_i, there exist l harmonic maps $\omega_k (k = 1, \ldots, l)$ from $\mathbb{R}^2$ to N such that

$$\lim_{i\to\infty} E(v_i) = E(v) + \sum_{k=1}^{l} E(\omega_k),$$

where v is the weak limit of v_i in $W^{1,2}$.

5.1. Review of bubble tree construction

Let us review the bubble tree construction of Parker [18]. We refer the reader to Section 1 of [18] for details.

Let v_i be the above minimizing sequence satisfying the equation

$$\triangle v_i + \frac{2(\alpha_i - 1)(\nabla^2 v_i, \nabla v_i)\nabla v_i}{1 + |\nabla v_i|^2} + A(\nabla v_i, \nabla v_i) = h_i.$$

By the $\varepsilon-$regularity theorem, we know that v_i converges to v on any compact subset of $M \setminus \{x_k\}$ for finitely many bubble points x_k. A certain amount of the energy concentrates on each bubble point x_k so that

$$\lim_{i\to\infty} |\nabla v_i|^2 \, dv = |\nabla v|^2 \, dv + \sum_k m_k \delta_{x_k}.$$

At each bubble point x_k, one may find $\varepsilon_{i,k} > 0$ such that $\varepsilon_{i,k} \leq \frac{1}{i}$ and

$$\int_{B_{2\varepsilon_{i,k}}(x_k)} |\nabla v|^2 \, dv \leq \frac{m_k}{16 i^2}.$$

A crucial step in the Park construction [18] is using a 'renormalization constant' $C_R > 0$, which can be chosen as small as we like. For our purposes, set $C_R = \varepsilon_1/6$, where ε_1 is the constant in Lemma 5.4.

By the exponential map at x_k, we may identify a point in a neighbourhood of x_k with its geodesic coordinates. For each k and i, the centre of

mass $c_{i,k} := (c^1_{i,k}, c^2_{i,k}) \in M$ is defined by

$$c^j_{i,k} = \frac{\int_{B_{2\varepsilon_{i,k}}} x^j |\nabla v_i|^2 \, dv}{\int_{B_{2\varepsilon_{i,k}}} |\nabla v_i|^2 \, dv}.$$

$\lambda_{i,k} > 0$ is defined by

$$\lambda_{i,k} = \left\{ \inf \lambda : \quad \int_{B_{\varepsilon_{i,k}}(c_{i,k}) \setminus B_\lambda(c_{i,k})} e(v_i) \le C_R \right\}.$$

This implies that

$$\int_{B_{\varepsilon_{i,k}}(c_{i,k}) \setminus B_{\lambda_{i,k}}(c_{i,k})} e(v_i) = C_R.$$

Set $A_{i,k} := B_{\varepsilon_{i,k}}(c_{i,k}) \setminus B_{i\lambda_i,k}(c_{i,k}) \subset B_{\varepsilon_{i,k}}(c_{i,k}) \setminus B_{\lambda_{i,k}}(c_{i,k})$. This implies that

$$\int_{A_{i,k}} e(v_i) \le C_R < \varepsilon_1/6. \tag{5.1}$$

Setting $\tilde{v}_i(\tilde{x}) = v_i(\lambda_{k,i}\tilde{x} + c_{i,k})$ on $B_i(0)$ and passing to the limit $i \to \infty$, we have a weak limit ω_k (a level 1 bubble) of $\tilde{v}_i$ and finitely many bubble points on $\mathbb{R}^2$ so that the blow-up happens again at a higher level. The process stops after finite steps. Suppose there are L such necks in total.

In conclusion, we have

$$\lim_{i\to\infty} E(v_i) = E(v) + \sum_{k=1}^{L} E(\omega_k) + \limsup_{i\to\infty} \sum_{k=1}^{L} \int_{A_{i,k}} e(v_i). \tag{5.2}$$

For simplicity, we label bubbles and necks by only one index. So $A_{i,k}$ could be a neck region at any level. An example of the bubble tree and the decomposition are illustrated in figure 1. The shadow parts in figure 1(b) stand for the neck regions.

5.2. An extension lemma

Let p be a point in N and r_{inj} the injectivity radius of N. Consider a map $f : S^1 \to N$ whose image lies in $B_{r_{inj}}(p)$. Define $\tilde{f} : B_2 \setminus B_1 \to N$ by

$$\tilde{f}(\theta, r) = \exp_p[(2 - r) \exp_p^{-1} f(\theta)] \tag{5.3}$$

(Here (θ, r) denotes the geodesic polar coordinates.) It is known from the definition that $\tilde{f}(\theta, 1) = f(\theta)$ and $\tilde{f}(\theta, 2) = p$. We need to compute the energy

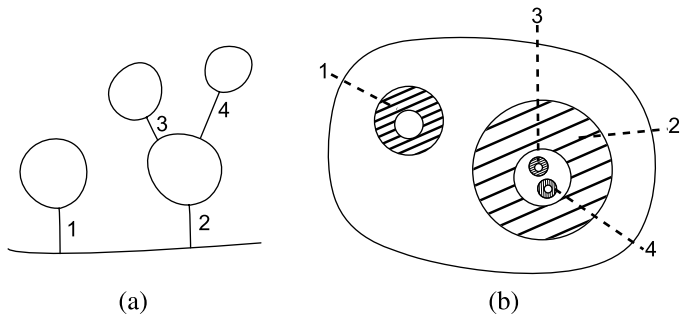

Figure 1: An example

of $\tilde{f}$. The chain rule implies

$$\left|\frac{\partial}{\partial r}\tilde{f}(\theta,r)\right| = \left|D(\exp_p)_{(2-r)\exp_p^{-1} f(\theta)} \cdot \exp_p^{-1} f(\theta)\right| \leq Cd(f(\theta),p)$$

and

$$\begin{aligned}\left|\frac{\partial}{\partial \theta}\tilde{f}(\theta,r)\right| &= \left|D(\exp_p)_{(2-r)\exp_p^{-1} f(\theta)} \cdot (2-r)D(\exp_p^{-1})_{f(\theta)} f'(\theta)\right| \\ &\leq C(2-r)|f'(\theta)|.\end{aligned}$$

Hence,

$$E(\tilde{f}) \leq C\int_{S^1}\int_1^2 \left(d(f(\theta),p)^2 + \frac{1}{r^2}(2-r)^2\left|f'(\theta)\right|^2\right) rdrd\theta.$$

We summarize the computation in a lemma.

Lemma 5.1. *For any $\varepsilon > 0$, there exists $\eta > 0$ depending only on ε and N such that if $f : S^1 \to N$ satisfies*

$$\max_\theta [d(f(\theta),p) + \left|f'(\theta)\right|] < \eta,$$

then the Dirichlet energy of $\tilde{f}$ defined in (5.3) on $B_2 \setminus B_1$ is smaller than ε.

Remark 5.2. We also need a similar result for

$$\tilde{f}(\theta,r) = \exp_p[(r-1)\exp_p^{-1} f(\theta)],$$

which is an extension of a map defined on ∂B_2 to $B_2 \setminus B_1$ such that it is a constant map on ∂B_1. The proof is the same as that of Lemma 5.1, so we omit it.

5.3. A reference map

For the minimizing sequence $\{v_i\}$, we will construct a new minimizing sequence from M to N whose image looks like the weak limit map v and bubbles ω_k connected by geodesics. It keeps a record of some topological information which will later be used to construct a new minimizing sequence.

Lemma 5.3. *There exists a positive constant σ depending only on N such that if f_1, f_2 are two continuous maps from Ω into N satisfying*

$$d(f_1(x), f_2(x)) \leq \sigma$$

for all $x \in \Omega$ and $f_1|_{\partial\Omega} = f_2|_{\partial\Omega}$, then there exists a homotopy deforming f_1 to f_2 with $\partial\Omega$ fixed.

We omit the proof since it is obvious.

Lemma 5.4. *There is a positive constant ε_1 depending only on N such that any $W^{1,2} \cap C^0$ map $g : S^2 \to N$ with $E(g) < \varepsilon_1$ is homotopic to a constant mapping.*

This result follows immediately from [24] and is well known.

The next lemma shows how to squeeze a little room on both ends of a long neck region without changing the energy much.

Lemma 5.5. *For i large so that $\delta >> \lambda_i R$, there is a diffeomorphism f from $B_\delta \setminus B_{\lambda_i R}$ to $B_{\delta/4} \setminus B_{4\lambda_i R}$ such that for any map $u : B_\delta \setminus B_{\lambda_i R} \to N$, we have*

$$1 - C\left(\frac{\delta}{\lambda_i R}\right) \leq \frac{E(u \circ f^{-1}, B_{\delta/4} \setminus B_{4\lambda_i R})}{E(u, E_\delta \setminus E_{\lambda_i R})} \leq 1 + C\left(\frac{\delta}{\lambda_i R}\right),$$

where $C(\frac{\delta}{\lambda_i R})$ is a constant which goes to zero if $\frac{\delta}{\lambda_i R}$ goes to infinity.

Proof. By the conformal invariance of the Dirichlet energy, u can be regarded as a map from $S^1 \times [0, K]$ to N, where $K = \delta/(\lambda_i R)$. Define $\tilde{f} : S^1 \times [0, K] \to$

$S^1 \times [4, K-4]$ by

$$\tilde{f}(\theta, \rho) = \left(\theta, 4 + \rho \frac{K-8}{K}\right).$$

It is straightforward to check

$$1 - C(K) \leq \frac{E(u \circ \tilde{f}^{-1}, S^1 \times [4, K-4])}{E(u, S^1 \times [0, K])} \leq 1 + C(K),$$

for some $C(K)$ such that $C(K)$ goes to zero if K goes to infinity. Set $s : S^1 \times [0, K] \to B_\delta \setminus B_{\lambda_i R}$ to be

$$s(\theta, \rho) = (\theta, \lambda_i R e^\rho).$$

Then $f : B_\delta \setminus B_{\lambda_i R} \to B_{\delta/4} \setminus B_{4\lambda_i R}$ defined by

$$f = s \circ \tilde{f} \circ s^{-1}$$

is the diffeomorphism we need. □

A remark about the notation is needed for the rest of the proof. For simplicity, we omit the subscript for necks as if there was only one neck. In fact, one should repeat the construction or proof for each neck region. By the nature of the following proof, this should cause no further difficulty.

In Section 5.1, a neck region refers to $A_{i,k} = B_{\varepsilon_i,k}(c_{i,k}) \setminus B_{i\lambda_{i,k}}(c_{i,k})$. For simplicity, we use a larger domain $B_\delta \setminus B_{\lambda_i R}$, omitting the centre, which is also called a neck region. By carefully choosing δ and R as below, we see that the energy on a neck region is small (see (5.4) below), which is crucial to our proof. This is in fact the main reason that we use Parker's construction instead of that of Ding and Tian.

We need δ, R and i to satisfy the following conditions. δ and R will then be fixed for the rest of this section. We assume the neck connects the weak limit and a level 1 bubble. For a general neck, one should replace v with some bubble map ω_j.

Recall that v_i and v satisfy the identity (5.2) and let ω be one of bubbles ω_k.

(A0) Owing to (a) and (b) in Section 5.1, choosing δ small and R large, we may require

$$\limsup_{i \to \infty} \int_{B_\delta \setminus B_{\lambda_{i,k} R}} e(v_i) < \varepsilon_1/3 \tag{5.4}$$

(A1) Set $\varepsilon = \min\{\varepsilon_1, \xi/L\}/8$, where ξ is a positive number to be chosen in the next subsection and L is the total number of necks. Lemma 5.1 gives an η. Choose δ and R so that for each neck region $B_\delta \setminus B_{\lambda_i R}$, v restricted to ∂B_δ satisfies

$$\max_\theta |d(v,p)| + \left|\frac{d}{d\theta} v\right| \le \min\{\eta/2, \sigma/2\}, \tag{5.5}$$

where p is the value of v at the bubble point and ω restricted to $\partial B_R \subset \mathbb{R}^2$ satisfies

$$\max_\theta |d(\omega,q)| + \left|\frac{d}{d\theta} \omega\right| \le \min\{\eta/2, \sigma/2\}, \tag{5.6}$$

where $q := \lim_{y\to\infty} \omega(y)$. Fix δ and R and choose some i sufficiently large so that

(A2) (5.5) and (5.6) remain true if we replace v by v_i, ω by $v_i(\lambda_i \cdot)$, $\eta/2$ by η and $\sigma/2$ by σ;

(A3) $|v_i - v| \le \sigma/3$ for $x \in M \setminus B_\delta$ and $|v_i(\lambda_i y) - \omega(y)| \le \sigma/3$ for all $y \in B_R$ where σ is the constant in Lemma 5.3;

(A4) Let $C(\frac{\delta}{\lambda_i R})$ be the function given in Lemma 5.5. For sufficiently large i, it satisfies $1 + C(\frac{\delta}{\lambda_i R}) < 2$.

Now we define a new map $w : M \to N$ so that

(1) $w = v_i$ on $M \setminus B_\delta$ and $B_{\lambda_i R}$;

(2) $w = v_i \circ f^{-1}$ on $B_{\delta/4} \setminus B_{4\lambda_i R}$, where f is the diffeomorphism constructed in Lemma 5.5;

(3) $w(\theta, r) = \exp_p[\frac{r-\delta/2}{\delta/2} \exp_p^{-1} v_i(\theta, \delta)]$ for $(\theta, r) \in B_\delta \setminus B_{\delta/2}$;

(4) $w(\theta, r) = w(\theta, \frac{\delta^2}{4r})$ for $(\theta, r) \in B_{\delta/2} \setminus B_{\delta/4}$;

(5) $w(\theta, r) = \exp_q[\frac{2\lambda_i R - r}{\lambda_i R} \exp_q^{-1} v_i(\theta, \lambda_i R)]$ for $(\theta, r) \in B_{2\lambda_i R} \setminus B_{\lambda_i R}$;

(6) $w(\theta, r) = w(\theta, \frac{(2\lambda_i R)^2}{r})$ for $(\theta, r) \in B_{4\lambda_i R} \setminus B_{2\lambda_i R}$.

It is easy to see that w is homotopic to v_i. Now let us consider the map w restricted to $B_{\delta/2} \setminus B_{2\lambda_i R}$. It follows from the construction that w maps $\partial B_{\delta/2}$ to p and $\partial B_{2\lambda_i R}$ to q. Hence, topologically it induces a map from S^2 to N if we identify $\partial B_{2\lambda_i R}$ with the north pole P_N and $\partial B_{\delta/2}$ with the south pole P_S. This is in fact a homotopically trivial map. To see this, consider a map $w_1 : \mathbb{R}^2 \to N$ which agrees with w on $B_{\delta/2} \setminus 2\lambda_i R$ and maps $B_{2\lambda_i R}$ to q and $\mathbb{R}^2 \setminus B_{\delta/2}$ to p. By the conformal invariance of E, we regard w_1 as a

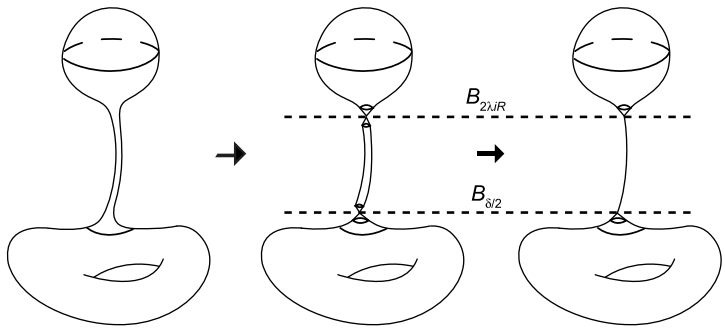

Figure 2: Reference map

map from S^2 to N. By (A1), (A2), (A4) and (5.4) , the energy of this map is smaller than ε_1 in Lemma 5.4, hence it is homotopic to a constant map. Let $F : S^2 \times [0,1] \to N$ be the homotopy such that $F(\cdot, 0) = w_1(\cdot)$, $F(\cdot, 1) \equiv p$ and $F(P_S, t) = p$ for any $t \in [0,1]$. The curve $F(P_N, \cdot)$ connects q and p. Let $\gamma : [0,1] \to N$ be the shortest geodesic homotopic to this curve connecting q and p.

We can now define the reference map $\tilde{w}$. Let $\tilde{w} = w$ outside of $B_{\delta/2} \setminus B_{2\lambda_i R}$, with

$$\tilde{w}(\theta, r) = \gamma\left(\frac{\log r - \log(2\lambda_i R)}{\log(\delta/2) - \log(2\lambda_i R)}\right) \quad \text{on } B_{\delta/2} \setminus B_{2\lambda_i R}.$$

The above construction is illustrated by Figure 2. We claim that w and $\tilde{w}$ are homotopic. Since w and $\tilde{w}$ are the same except on $B_{2\delta} \setminus B_{2\lambda_i R}$, we construct the homotopy explicitly. For $(\theta, r) \in B_{\delta/2} \setminus B_{2\lambda_i R}$,

$$\begin{aligned} &G((\theta, r), s) \\ &\quad = \begin{cases} F\left(P_N, \dfrac{r - 2\lambda_i R}{\delta/2 - 2\lambda_i R}\right), & 2\lambda_i R \le r \le r(s) \\ F\left(\left(\theta, (\delta/2 - 2\lambda_i R)\dfrac{r - r(s)}{\delta/2 - r(s)} + 2\lambda_i R\right), s\right), & r(s) \le r \le \delta/2, \end{cases} \end{aligned}$$

where $r(s) = 2\lambda_i R + s(\delta/2 - 2\lambda_i R)$. From here, it is obvious that one can further deform $F(P_N, t)$ to γ.

If there is more than one neck region, we repeat the above construction for each neck.

5.4. Proof of the energy identity

We prove the energy identity (1.2) by contradiction. Assume that (1.2) is not true. By (5.2), there exists $\xi > 0$ such that

$$\limsup_{i\to\infty} \sum_{k=1}^{L} \int_{A_{i,k}} e(v_i) \geq \xi.$$

Passing to a subsequence, we have

$$\lim_{i\to\infty} \sum_{k=1}^{L} \int_{B_\delta \setminus B_{R\lambda_i}} e(v_i) \geq \xi.$$

Let $\tilde{w}_i$ be the corresponding map of the map v_i constructed in last subsection. We can construct a new sequence $\tilde{v}_i$ in the same homotopy class such that

$$\lim_{i\to\infty} E(\tilde{v}_i) < \lim_{i\to\infty} E(v_i).$$

More precisely, we define $\tilde{v}_i$ in the follows:

(1) $\tilde{v}_i = v_i$ outside the neck region $(B_\delta \setminus B_{\lambda_i R})$;

(2) $\tilde{v}_i(\theta, r) = \exp_p[\frac{r-\delta/2}{\delta/2} \exp_p^{-1} v_i(\theta, \delta)]$ for $(\theta, r) \in B_\delta \setminus B_{\delta/2}$;

(3) $\tilde{v}_i(\theta, r) = \exp_q[\frac{2\lambda_i R - r}{\lambda_i R} \exp_p^{-1} v_i(\theta, \lambda_i R)]$ for $(\theta, r) \in B_{2\lambda_i R} \setminus B_{\lambda_i R}$;

(4) $\tilde{v}_i(\theta, r) = \gamma(\frac{\log r - \log(2\lambda_i R)}{\log(\delta/2) - \log(2\lambda_i R)})$ for $(\theta, r) \in B_{\delta/2} \setminus B_{2\lambda_i R}$.

It remains to show that firstly the energy of $\tilde{v}_i$ in the neck region is smaller than $\xi/2$ for i large, and secondly $\tilde{v}_i$ is homotopic to $\tilde{w}$. For the first part, (A1) and Lemma 5.1 implies that the energy in $B_\delta \setminus B_{\delta/2}$ and $B_{2\lambda_i R} \setminus B_{\lambda_i R}$ for all necks is no more than $\xi/4$. Moreover, the energy on the geodesic part is

$$E(\tilde{v}_i; B_{\delta/2} \setminus B_{2\lambda_i R}) = \frac{\mathrm{Length}(\gamma)^2}{\log(\delta/2) - \log(2\lambda_i R)},$$

which can be made as small as we like when i is sufficiently large due to the fact that the length of γ is fixed (see Section 5.3).

For the second point, if there is only one neck, then M is decomposed into three parts; $M \setminus B_{\delta/2}$, $B_{\delta/2} \setminus B_{2\lambda_i R}$ and $B_{2\lambda_i R}$. For the reference map $\tilde{w}$, the decomposition is fixed because we chose and fixed a large i in the construction. For $\tilde{v}_i$, the centre of B_δ and $B_{\lambda_i R}$ changes with i and so does

the radius of $B_{2\lambda_i R}$. However, when restricted to the separating circles $\partial B_{\delta/2}$ and $\partial B_{2\lambda_i R}$, $\tilde{w}$ and $\tilde{v}_i$ are constant maps to the same points. This allows us to prove the homotopy between $\tilde{w}$ and $\tilde{v}_i$ on each of the three parts separately.

On $M \setminus B_{\delta/2}$, (A1) and (A3) show that $d(\tilde{w}, \tilde{v}_i) < \sigma$, which implies that there is a homotopy between them leaving $\partial B_{\delta/2}$ fixed. The bubble part $B_{2\lambda_i R}$ of $\tilde{w}_i$ and $\tilde{v}_i$ are homotopic for the same reason. On the $B_{\delta/2} \setminus B_{2\lambda_i R}$ part, they are different parametrizations of the same curve. This concludes our proof for (1.2) in the case of only one neck.

For the general case of multiple necks, the decomposition is more complicated but the same argument applies. To illustrate the idea, consider a bubble tree as shown in Figure 1(a). The necks are labelled from 1 to 4. The corresponding decomposition of M is illustrated in Figure 1(b). The shadow parts are neck regions. There are eight circles separating the neck regions and the bubble regions. Both $\tilde{w}_i$ and $\tilde{v}_i$ map these circles to p_i's and q_i's. The homotopy for bubble regions (including $M \setminus (B_{\delta/2}(x_1) \cup B_{\delta/2}(x_2))$ follows from Lemma 5.3 and the homotopy for neck regions is obvious because they are different parametrization of the same geodesics.

A. Appendix

In this appendix, we prove the local existence and uniqueness of the α-flow equation

$$\partial_t u = \triangle_M u + (\alpha - 1)\frac{\nabla |\nabla u|^2 \cdot \nabla u}{1 + |\nabla u|^2} + A(u)(\nabla u, \nabla u), \tag{A.1}$$

for any smooth initial value $u_0 : M \to N \subset \mathbb{R}^k$.

As in the case of harmonic maps, there is an intrinsic way of writing the right-hand side of (A.1). If we regard $(1 + |\nabla u|^2)^{\alpha-1}\nabla u$ as a section of the pull-back bundle u^*TN and denote the induced connection of u^*TN by $\tilde{D}$, the right-hand side can be written as

$$\tau_\alpha(u) := \frac{1}{(1 + |\nabla u|^2)^{\alpha-1}} \sum_i \tilde{D}_{e_i} \left((1 + |\nabla u|^2)^{\alpha-1} \nabla_{e_i} u \right),$$

where $\{e_i\}$ is an orthonormal basis of TM in some open set. In the following, we will call τ_α the $\alpha-$tension field of u. It is obvious that smooth solution to (A.1) is equivalent to that of

$$\frac{\partial u}{\partial t} = \tau_\alpha(u),$$

where we regard $u(\cdot,t)$ as maps into N. This allows us to consider a totally geodesic isometric embedding of N into $\mathbb{R}^k$ with some non-flat metric as in [10]. Precisely, there is a metric h on $\mathbb{R}^k$ and an embedding N into $\mathbb{R}^k$, such that the isometric embedding is totally geodesic, and there exists an isometric involution i which acts on a tubular neighbourhood $\mathcal{T}$ of N, leaving points of N fixed. By composing with this embedding, we regard u as a map from M into $(\mathbb{R}^k,h)$, and denote by τ_α^N and $\tau_\alpha^{\mathbb{R}^k}$ the $\alpha-$tension fields of u as a map into N and into $\mathbb{R}^k$ respectively. Since N is a totally geodesic submanifold, we have $\tau_\alpha^N(u)=\tau_\alpha^{\mathbb{R}^k}(u)$ (see page 108 in [10]). Therefore, it suffices to study

$$\frac{\partial u^\beta}{\partial t}=\triangle_M u^\beta+(\alpha-1)\frac{\nabla\left|\nabla u\right|^2\cdot\nabla u^\beta}{1+\left|\nabla u\right|^2}+\Gamma(u)\#\nabla u\#\nabla u, \tag{A.2}$$

where u^β are the components of u, $\triangle_M$ is the usual Laplacian on M and Γ is the Christoffel symbol of the Levi–Civita connection of $(\mathbb{R}^k,h)$. Since $\left|\nabla u\right|^2=h_{\beta\gamma}(u)\nabla u^\beta\cdot\nabla u^\gamma$, when we expand $\nabla\left|\nabla u\right|^2$, there is an extra term involving the gradient of $h(u)$ which can be absorbed into the last term. The equation can be rewritten as

$$\frac{\partial u^\beta}{\partial t}=\triangle_M u^\beta+2(\alpha-1)\frac{\nabla^2_{ij}u^\gamma\nabla_i u^\gamma\nabla_j u^\beta}{1+\left|\nabla u\right|^2}+\tilde{\Gamma}(u)\#\nabla u\#\nabla u. \tag{A.3}$$

Next, we start a routine iteration procedure to show the local existence of (A.3). Consider the following linear parabolic system:

$$\frac{\partial u}{\partial t}=\triangle u+(\alpha-1)\frac{(\nabla^2u,\nabla v)\nabla v}{1+\left|\nabla v\right|^2}+\tilde{\Gamma}(v)\#\nabla v\#\nabla v. \tag{A.4}$$

Let $w=u-u_0$. Then w satisfies

$$\begin{aligned}\frac{\partial w}{\partial t}&=\triangle w+(\alpha-1)\frac{(\nabla^2w,\nabla v)\nabla v}{1+\left|\nabla v\right|^2}+\tilde{\Gamma}(v)\#\nabla v\#\nabla v\\&\quad+\triangle u_0+(\alpha-1)\frac{(\nabla^2u_0,\nabla v)\nabla v}{1+\left|\nabla v\right|^2}.\end{aligned} \tag{A.5}$$

For some $\mu\in(1,2)$, set

$$V=\left\{v\in C^{\mu,\mu/2}(M\times[0,T])|\quad \|v\|_{C^{\mu,\mu/2}(M\times[0,T])}\le 1, v(0)\equiv 0\right\}.$$

For each $v \in V$, the Schauder estimate of linear parabolic systems (see [8] and [20]) implies

$$\|w\|_{C^{\mu+1,\frac{\mu+1}{2}}(M\times[0,T])} \leq C(\|w\|_{C^0(M\times[0,T])} + \|u_0\|_{C^{\mu+1}(M)} + 1). \tag{A.6}$$

Since $w(\cdot, 0) \equiv 0$, we have

$$\|w\|_{C^0(M\times[0,T])} \leq TC(\|w\|_{C^0(M\times[0,T])} + \|u_0\|_{C^{\mu+1}(M)} + 1).$$

Choose T small so that

$$\|w\|_{C^0(M\times[0,T])} \leq TC(\|u_0\|_{C^{\mu+1}(M)} + 1).$$

By the interpolation of the Hölder space, for $\sigma > 0$ such that $\mu = (1-\sigma)(\mu+1)$, we have

$$\|w\|_{C^{\mu,\mu/2}(M\times[0,T])} \leq \|w\|^{\sigma}_{C^0(M\times[0,T])} \|w\|^{1-\sigma}_{C^{\mu+1,\frac{\mu+1}{2}}(M\times[0,T])}.$$

This implies that we can choose T sufficiently small so that $w \in V$.

Now start from any $v_0 \in V$. Let v_{k+1} be the solution of (A.4) with $v = v_k$ and zero initial value. We then have a uniform bound

$$\|v_k\|_{C^{\mu+1,\frac{\mu+1}{2}}(M\times[0,T])} \leq C.$$

Using Schauder estimate again, we obtain a uniform $C^{\mu+3,\frac{\mu+3}{2}}$ estimate for v_k. By taking a subsequence, we know v_k converges to some w in $C^{\mu+1,\frac{\mu+1}{2}}$ $(M \times [0,T])$. Then $w + u_0$ is a solution to the α-flow with initial value u_0.

If there are two smooth solutions u and w to Equation (A.3) defined on $[0,T]$. We can subtract the two equations, multiply both sides by $u - w$ and integrate over M to get

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\int_M |w-u|^2\,dv \leq &-\int_M |\nabla(u-w)|^2\,dv \\ &+ 2(\alpha-1)\int_M (\nabla^2_{ij}(u^\beta - w^\beta))\frac{u^\beta_i u^\gamma_j}{1+|\nabla u|^2}(u^\gamma - w^\gamma)dv \\ &+ C\int_M |u-w|^2\,dv + C\int_M |\nabla u - \nabla w|\,|u-w|\,dv.\end{aligned}$$

For the second term on the right-hand side, integrating by parts and noticing that

$$\frac{u^\beta_i u^\gamma_j}{1+|\nabla u|^2}\nabla_i(u^\beta - w^\beta)\nabla_j(u^\gamma - w^\gamma) \geq 0$$

yields

$$\frac{d}{dt}\int_M |u-w|^2\,dv \le C\int_M |u-w|^2\,dv,$$

which implies $u \equiv w$ on $M \times [0,T]$ if $u = w$ for $t = 0$. It follows immediately from the uniqueness and the involution isometry $i : \mathcal{T} \to \mathcal{T}$ that if u_0 is a map from M to N, the image of the local solution $u(x,t)$ obtained above lies in N.

Acknowledgments

After finishing an early version in 2010, the authors have received useful comments and suggestions from Professors R. Schoen, R. Hamilton and M. Struwe and would like to thank them for their interests and encouragements. The research of the first author was supported by the Australian Research Council grant no. DP0985624. The second author was supported by a postdoctoral fellowship at the University of Queensland through the grant no. DP0985624. The second author was also supported by NSFC No. 11101272.

References

[1] M. Bertsch, R. Dal Passo and R. Van der Hout, *Nonuniqueness for the heat flow of harmonic maps on the disk*, Arch. Ration. Mech. Anal., **161**(2) (2002), 93–112.

[2] K.C. Chang, W.Y. Ding and R. Ye, *Finite-time blow-up of the heat flow of harmonic maps from surfaces*, J. Differ. Geom. **36**(2) (1992), 507–515.

[3] J.Y. Chen and G. Tian, *Compactification of moduli space of harmonic mappings*, Comment. Math. Helv. **74** (1999), 201–237.

[4] Y.M. Chen and M. Struwe, *Existence and partial regularity for heat flow for hamonic maps*, Math. Z. **201** (1989), 83–103.

[5] W.Y. Ding and G. Tian, *Energy identity for a class of approximate harmonic maps from surfaces*, Commun. Anal. Geom. **3** (1996), 543–554.

[6] F. Duzaar and E. Kuwert, *Minimization of conformal invariant energies in homotopy classes*, Calc. Var. Partial Differ. Eq. **6**(4) (1998), 285–313.

[7] J. Eells and J.H. Sampson, *harmonic mappings of Riemannian manifolds,* Amer. J. Math. **86** (1964) 109–160.

[8] S.D. Eidel'man, *Parabolic systems*, Translated from the Russian by Scripta Technica, London North-Holland Publishing Co., Amsterdam-London; Wolters–Noordhoff Publishing, Groningen 1969.

[9] M. Giaquinta, M.-C. Hong and H. Yin, *A New Approximation of Relaxed Energies for Harmonic Maps and the Faddeev Model,* Calc. Var. Partial Differ. Eq. **41** (2011) 45–69.

[10] R.S. Hamilton, *Harmonic maps of manifolds with boundary*, Lecture Notes in Mathematics, **471**, Springer–Verlag, Berlin–New York, 1975.

[11] T. Lamm, *Energy identity for approximations of harmonic maps from surfaces*, Trans. Amer. Math. Soc. **362** (2010), 4077–4097.

[12] L. Lemaire, *Applications harmoniques des surfaces Riemanniennes*, J. Differ. Geom. **13** (1978), 51–78.

[13] J.Y. Li and G. Tian, *The blow-up locus of heat flows for harmonic maps*, Acta Math. Sin. (Engl. Ser.), **16**(1) (2000), 29–62.

[14] Y.X. Li and Y.D. Wang, *A weak energy identity and the length of necks for a sequence of Sacks-Uhlenbeck α-harmonic maps*, Adv. Math. **225**(3) (2010), 1134–1184, doi:10.1016/j.aim.2010.03.020.

[15] Y.X. Li and Y.D. Wang, *Bubbling location for sequences of approximate $f-$harmonic maps from surfaces*; Internat. J. Math. **21**(4) (2010), 475–495.

[16] F.H. Lin and C.Y. Wang, Harmonic and quasi-harmonic spheres, Comm. Anal. Geom. **7**(2) (1999), 397–429.

[17] F.H. Lin and C.Y. Wang, The analysis of harmonic maps and their heat flows, World Scientific Publishing Co. Pte. Ltd. (2008).

[18] T. Parker, Bubble tree convergence for harmonic maps, J. Differ. Geom. **44** (1996), 595–633.

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

[20] W. Schlag, *Schauder and L^p estimates for parabolic systems via companato spaces*, Commun. Partial Differ. Eq. **21**(7–8) (1996), 1141–1175.

[21] R. Schoen, *Analytic aspects of the harmonic map problem*, Seminar on Nonlinear Partial Differential Equations (Berkeley, California, 1983), 321–358, Math. Sci. Res. Inst. Publ., 2, Springer, New York, 1984.

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

[23] M. Struwe, *On a free boundary problem for minimal surfaces*, Invent. Math. **75** (1984), 547560.

[24] M. Struwe, *On the evolution of harmonic maps of Riemannian surfaces*, Comm. Math. Helv. **60** (1985), 558–581.

[25] M. Struwe, *On the evolution of harmonic maps in higher dimensions*, J. Differ. Geom. **28** (1988), 485–502.

[26] P. Topping, *Reverse bubbling and nonuniqueness in the harmonic map flow*, Int. Math. Res. Not. (10) (2002), 505–520.

[27] K. Uhlenbeck, *Regularity for a class of non-linear elliptic systems*, Acta Math. **138** (1977), 219–240.

[28] K. Uhlenbeck, *Morse theory by perturbation methods with applications to harmonic maps*, Trans. Amer. Math. Soc **267** (1981), 569–583.

Department of Mathematics,
The University of Queensland
Brisbane, QLD 4072,
Australia
E-mail address: hong@maths.uq.edu.au

School of Mathematical Sciences,
University of Science and Technology of China,
Hefei, China
E-mail address: haoyin@ustc.edu.cn

Received July 7, 2011