MINIMIZING ENERGY AMONG HOMOTOPIC MAPS

PENGZI MIAO

Received 6 May 2003

We study an energy minimizing sequence \{u_i\} in a fixed homotopy class of smooth maps from a 3-manifold. After deriving an approximate monotonicity property for \{u_i\} and a continuous version of the Luckhaus lemma (Simon, 1996) on \(S^2\), we show that, passing to a subsequence, \{u_i\} converges strongly in \(W^{1,2}\) topology wherever there is small energy concentration.

2000 Mathematics Subject Classification: 53C43, 58E20.

1. Introduction. Let \(\phi : M \to N\) be a continuous map between two compact Riemannian manifolds. In general, there may not exist a harmonic map homotopic to \(\phi\) (see [2]). Hence, a map \(u\) that minimizes energy among all smooth maps homotopic to \(\phi\) may not exist. However, it is still a basic question to understand the analytical property of a minimizing sequence. If the domain \(M\) is a compact surface, it is known to experts that any minimizing sequence that converges weakly indeed converges strongly in \(W^{1,2}\) topology away from a finite number of points, where energy concentrates and bubble forms (see [3, 7]). If the domain is of higher dimension, B. White showed that the infimum of the energy functional over the homotopy class of \(\phi\) is determined only by the restriction of \(\phi\) to a 2-skeleton of \(M\) (see [8]). It is our goal in this paper to apply White’s result to derive a similar theorem for 3-manifolds.

**Theorem 1.1.** Let \(\phi : M \to N\) be a smooth map between two compact Riemannian manifolds without boundary. Assume that \(M\) has dimension 3. Then there exists a constant \(\epsilon_0 = \epsilon_0(M,N) > 0\) such that, for any sequence of maps \{u_i\} which minimizes energy among all smooth maps homotopic to \(\phi\) and converges weakly in \(W^{1,2}(M,N)\), if

\[
\liminf_{i \to \infty} \frac{1}{\sigma} \int_{B_{\sigma}(x)} |du_i|^2 dV \leq \epsilon_0,
\]

then \{u_i\} converges strongly in \(W^{1,2}(B_{\sigma/4}(x),N)\).

As an application of Theorem 1.1, we prove a partial regularity result for the weak limit of an energy minimizing sequence.

**Corollary 1.2.** Let \(\phi : M \to N\) be a smooth map between two compact Riemannian manifolds without boundary, where \(M\) has dimension 3. Let \{u_i\} be a sequence of maps which minimizes energy among all smooth maps homotopic to \(\phi\). Suppose that \{u_i\} converges weakly to some \(u \in W^{1,2}(M,N)\); then there exists a closed set \(\Sigma \subset M\) with finite
1-dimensional Hausdorff measure such that \( u \) is a smooth harmonic map from \( M \setminus \Sigma \) to \( N \). In particular, \( u \) is a weakly harmonic map from \( M \) to \( N \).

We remark that the dimension restriction of the domain space comes only from the lemma proved in Section 4. If a similar lemma could be established on a general sphere \( S^{n-1} \subset \mathbb{R}^n \), the rest of the argument in this paper would imply that Theorem 1.1 holds for arbitrary dimension \( n \) with the small energy concentration assumption (1.1) replaced by

\[
\liminf_{i \to \infty} \frac{1}{\sigma^{n-2}} \int_{B_{\sigma}(x)} |d\nu_i|^2 dV \leq \epsilon_0.
\]

2. Preliminaries. Let \( (M^3, g) \) and \( (N^m, h) \) be compact Riemannian manifolds of dimensions 3 and \( m \). We assume that \( M \) and \( N \) have no boundary. By the Nash embedding theorem, it is convenient to regard \( N \) as isometrically embedded in some Euclidean space \( \mathbb{R}^K \). We define

\[
W^{1,2}(M, N) = \{ u \in W^{1,2}(M, \mathbb{R}^K) \mid u(x) \in N \text{ a.e. } x \in M \},
\]

where \( W^{1,2}(M, \mathbb{R}^K) \) is the separable Hilbert space of maps \( u : M \to \mathbb{R}^K \) whose component functions are \( W^{1,2} \) Sobolev functions on \( M \). We note that \( W^{1,2}(M, N) \) inherits both strong and weak topologies from \( W^{1,2}(M, \mathbb{R}^K) \). Moreover, it is a strongly closed set with the property that, for any \( C > 0 \),

\[
\{ u \in W^{1,2}(M, N) \mid \|u\|_{W^{1,2}} \leq C \}
\]

is weakly compact in \( W^{1,2}(M, N) \) (see [4]).

For any \( u \in W^{1,2}(M, N) \), the energy of \( u \) is defined by

\[
E(u) = \int_M \text{Tr}_g(u^*h) dV = \int_M |du|^2 dV,
\]

where \( u^*h \) is the pullback of \( h \) by \( u \) and \( dV \) is the volume measure determined by \( g \) on \( M \).

Let \( C^\infty(M, N) \subset W^{1,2}(M, N) \) be the space of smooth maps. For any \( \phi \in C^\infty(M, N) \), we define

\[
\mathcal{F}_\phi = \{ u \in C^\infty(M, N) \mid u \text{ is homotopic to } \phi \},
\]

\[
E_\phi = \inf \{ E(u) \mid u \in \mathcal{F}_\phi \}.
\]

The following result, which is due to White [8], gives a fundamental characterization of \( E_\phi \).

**White’s Theorem.** Let \( \mathcal{F}_\phi^{(2)} = \{ u \in C^\infty(M, N) \mid u \text{ is 2-homotopic to } \phi \} \), where two continuous maps \( v \) and \( w \) are said to be 2-homotopic if their restrictions to the 2-dimensional skeleton of some triangulation of \( M \) are homotopic. Then

\[
\inf \{ E(u) \mid u \in \mathcal{F}_\phi^{(2)} \} = E_\phi.
\]
Let \( \{u_i\} \subset \mathcal{F}_\phi \) be an arbitrary sequence which minimizes the energy functional, that is,

\[
\lim_{i \to \infty} E(u_i) = E_\phi.
\]  

(2.6)

Then the above theorem suggests that \( \{u_i\} \) is also a minimizing sequence in \( \mathcal{F}^{(2)}_\phi \). This fact is very useful since it allows more competitors to be compared with \( u_i \).

By the weak compactness of bounded sets in \( W^{1,2}(M,N) \), we may assume that, passing to a subsequence, \( \{u_i\} \) converges weakly in \( W^{1,2}(M,N) \), strongly in \( L^2(M,N) \), and pointwise almost everywhere to some \( u \in W^{1,2}(M,N) \), which has the property that

\[
E(u) \leq \lim_{i \to \infty} E(u_i) = E_\phi.
\]  

(2.7)

Moreover, by the Riesz representation theorem, we know that there exists a Radon measure \( \mu \) on \( M \) so that

\[
\|du_i\|^2(x) dV \rightharpoonup \mu.
\]  

(2.8)

3. Approximate monotonicity of \( \{|du_i(x)|^2 dV\} \). Given a \( C^1 \) vector field \( X \) on \( M \), we let \( \{F_t\} \) denote the one-parameter group of diffeomorphism on \( M \) generated by \( X \). For any \( v \in W^{1,2}(M,N) \), we define \( E_v(t,X) = E(v \circ F_t) \), where \( v \circ F_t(x) = v(F_t(x)) \). The first variation formula for the energy functional (see [4]) then gives that

\[
\frac{d}{dt} E_v(t,X) = \int_M \left\{ v^* h, -g'(t) + \frac{1}{2} \{ \text{Tr} g(t) g'(t) \} g(t) \right\} g(t) dV(t),
\]  

(3.1)

where \( v^* h \) is the pullback of \( h \) by \( v \), \( g(t) = E^*_t(g) \), and \( dV(t) \) is the volume measure determined by \( g(t) \). In particular, at \( t = 0 \), we have that

\[
\frac{d}{dt} E_v(0,X) = \int_M \left\{ v^* h, \mathcal{L}_X g - \frac{1}{2} \{ \text{Tr}_g (\mathcal{L}_X g) \} g \right\} g) dV.
\]  

(3.2)

The following lemma says that, for large \( n \), \( u_n \) is “almost stationary” with respect to a large class of domain variations.

**Lemma 3.1.** Given \( \Lambda > 0 \), let \( V_\Lambda = \{ \text{\( C^1 \) vector field \( X \) with } \|X\|_{C^1} \leq \Lambda \} \). Then

\[
\sup_{X \in V_\Lambda} \left\{ \frac{d}{dt} E_n(0,X) \right\} \to 0 \quad \text{as } n \to \infty,
\]  

(3.3)

where \( E_n(t,X) = E_{u_n}(t,X) \).

**Proof.** Let \( \sigma_0 \) be a sufficiently small positive constant depending only on \( (M,g) \) such that for any geodesic ball \( B_\sigma(x_0) \subset M \) with \( \sigma \leq \sigma_0 \) and any geodesic normal coordinate chart \( \{x^1, x^2, x^3\} \) in \( B_\sigma(x_0) \), all the eigenvalues of the matrix \( [g_{ij}(x)]_{3 \times 3} \) lie
in $[1/2,2]$ for each $x \in B_\sigma(x_0)$. With such a choice of $\sigma$, we have that

$$
\int_{B_\sigma(x_0)} |dv|^2dV = \sum_{\alpha=1}^K \int_{B_\sigma(x_0)} \frac{\partial v^\alpha}{\partial x^i} \frac{\partial v^\alpha}{\partial x^j} g^{ij}(x) dV \\
= \frac{1}{2} \sum_{\alpha=1}^K \sum_{i=1}^3 \left| \frac{\partial v^\alpha}{\partial x^i} \right|^2(x) dV
$$

(3.4)

for any $v \in W^{1,2}(M,\mathbb{R}^K)$, where $dx$ denotes the Lebesgue measure in $\mathbb{R}^3$.

To prove the lemma, we first consider $V_{\Lambda,\sigma}$ instead of $V_{\Lambda}$, where $\sigma \leq \sigma_0$ and $V_{\Lambda,\sigma} = \{ X \in V_{\Lambda} | \text{support}\{X\} \subset B_\sigma(x_0) \text{ for some } x_0 \in M \}$.

(3.5)

For any $X \in V_{\Lambda,\sigma}$, we write

$$
G(t) = -g'(t) + \frac{1}{2} \{ \text{Tr} g(t) g'(t) \} g(t),
$$

$$
H^{ij}(t,x) = G_{kl}(t,x) g^{ik}(t,x) g^{jl}(t,x) \sqrt{\det(g_{ij}(t,x))}.
$$

(3.6)

It follows from (3.1) that

$$
\frac{d}{dt} E_m(t,X) - \frac{d}{dt} E_m(0,X) = \int_{B_\sigma(x_0)} (u^*_m h)_{ij}(x) (H^{ij}(t,x) - H^{ij}(0,x)) dx,
$$

(3.7)

where

$$(u^*_m h)_{ij}(x) = \sum_{\alpha=1}^K \frac{\partial u^\alpha_m}{\partial x^i}(x) \frac{\partial u^\alpha_m}{\partial x^j}(x).$$

(3.8)

Hence,

$$
\left| \frac{d}{dt} E_m(t,X) - \frac{d}{dt} E_m(0,X) \right|
\leq 6 \sum_{i,j=1}^3 \left( \sup_{x \in B_\sigma(x_0)} |H^{ij}(t,x) - H^{ij}(0,x)| \right) \cdot \left( \int_{B_\sigma(x_0)} |d u_m|^2 dV \right)
$$

(3.9)

by the Cauchy-Schwartz inequality and (3.4). We note that $H^{ij}(t,x)$ is a known function of $\{g_{ij}(t,x)\}$ and $\{(d/dt) g_{ij}(t,x)\}$, while $g(t,x) = F_\epsilon^x g(x)$ and $(d/dt) g(t,x) = F_\epsilon^x (\mathcal{L} g)(x)$. Since $\|X\|_{C^1} \leq \Lambda$, it follows from the standard ODE theory that, for any $\epsilon > 0$, there exists $t_0$ depending only on $\epsilon, \Lambda$, and $g$ so that, for any $t \in [-t_0, t_0]$, we have that $\|g(t) - g\|_{C^1} \leq \epsilon$, hence $|H^{ij}(t,x) - H^{ij}(0,x)| \leq C \epsilon$ for some constant $C$ depending only on the algebraic expression of $H^{ij}$.

Now assume that the lemma is not true for $V_{\Lambda,\sigma}$; then there exist $\delta_0 > 0$, a sequence of $\{X_k\} \subset V_{\Lambda,\sigma}$, and a subsequence $\{u_{ik}\}$ of $\{u_i\}$ such that

$$
\left| \frac{d}{dt} E_{ik}(0,X_k) \right| > \delta_0.
$$

(3.10)
Our above analysis then shows that there exists \( t_0 = t_0(\delta_0, g, \Lambda, E_\phi) \) such that

\[
\left| \frac{d}{dt} E_{ik}(t, X_k) \right| > \frac{1}{2} \delta_0 \quad \forall t \in [-t_0, t_0].
\] (3.11)

Since \( \lim_{k \to \infty} E(u_{ik}) = E_\phi \), we conclude that, for some \( k \) large enough and some \( t \in [-t_0, t_0] \), \( E(u_{ik} \circ F_t) < E_\phi - (1/4)\delta_0 t_0 \), which is a contradiction to the fact \( u_{ik} \circ F_t \in \mathcal{F}_\phi \) and the definition of \( E_\phi \).

To replace \( V_{\Lambda, \sigma} \) by \( V_{\Lambda, \sigma} \), we can simply apply a partition of unity argument considering that \( (d/dt) E_n(0, X) \) is linear in \( X \). Hence, the lemma is proved. \( \square \)

Now we are ready to derive an approximate monotonicity property for \( \{u_i\} \). Let \( \xi(t) \) be any \( C^1 \) decreasing function on \([0, +\infty)\) whose support lies in \([0, 1]\). We fix \( x_0 \in M \) and let \( \{x^1, x^2, x^3\} \) be a geodesic normal coordinate chart in \( B_\sigma(x_0) \). For \( 0 < \rho < \sigma \leq \sigma_0 \) and \( x \in B_\sigma(x_0) \), we define \( X_\rho(x) = \xi(|x|/\rho) x^i (\partial / \partial x^i) \) and view \( X_\rho \) as a vector field defined globally on \( M \). It is easily checked that \( \|X_\rho\|_{C^1} \leq \Lambda \) for some constant \( \Lambda = \Lambda(\xi) > 0 \).

Thus Lemma 3.1 implies that there exists a sequence \( \{\kappa_i\} \) depending on \( \Lambda(\xi) \) but not on \( \rho \) such that

\[
\left| \frac{d}{dt} E_i(0, X_\rho) \right| \leq \kappa_i, \quad \lim_{i \to \infty} \kappa_i = 0.
\] (3.12)

A direct calculation shows that

\[
\begin{align*}
\frac{d}{dt} E_i(0, X_\rho) &= \text{error}(\rho) + (-1) \int_{B_\sigma(x_0)} \xi\left(\frac{|x|}{\rho}\right) |du_i|^2 dV \\
&\quad + (-1) \int_{B_\sigma(x_0)} \xi'\left(\frac{|x|}{\rho}\right) \left(\frac{|x|}{\rho}\right) |du_i|^2 dV \\
&\quad + 2 \int_{B_\sigma(x_0)} \xi'\left(\frac{|x|}{\rho}\right) \left(\frac{|x|}{\rho}\right) \frac{\partial u_i}{\partial y}^2 dV,
\end{align*}
\] (3.13)

where \( v = (x^i/|x|)(\partial / \partial x^i), |\text{error}(\rho)| \leq \tilde{c} \rho^2 (\int_{B_\sigma(x_0)} |du_i|^2 dV) \), and \( \tilde{c} = \tilde{c}(\xi, g) \). We then define

\[
E_i(\rho) = \frac{1}{\rho} \int_{B_\sigma(x_0)} \xi\left(\frac{|x|}{\rho}\right) |du_i|^2 dV.
\] (3.14)

It follows from (3.12) and (3.13) that

\[
E'_i(\rho) + \tilde{c} \int_{B_\sigma(x_0)} |du_i|^2 dV \geq -\kappa_i \frac{1}{\rho^2},
\] (3.15)

which gives that

\[
E_i(\tau) \leq E_i(\rho) + \tilde{c}(\rho - \tau) \int_{B_\sigma(x_0)} |du_i|^2 dV + \kappa_i \left(\frac{1}{\tau} - \frac{1}{\rho}\right)
\] (3.16)

for any \( 0 < \tau < \rho < \sigma \). Hence, we have proved the following proposition.
\textbf{Proposition 3.2.} For any $C^1$ decreasing function $\xi(t)$ with support in $[0,1]$, there exists a sequence $\{\kappa_i\}$ such that $\lim_{i \to 0} \kappa_i = 0$ and
\begin{equation}
\frac{1}{\tau} \int_{B_\sigma(x_0)} \xi\left(\frac{|x|}{\tau}\right) |d\mu_i|^2 dV \leq \frac{1}{\rho} \int_{B_\sigma(x_0)} \xi\left(\frac{|x|}{\rho}\right) |d\mu_i|^2 dV + \bar{c}(\rho - \tau) \int_{B_\sigma(x_0)} |d\mu_i|^2 dV + \kappa_i \left(1 - \frac{1}{\rho}\right) \tag{3.17}\end{equation}
for any $x_0 \in M$ and any $0 < \tau < \rho < \sigma \leq \sigma_0$. Here $\{\kappa_i\}$ is independent of $\rho$ and $\tau$, and $\bar{c}$ is a constant depending only on $\xi$ and $g$.

Letting $i$ go to $\infty$, we have the following “monotonicity” formula for the limiting measure $\mu$.

\textbf{Corollary 3.3.} For any $C^1$ decreasing function $\xi(t)$ with its support in $[0,1]$,
\begin{equation}
\frac{1}{\tau} \int_{B_\sigma(x_0)} \xi\left(\frac{|x|}{\tau}\right) d\mu \leq \frac{1}{\rho} \int_{B_\sigma(x_0)} \xi\left(\frac{|x|}{\rho}\right) d\mu + \bar{c} \mu(B_\sigma(x_0)) \tag{3.18}\end{equation}
for any $x_0 \in M$, $0 < \tau < \rho < \sigma \leq \sigma_0$, and some constant $\bar{c} = \bar{c}(\xi,g)$. Choosing $\xi$ to be 1 on $[0,1/2]$,
\begin{equation}
\frac{1}{\tau} \mu(B_\tau(x_0)) \leq \frac{2}{\rho} \mu(B_\rho(x_0)) + \bar{c} \mu(B_\sigma(x_0)) \tag{3.19}\end{equation}
for any $0 < 2\tau < \rho < \sigma \leq \sigma_0$, where $\bar{c} = \bar{c}(g)$.

As an application of this “monotonicity” property of $\mu$, we show that $u$ can be well approximated by smooth maps into $N$ from the region where $\{u_i\}$ has small energy concentration.

\textbf{Proposition 3.4.} There exists a number $\epsilon_1$ depending only on $M$ and $N$ such that if $\mu(B_\sigma(x_0))/\sigma < \epsilon_1$, then there exists a sequence of smooth maps $\{u_\tau\}_{0 < \tau < \tau_0}$ from $B_{\sigma/2}(x_0)$ to $N$ such that $\lim_{\tau \to 0} \|u_\tau - u\|_{W^{1,2}(B_{\sigma/2}(x_0))} = 0$.

\textbf{Proof.} We use the idea in [5] to mollify $u$. Let $\varphi : \mathbb{R}^3 \to \mathbb{R}^+$ be a smooth radial mollifying function so that support $(\varphi) \subset B_1$ and $\int_{\mathbb{R}^3} \varphi \, dx = 1$. Assume that $\mu(B_\sigma(x_0))/\sigma < \epsilon_1$ for some $\epsilon_1$ to be determined later; by \textbf{Corollary 3.3}, we have that
\begin{equation}
\frac{\mu(B_\tau(y))}{\tau} \leq 4 \frac{\mu(B_{\sigma/2}(y))}{\sigma} + \bar{c} \sigma \epsilon_1 \leq 4 \frac{\mu(B_\sigma(x_0))}{\sigma} + \bar{c} \sigma \epsilon_1 \leq 5 \epsilon_1 \tag{3.20}\end{equation}
for any $y \in B_{\sigma/2}(x_0)$ and $0 < 2\tau < \sigma/2$ provided $\bar{c} \sigma \leq 1$. Now define $u^\tau(y) = (1/\tau^3) \int_{B_\tau(y)} \varphi(|y-z|/\tau) u(z) \, dz$ inside a normal coordinate chart around $x_0$; we can apply a version of the Poincare inequality to assert that
\begin{equation}
\frac{1}{\tau^3} \int_{B_\tau(y)} |u(x) - u^\tau(y)|^2 \, dx \leq c_2 \frac{1}{\tau} \int_{B_\tau(y)} |du|^2 \, dx \leq c_2 \frac{\mu(B_\tau(y))}{\tau}, \tag{3.21}\end{equation}
where the last inequality holds because of the lower semicontinuity of energy with respect to weak convergence. It follows from (3.20) and (3.21) that $u^\tau(y)$ lies near...
many values of $u(z)$ for $z \in B_\tau(y)$. In particular, we see that
\[ \text{dist}(u^T(y), N) \leq c_3 \varepsilon_1^{1/2}. \] (3.22)

Let $\mathcal{C}_\varepsilon$ be a $\varepsilon$-tubular neighborhood of $N$ in $\mathbb{R}^K$, and let $\Phi : \mathcal{C}_\varepsilon \to N$ denote the smooth nearest point projection map. We see that if $c_3 \varepsilon_1^{1/2} < \varepsilon$, then $u^T(y) \in \mathcal{C}_\varepsilon$ for all $y \in B_{\sigma/2}(x_0)$. Hence, we can define a smooth map $u : B_{\sigma/2}(x_0) \to N$ by $u^T(y) = \Phi \circ u^T(y)$. Since $u^T(y)$ is the standard mollification of $u$ by $\varphi$ with a scaling factor $\tau$, we see immediately that $\lim_{\tau \to 0} \|u^T - u\|_{W^{1,2}(B_{\sigma/2}x_0)} = 0$.

4. A continuous version of Luckhaus lemma. In this section, we use $\nabla (\cdot)$ to denote the gradient operator on $S^2 \subset \mathbb{R}^3$ and $d\omega$ to denote the Euclidean surface measure on $S^2$.

For a map $u$ defined on a cylinder $[a, b] \times S^2$, we use $\nabla_x u$, $\nabla_t u$ to denote the partial $x$, $t$ gradient of $u$, where $(t, x) \in [a, b] \times S^2$. The following technical lemma, which may be viewed as a continuous version of the 2-dimensional Luckhaus lemma (see [6]) in the study of energy minimizing maps, will help us construct comparison maps in the proof of the main theorem.

**Lemma 4.1.** Assume that $N \subset \mathbb{R}^K$ is an isometrically embedded compact manifold. Then there exists $\varepsilon_2 = \varepsilon_2(N) > 0$ such that if $v, w \in W^{1,2}(S^2, N) \cap C^0(S^2, N)$ and
\[ \int_{S^2} |\nabla v|^2 d\omega \leq \varepsilon_2, \quad \int_{S^2} |\nabla w|^2 d\omega \leq \varepsilon_2, \] (4.1)
then for all $\beta > 0$, there exists $\eta = \eta(\varepsilon_2, \beta) > 0$, where $\eta$ does not depend on the choice of $v$ and $w$, such that if
\[ \int_{S^2} |v - w|^2 d\omega < \eta, \] (4.2)
then there exist $\beta' \in [0, \beta)$ and $v' \in W^{1,2}([0, \beta'] \times S^2, N) \cap C^0([0, \beta'] \times S^2, N)$ with properties that
\[ v'(0, x) = v(x), \quad v'((0, \beta'), x) = w(x), \]
\[ \int_{[0, \beta'] \times S^2} |\nabla(u(x) v)|^2 d\omega dt \leq \beta. \] (4.3)

**Proof.** Let $v, w \in W^{1,2}(S^2, N) \cap C^0(S^2, N)$ such that (4.1) holds for some $\varepsilon_2$ to be determined later. Let $\varphi : \mathbb{R}^2 \to \mathbb{R}^+$ be a smooth radial mollifying function so that support$(\varphi) \subset B_1$ and $\int_{\mathbb{R}^2} \varphi d\omega = 1$. For any $0 < h \ll \pi/2$ and any $(t, x) \in (0, h] \times S^2$, we define
\[ v(t, x) = \int_{S^2} v(y) \varphi^t(\text{dist}(x, y)) d\omega(y), \] (4.4)
where dist$(x, y)$ represents the sphere distance between $x$ and $y$ on $S^2$ and $\varphi^t(r) = (1/t^2) \varphi(r/t)$. Let $\mathcal{C}_2\varepsilon$ be a $2\varepsilon$-tubular neighborhood of $N$ in $\mathbb{R}^K$; by the argument used in the proof of Proposition 3.4, we know that, if we choose $\varepsilon_2 = \varepsilon_2(N)$ to be sufficiently small, then $v(t, x) \in \mathcal{C}_\varepsilon$ for all $(t, x) \in (0, h] \times S^2$. (We note that the monotonicity of the
energy of \( v \), which is crucial in the proof of Proposition 3.4, is automatically satisfied in this case because the domain of \( v \) is of 2-dimensional. Since \( v \) is continuous on \( S^2 \), we have that
\[
\lim_{{(t,z) \to (0,\infty)}} v(t,z) = v(x).
\]
(4.5)

Thus \( v(t,x) \) is a continuous map on the closed cylinder \([0,h] \times S^2\) with \( v(0,x) = v(x) \). On the other hand, if \( B_{\sigma}(z) \) is a geodesic ball with a normal coordinate chart such that \( x \in B_{\sigma/2}(z) \), then for \( 0 < h = h(\epsilon_2) \ll 1 \), we have that
\[
v(t,x) = \int_{S^2} v(y) \varphi^t(\text{dist}(x,y))\,d\omega(y) \approx \int_{B_{\sigma}(0) \subset \mathbb{R}^2} v(x-ty)\varphi(|y|)\,dy,
\]
(4.6)

which implies that
\[
\left| \nabla_x v(t,x) \right|^2 \leq c_4 \int_{S^2} \left| \nabla v(y) \right|^2 \varphi^t(\text{dist}(x,y))\,d\omega(y) + \epsilon_2,
\]
\[
\left| \nabla_t v(t,x) \right|^2 \leq c_4 \int_{S^2} \left| \nabla v(y) \right|^2 \varphi^t(\text{dist}(x,y))\,d\omega(y) + \epsilon_2
\]
(4.7)

by the Cauchy-Schwarz inequality. Then it follows from (4.7) that
\[
\int_{[0,h] \times S^2} \left| \nabla_{(t,x)} v(t,x) \right|^2\,d\omega(x)\,dt \leq c_5 h \left( \int_{S^2} \left| \nabla v(y) \right|^2\,d\omega(y) + \epsilon_2 \right)
\]
(4.8)

by the Fubini theorem. Similarly, we define \( w(t,x) : [l+h, l+2h] \times S^2 \to \mathcal{C}_\ell \) by
\[
w(t,x) = \int_{S^2} w(y) \varphi^{l+2h-t}(\text{dist}(x,y))\,d\omega(y)
\]
(4.9)

for some \( l \) determined later.

Now we want to connect \( v(h,x) \) and \( w(l+h,x) \) on \([h, l+h] \times S^2\). We first estimate \( |v(h,x) - w(l+h,x)| \) pointwise. It follows from the definition and the Cauchy-Schwarz inequality that
\[
|v(h,x) - w(l+h,x)| = \left( \int_{S^2} \left| \nabla v(y) - \nu(y) \right| \varphi^h(\text{dist}(x,y))\,d\omega(y) \right)^{1/2}
\]
\[
\leq c_6 \frac{1}{h} \left( \int_{S^2} \left| \nabla v(y) - \nu(y) \right|^2\,d\omega(y) \right)^{1/2}.
\]
(4.10)

We define \( z(t,x) \) on \([h, l+h] \times S^2\) to be
\[
z(t,x) = \left( \frac{t-h}{l} \right) w(l+h,x) + \left( \frac{l+h-t}{l} \right) v(h,x).
\]
(4.11)

Then (4.7) and (4.10) imply that
\[
\int_{[h, l+h] \times S^2} \left| \nabla_{(t,x)} z(t,x) \right|^2\,d\omega(x)\,dt \leq c_7 l \left\{ \int_{S^2} \left[ |\nabla v|^2 + |\nabla w|^2 \right] d\omega + 2\epsilon_2 \right\}
\]
\[
+ c_7 \frac{1}{l} \frac{1}{h^2} \int_{S^2} \left| \nabla v(y) - \nu(y) \right|^2\,d\omega(y).
\]
(4.12)
Now we consider

\[ \tilde{v} = \begin{cases} 
  v(t, x), & 0 \leq t \leq h, \\
  z(t, x), & h \leq t \leq l + h, \\
  w(t, x), & l + h \leq t \leq l + 2h.
\] (4.13)

Clearly, \( \tilde{v} \in C^0 \cap W^{1,2}([0, l + 2h] \times S^2, \mathbb{R}^K) \). Furthermore,

\[
\int_{[0,l+2h] \times S^2} | \nabla_{(t,x)} \tilde{v}(t,x) |^2 \, d\omega dt \leq c_8 \frac{1}{h} \int_{S^2} |v(y) - w(y)|^2 \, d\omega(y) 
+ c_8 (h + l) \epsilon_2.
\] (4.14)

For any \( \beta > 0 \), we first choose \( h = h(\beta, \epsilon_2) \) and \( l = l(\beta, \epsilon_2) \) such that

\[ l + 2h < \beta, \quad c_8 (h + l) \epsilon_2 < \frac{\beta}{2}, \] (4.15)

then we let \( \|v - w\|_{L^2(S^2)} < \eta \), where \( \eta = \eta(l, h, \beta, \epsilon_2) \) is so small that

\[ c_8 \frac{1}{l} \frac{1}{h^2} \eta < \frac{\beta}{2}, \quad c_0 \frac{1}{h} \eta^{1/2} < \epsilon. \] (4.16)

It follows from (4.10) and (4.14) that \( \tilde{v}(t,x) \in \mathcal{O}_{2\epsilon} \) for all \( (t,x) \in [0, l + 2h] \times S^2 \) and the total energy of \( \tilde{v} \) is bounded by \( \beta \). To get \( v' \) finally, we compose \( \tilde{v} \) with the nearest point projection map \( \Phi : \mathcal{O}_{2\epsilon} \to N \). Hence, the lemma is proved. \( \Box \)

5. Proof of Theorem 1.1. Throughout this section, we fix a geodesic ball \( B_\sigma(x_0) \), where

\[
\frac{1}{\sigma} \mu(B_\sigma) = \lim_{i \to \infty} \frac{1}{\sigma} \int_{B_\sigma} |du_i|^2 \, dV < \epsilon_0
\] (5.1)

for some \( \epsilon_0 \) to be determined. For each \( \tau \), we let \( B_\tau \) denote \( B_\tau(x_0) \).

Assume that \( \epsilon_0 < \epsilon_1 \); Proposition 3.4 implies that there exists a sequence \( \{v_i\} \subset C^\infty(B_{\sigma/2}, N) \) such that

\[
\lim_{i \to \infty} \|v_i - u\|_{W^{1,2}(B_{\sigma/2})} = 0.
\] (5.2)

We then choose \( \rho \in (\sigma/4, \sigma/2) \) such that \( u \mid_{\partial B_\rho}, u_i \mid_{\partial B_\rho} \in W^{1,2}(\partial B_\rho, N), \)

\[
\lim_{i \to \infty} \|v_i - u\|_{W^{1,2}(\partial B_\rho)} = 0, \\
\lim_{i \to \infty} \|u_i - u\|_{L^2(\partial B_\rho)} = 0,
\]

\[
\int_{\partial B_\rho} |du|^2 \, d\Sigma \leq c_9 \frac{1}{\sigma} \int_{B_\rho} |du|^2 \, dV, \\
\liminf_{i \to \infty} \int_{\partial B_\rho} |du_i|^2 \, d\Sigma \leq c_9 \frac{1}{\sigma} \lim_{i \to \infty} \int_{B_\rho} |du_i|^2 \, dV,
\] (5.3)
where $d\Sigma$ is the induced surface measure on $\partial B_\rho \subset M$. After fixing a geodesic normal coordinate chart centered at $x_0$, we may also view $u|_{\partial B_\rho}, u_i|_{\partial B_\rho}, v_i|_{\partial B_\rho}$ as defined on the Euclidean sphere $S_\rho$ with radius $\rho$; then we have that

\begin{align}
\int_{S_\rho} |\nabla u|^2 d\omega_\rho &\leq c_{10} \int_{\partial B_\rho} |du|^2 d\Sigma, \\
\int_{S_\rho} |\nabla u_i|^2 d\omega_\rho &\leq c_{10} \int_{\partial B_\rho} |du_i|^2 d\Sigma, \\
\int_{S_\rho} |\nabla v_i|^2 d\omega_\rho &\leq c_{10} \int_{\partial B_\rho} |dv_i|^2 d\Sigma,
\end{align}

(5.4)

where $d\omega_\rho$ represents the Euclidean surface measure on $S_\rho$. Now we choose $\epsilon_0 < \min\{\epsilon_1, (c_9 c_{10})^{-1} \epsilon_2\}$ and we claim that

\begin{align}
\lim_{i \to \infty} \int_{B_\rho} |du_i|^2 dV = \int_{B_\rho} |du|^2 dV.
\end{align}

(5.5)

We remark that, once (5.5) is established, it will readily imply that

\begin{align}
\lim_{i \to \infty} \|u_i - u\|_{W^{1,2}(B_\rho)} = 0
\end{align}

(5.6)

by the fact that $\{u_i\}$ converges to $u$ weakly and the standard Hilbert space theories.

Assume that (5.5) does not hold; we have that

\begin{align}
\lim_{i \to \infty} \int_{B_\rho(x_0)} |du_i|^2 dV > \int_{B_\rho(x_0)} |du|^2 dV + \delta
\end{align}

(5.7)

for some $\delta > 0$. The idea for the rest of the proof is to construct a sequence of comparison maps which are almost $v_i$ inside $B_\rho$ and $u_i$ outside $B_\rho$. For that purpose, we need to connect $u_i$ and $v_i$ on the boundary of $B_\rho$ using Lemma 4.1. We first note that (5.3) imply that there exists a subsequence $\{u_{ik}\}, \{v_{ik}\}$ such that

\begin{align}
\int_{S_\rho} |\nabla u_{ik}|^2 d\omega_\rho < \epsilon_2, \quad \int_{S_\rho} |\nabla v_{ik}|^2 d\omega_\rho < \epsilon_2 \quad \forall k, \\
\lim_{k \to \infty} \|v_{ik} - u_{ik}\|_{L^2(S_\rho)} = 0.
\end{align}

(5.8)

We then consider $\tilde{u}_{ik}(\omega) = u_{ik}(\rho \omega)$ and $\tilde{v}_{ik}(\omega) = v_{ik}(\rho \omega)$, where $\omega$ denotes the point on $S^2$. It follows from (5.8) and Lemma 4.1 that for all $\beta > 0$, there exist $k_0 = k_0(\beta, \epsilon_2)$ and $\beta' = \beta'(\beta, \epsilon_2) < \beta$ such that for all $k > k_0$, there exists $\tilde{w}_k \in W^{1,2} \cap C^0([0, \beta'] \times S^2, N)$ such that

\begin{align}
\tilde{w}_k(0, x) &= \tilde{u}_{ik}(x), \\
\tilde{w}_k(\beta', x) &= \tilde{v}_{ik}(x), \\
\int_{[0, \beta'] \times S^2} |\nabla (t, x) \tilde{w}_k|^2 d\omega dt &\leq \beta.
\end{align}

(5.9)
Next, we use polar coordinates to transplant \( \bar{w}_k \) to the shell region between \( S_\rho \) and \( S(1-\beta')\rho \) by defining \( w_k((1-t)\rho \omega) = \bar{w}_k(t,\omega) \). Rescaling \( v_{ik}(x) \) on \( B_\rho \) to \( v'_{ik}(x) = v_{ik}(r \omega/(1-\beta')) \) on \( B(1-\beta')\rho \), we then have that

\[
\begin{align*}
    w_k(x) &= u_{ik}(x), \quad x \in \partial B_\rho, \\
    w_k(x) &= v'_{ik}(x), \quad x \in \partial B(1-\beta')\rho, \\
    \int_{B_\rho \setminus B(1-\beta')\rho} |dw_k|^2 \, dV &\leq c_{10} \rho \beta. 
\end{align*}
\]

Now we consider a new sequence \( \{\hat{u}_k\} \subset W^{1,2} \cap C^0(M,N) \) given by

\[
\hat{u}_k = \begin{cases} 
    u_{ik}, & x \notin B_\rho, \\
    w_k, & x \in B_\rho \setminus B(1-\beta')\rho, \\
    v'_{ik}, & x \in B(1-\beta')\rho. 
\end{cases}
\]

First, we note that the fact that \( \hat{u}_k = u_{ik} \) outside \( B_\rho \) and \( \pi_2(S^3) = 0 \) implies that \( \hat{u}_k \) is 2-homotopic to \( u_{ik} \). Second, we have the following energy estimate:

\[
E(\hat{u}_k) = \int_{M \setminus B_\rho} |du_{ik}|^2 \, dV + \int_{B_\rho \setminus B(1-\beta')\rho} |dw_k|^2 \, dV + \int_{B(1-\beta')\rho} |dv'_{ik}|^2 \, dV 
\leq E(u_{ik}) - \int_{B_\rho} |du_{ik}|^2 \, dV + c_{10} \rho \beta + c(\beta') \int_{B_\rho} |dv'_{ik}|^2 \, dV, 
\]

where \( c(\beta') \) is the supremum of the Jacobian of the scaling diffeomorphism from \( B(1-\beta')\rho \) to \( B_\rho \), which satisfies \( \lim_{\beta' \to 0} c(\beta') = 1 \) with the convergence only depending on \((M,g)\). We now fix \( \beta \) such that

\[
\beta < (c_{10} \rho)^{-1} \delta, \quad |c(\beta') - 1| \int_{B_\rho(x_0)} |du|^2 \, dV < \frac{\delta}{2}. 
\]

Letting \( k \to \infty \), we then have that

\[
\limsup_{k \to \infty} E(\hat{u}_k) \leq E_\Phi - \frac{\delta}{2}. 
\]

Finally, we note that the fact that \( \hat{u}_k \in W^{1,2} \cap C^0(M,N) \) implies that \( \hat{u}_k \) can be well approximated in \( W^{1,2} \) norm by smooth maps from \( M \) to \( N \) which are homotopic to \( \bar{u}_k \). One way to see this is to consider the standard mollification of \( \hat{u}_k \) into \( \mathbb{R}^K \), where the uniform continuity of \( \hat{u}_k \) on \( M \) will guarantee that the image of the mollification will be inside a tubular neighborhood of \( N \). Composing it with the nearest point projection map, we then have the desired approximation. Hence, we know that there exists another sequence \( \{\tilde{u}_k\} \subset C^\infty(M,N) \) such that \( \tilde{u}_k \) is homotopic to \( \hat{u}_k \) and

\[
\limsup_{k \to \infty} E(\tilde{u}_k) = \limsup_{k \to \infty} E(\hat{u}_k) < E_\Phi - \frac{\delta}{2}. 
\]
Since $\tilde{u}_k$ and $u_{i_k}$ are 2-homotopic, we know that $\{\tilde{u}_k\} \subset \mathcal{F}_\phi^{(2)}$. Thus (5.15) gives a contradiction to the fact that $E_\phi = \inf\{E(u) \mid u \in \mathcal{F}_\phi^{(2)}\}$ by White’s theorem. Therefore, (5.5) holds and Theorem 1.1 is proved.

Next, we prove Corollary 1.2 on the partial regularity of the weak limit of $\{u_t\}$. We recall the following $\epsilon$-regularity theorem for stationary harmonic maps obtained by Bethuel [1].

**Bethuel’s theorem.** There exists a number $\epsilon_3 = \epsilon_3(M,N) > 0$ such that if $u : B_\sigma(x_0) \subset M \to N$ is a stationary harmonic map and $(1/\sigma) \int_{B_\sigma(x_0)} |du|^2 dV \leq \epsilon_3$, then $u$ is smooth inside $B_{\sigma/2}(x_0)$.

**Proof of Corollary 1.2.** Let $\bar{\epsilon}$ be a number to be determined and let $B_\sigma(x_0)$ be a geodesic ball, where $(1/\sigma) \mu(B_\sigma(x_0)) < \bar{\epsilon}$. Assume that $\epsilon < \epsilon_0$; our main theorem implies that $\{u_k\}$ converges strongly to $u$ in $W^{1,2}(B_{\sigma/4}(x_0), N)$. Then it follows from this $W^{1,2}$ strong convergence and the fact that $\{u_i\}$ is a minimizing sequence that $u$ is stationary with respect to both the first and the second variation (see [4]) inside $B_{\sigma/4}(x_0)$, hence $u : B_{\sigma/4}(x_0) \to N$ is a stationary harmonic map. We note that

$$
\frac{4}{\sigma} \int_{B_{\sigma/4}(x_0)} |du|^2 dV \leq \frac{4}{\sigma} \int_{B_\sigma(x_0)} |du|^2 dV \leq \frac{4}{\sigma} \mu(B_\sigma(x_0)).
$$

Hence, assuming that $\bar{\epsilon} < (1/4)\epsilon_3$ and applying Bethuel’s theorem, we know that $u$ is smooth inside $B_{\sigma/8}(x_0)$. With such a choice of $\bar{\epsilon}$, we define

$$
\Sigma = \left\{ x \in M \mid \lim_{\sigma \to 0} \frac{1}{\sigma} \mu(B_\sigma(x)) \geq \bar{\epsilon} \right\}.
$$

A standard covering argument (see [4]) then shows that $\Sigma$ is a closed set with finite 1-dimensional Hausdorff measure. Hence, we conclude that $u$ is a smooth harmonic map from $M \setminus \Sigma$ to $N$, where $\Sigma$ is a close set with $\mathcal{H}^1(\Sigma) < \infty$.

**Acknowledgments.** The author wants to thank Professor Richard Schoen for suggesting this problem. The author also wants to thank the referees for careful reading of the original manuscript.

**References**


Pengzi Miao: Department of Mathematics, Stanford University, Palo Alto, CA 94305, USA

*E-mail address: mpengzi@math.stanford.edu*