RIGIDITY OF $\varepsilon$-HARMONIC MAPS OF LOW DEGREE

JASMIN HÖRTER, TOBIAS LAMM, AND MARIO MICALLEF

Abstract. In 1981, Sacks and Uhlenbeck introduced their famous $\alpha$-energy as a way to approximate the Dirichlet energy and produce harmonic maps from surfaces into Riemannian manifolds. However, the second and third authors together with Malchiodi ([11], [12]) showed that for maps between two-spheres this method does not capture every harmonic map. They established a gap theorem for $\alpha$-harmonic maps of degree zero and also showed that below a certain energy bound $\alpha$-harmonic maps of degree one are rotations. We establish similar results for $\varepsilon$-harmonic maps $u_\varepsilon: S^2 \to S^2$, which are critical points of the $\varepsilon$-energy introduced by the second author in [9]. In particular, we similarly show that $\varepsilon$-harmonic maps of degree zero with energy below $8\pi$ are constant and that maps of degree $\pm 1$ with energy below $12\pi$ are of the form $R\pi$ with $R \in O(3)$. Moreover, we construct non-trivial $\varepsilon$-harmonic maps of degree zero with energy $> 8\pi$.

1. Introduction

The Dirichlet energy $E(u)$ of a map $u \in W^{1,2}(M, N)$ from a smooth, closed two-dimensional Riemannian manifold $(M^2, g)$ to a smooth, closed Riemannian manifold $(N^n, h)$ which is isometrically embedded into some $\mathbb{R}^k$ is defined as

$$E(u) := \frac{1}{2} \int_M |\nabla u|^2 dA_M.$$  

Critical points of this functional are called harmonic maps and satisfy

$$\Delta u \perp T_uN.$$  

In 1981 Sacks and Uhlenbeck [14] introduced their famous $\alpha$-energy approximation

$$E_\alpha(u) = \frac{1}{2} \int_M (1 + |\nabla u|^2)^\alpha dA_M, \quad \alpha > 1.$$  

Since $\alpha > 1$, this functional has better compactness properties than the Dirichlet energy, which allowed Sacks and Uhlenbeck to show that, as $\alpha \to 1$, a sequence of critical points $(u_\alpha)$ converges to a harmonic map and finitely many bubbles, i.e. non-trivial two-spheres. Now one can ask whether every harmonic map can be captured by this procedure and the answer is no under suitable energy assumptions. Together with Malchiodi, the second and third authors showed ([11], [12]) that the only $\alpha$-harmonic maps $u: S^2 \to S^2$ of degree $\pm 1$ with energy $E_\alpha$ below $8\alpha^2\pi$ are of the form $R\pi$, $R \in O(3)$ and $\alpha$-harmonic maps of degree zero with energy below $6\alpha^2\pi$ are constant. Moreover, they constructed non-constant $\alpha$-harmonic maps of degree zero with energy slightly above $6\alpha^2\pi$, establishing a gap theorem.

In the following we pose the same question for the fourth order approximation of the Dirichlet energy

$$E_\varepsilon(u) = \frac{1}{2} \int_M (|\nabla u|^2 + \varepsilon |\Delta u|^2) dA_M, \quad \varepsilon > 0.$$  

Date: June 1, 2022.

The first and second author gratefully acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 281869850 (RTG 2229).
This approximation was first studied by the second author in [9], where he showed that critical points \( u_\varepsilon \in W^{2,2}(M, N^n) \) exist for every \( \varepsilon > 0 \) and that sequences of critical points satisfy the same bubbling picture as the \( \alpha \)-harmonic maps studied earlier (see Theorem 1.1 in [9]). Further, critical points are smooth and satisfy

\[
\Delta u - \varepsilon \Delta^2 u = \varepsilon \sum_{i=n+1}^k \left( \Delta ((\nabla u, (d\nu_i \circ u)\nabla u)) + \text{div}(\Delta u, (d\nu_i \circ u)\nabla u) \right) 
+ (\nabla \Delta u, (d\nu_i \circ u)\nabla u)) \nu_i \circ u - A(u)(\nabla u, \nabla u),
\]

where \( \{\nu_i\}_{i=n+1}^k \) is a smooth orthonormal frame of the normal space of \( N \) and \( A \) is the second fundamental form of the embedding \( N \to \mathbb{R}^k \). In the following we are only interested in maps \( u: S^2 \to S^2 \). Thus the equation simplifies to

\[
\Delta u - \varepsilon \Delta^2 u = -u|\nabla u|^2 + \varepsilon u \left( \Delta |\nabla u|^2 + \text{div}(\Delta u, \nabla u) + \langle \nabla \Delta u, \nabla u \rangle \right)
\]

The degree of every map \( u: S^2 \to S^2 \) is defined by

\[
\text{deg}(u) := \frac{1}{4\pi} \int_{S^2} J(u) dA_{S^2} \quad \text{with} \quad J(u) = u \cdot e_1(u) \wedge e_2(u),
\]

where \( (e_1, e_2) \) is a local oriented orthonormal frame of \( TS^2 \). For every \( u \in W^{2,2}(S^2, S^2) \) with \( \text{deg}(u) = 1 \) we have

\[
4\pi(1 + 2\varepsilon) = \int_{S^2} J(u) dA_{S^2} + \frac{\varepsilon}{2\pi} \left( \int_{S^2} J(u) dA_{S^2} \right)^2 
\leq E(u) + \frac{\varepsilon}{8\pi} \left( \int_{S^2} |\nabla u|^2 dA_{S^2} \right)^2 
\leq E(u) + \frac{\varepsilon}{2} \int_{S^2} |\nabla u|^4 dA_{S^2} 
\leq E_\varepsilon(u), \tag{1.3}
\]

where we used that \( \Delta u = (\Delta u)^T - u|\nabla u|^2 \) and therefore

\[
\int_{S^2} |(\Delta u)^T|^2 dA_{S^2} + \int_{S^2} |\nabla u|^4 dA_{S^2} = \int_{S^2} |\Delta u|^2 dA_{S^2}
\]

in the last step. Thus equality holds in (1.3) if and only if \( u \) is a harmonic map (third inequality and see (1.2)), which is conformal (first inequality) and with constant energy density (second inequality)

\[
e(u) := \frac{1}{2} |\nabla u|^2 \equiv 1.
\]

For every \( R \in SO(3) \) and map \( u^R(x) = Rx \) we have

\[
E_\varepsilon(u^R) = 4\pi + 8\pi \varepsilon. \tag{1.4}
\]

Hence the rotations are the only minimizers of \( E_\varepsilon \) among all maps of degree 1. Note that it was shown by Wood and Lemaire ([(11.5) in [5]]) that all harmonic maps between 2-spheres are precisely the rational maps and their complex conjugates (i.e., rational in \( z \) or \( \bar{z} \)).

A rational map \( u \) has Dirichlet energy \( E(u) = 4\pi|\text{deg}(u)| \), which is the least energy that a map of this degree can have. However one can verify by direct calculation that dilations, which are rational maps of degree one, are not critical points of \( E_\varepsilon \) for \( \varepsilon > 0 \). This is a very special case of the second of the following two main results in this paper, whose proofs occupy the next four sections.
Theorem 1.1. For any $\delta > 0$ there exists $\bar{\varepsilon} > 0$ such that the only critical points $u_\varepsilon$ of $E_\varepsilon$ of degree zero which satisfy $E_\varepsilon(u_\varepsilon) \leq 8\pi - \delta$ and $\varepsilon \leq \bar{\varepsilon}$ are the constant maps.

Theorem 1.2. For any $\mu > 0$ there exists $\bar{\varepsilon} > 0$ such that the only critical points $u_\varepsilon$ of $E_\varepsilon$ of degree $\pm 1$ which satisfy $E_\varepsilon(u_\varepsilon) \leq 12\pi - \mu$ and $\varepsilon \leq \bar{\varepsilon}$ are maps of the form $u^R(x) = Rx$ with $R \in O(3)$.

Note that we have to include reflections if $\deg u = -1$. The proof of Theorem 1.1 follows analogously to [11],[12]. We use the energy identity for $\varepsilon$-harmonic maps (see Theorem 1.1 in [9]) and a result by Duzaar and Kuwert [4], which shows that the degree of a sequence $(u_\varepsilon)$ is preserved in the limit. The gap theorem for $\varepsilon$-harmonic maps with small energy (Lemma 6.1) concludes the proof.

To prove Theorem 1.2, we use the group of conformal transformations of the sphere, which is called the Möbius group. In section 2 we will see that these transformations correspond to $M \in PSL(2, \mathbb{C})$ via stereographic projection to the complex plane. We follow Malchiodi’s and the second and third authors’ idea [11] and apply a Möbius transformation $M$ to a critical point $u_\varepsilon$.

The goal is to show that for every $\varepsilon > 0$ small enough, there exists $M \in PSL(2, \mathbb{C})$ such that $(u_\varepsilon) \mapsto u_\varepsilon \circ M$ is equal to the identity. Moreover, we show further that this $M$ defines a rotation on the sphere.

In a first step we investigate how $E_\varepsilon(u_M)$ changes as we vary $M$. We will see that the transformation relation depends only on the larger eigenvalue $\lambda$ of $MM^*$ and it is therefore enough to demonstrate that $\lambda = 1$. To do this we show that critical points $u_\varepsilon$ are close to a Möbius transformation in the $\sqrt{\varepsilon}W^{2,2}$-norm and simultaneously establish a bound on $\lambda$ (Proposition 3.2). Thanks to the structure of the $\varepsilon$-approximation, the derivative of $E_\varepsilon,\lambda$ with respect to $\log \lambda$ is easy to calculate. This simplifies the proof of the bound on the eigenvalue $\lambda$ significantly compared to the $\alpha$-harmonic case (see Proposition 3.1 in [11]). However, to improve the bound of $\hat{u}_{\varepsilon,\lambda}$ in the $\sqrt{\varepsilon}W^{2,2}$-norm we have to employ rather technical tools and estimates like Sobolev embeddings and the Poincaré inequality.

In the last step we choose a Möbius transformation $M^*$ (not to be confused with the adjoint $M^*$ in the definition of $\lambda$) which is optimal in the sense that $E_\varepsilon(u_M)$ attains a minimum at $M^*$ as $M$ varies over $PSL(2, \mathbb{C})$. This optimal property of $M^*$ is then used to show that the corresponding eigenvalue $\lambda^*$ is equal to one. To do this, we consider the tangential component (which we will call $\hat{\psi}$ in the proof) of $(u_{M^*} - \text{Id})$ at the identity and decompose it into the eigenspaces of the tangential Laplacian $(\Delta)^T$ on vector fields of $S^2$. Then we adapt the ideas in [11] to complete the proof; the key point is that the kernel of the Jacobi operator $J_\varepsilon$ is the same as the kernel of $(\Delta)^T + 2$, which is the tangent space of the Möbius group at the identity (and which we will call $Z$ in the proof). The optimality of $M^*$ allows us to show that the $Z$-component $\hat{\psi}_0$ of $\hat{\psi}$ is much smaller than $\hat{\psi}$; indeed, in suitable norms, $\hat{\psi}_0$ will be shown to be of quadratic order in $\hat{\psi}$. This is a key ingredient in the proof that $\lambda^* = 1$.

In the final section of the paper we prove the following theorem which shows that the bound in Theorem 1.1 is optimal.

Theorem 1.3. For every $\delta > 0$ there exists $\varepsilon_0 > 0$ depending only on $\delta$ such that, if $0 < \varepsilon < \varepsilon_0$, there exists an $\varepsilon$-harmonic map $u_\varepsilon: S^2 \to S^2$ with $\deg(u_\varepsilon) = 0$ and

$$8\pi < E_\varepsilon(u_\varepsilon) < 8\pi + \delta.$$

The map $u_\varepsilon$ in the above theorem is constructed by minimising $E_\varepsilon$ among a suitable class of rotationally symmetric maps.

2. The Möbius group

First we turn our attention to maps of degree one. As mentioned in the introduction, all harmonic maps between two-spheres projected to the complex plane are rational with Dirichlet energy
$E(u) = 4\pi \deg(u)$. Moreover, all rational maps of degree one form a group under composition. This group is called the Möbius group and we shall now describe those of its features that are relevant to us.

Let $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. A holomorphic function $m: \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ of the form

$$m(\xi) = \frac{a\xi + b}{c\xi + d} \quad \text{with} \quad ad - bc = 1, \ a, b, c, d \in \mathbb{C},$$

is called a Möbius transformation. All rational functions of degree one are of this form. We write the coefficients $a, b, c, d$ in matrix form

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{with} \quad \det M = 1.$$ 

The group of such matrices is called the special linear group and it is denoted by $SL(2, \mathbb{C})$. Note that $M$ and $N$ in $SL(2, \mathbb{C})$ represent the same rational map of degree one if, and only if, $M = \pm N$.

Thus, the Möbius group can be identified with the projective special linear group $PSL(2, \mathbb{C})$ defined by

$$PSL(2, \mathbb{C}) := SL(2, \mathbb{C})/\{\pm I_2\}$$

where $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

It is well known that, if we identify $\hat{\mathbb{C}}$ with $S^2$ via stereographic projection, a Möbius transformation can be expressed as a rotation followed by a dilation followed by another rotation. This can be seen from the singular value decomposition of $M$ according to which there exists $U, V \in SU(2)$ such that

$$M = UDV^*$$

where $D$ is the diagonal matrix whose entries $D_{11}, D_{22} > 0$ are the square roots of the eigenvalues of $MM^*$, where $M^*$ is the adjoint of $M$. Since $\det M = 1$ it follows that $\det D = 1$ and thus $D_{11} = \frac{1}{D_{22}}$. By relabelling $D_{11}$ and $D_{22}$ if necessary, we see that $D$ can be written in the form

$$D = \begin{pmatrix} \lambda^2 & 0 \\ 0 & \lambda^{-2} \end{pmatrix}, \quad \lambda \geq 1. \quad (2.2)$$

If $M = D$, the corresponding Möbius transformation is a dilation

$$m_\lambda(\xi) := \lambda \xi.$$ 

We shall now explain the well-known fact that elements of $SU(2)$ in the Möbius group can be identified with rotations. We start by viewing $S^2$ as the complex projective line $\mathbb{CP}^1$ where, if $\sim$ denotes the equivalence relation on $\mathbb{C}^2 \setminus \{(0,0)\}$ defined by $(\sigma, \tau) \sim (p\sigma, p\tau), \ p \neq 0$, then $\mathbb{CP}^1 := (\mathbb{C}^2 \setminus \{(0,0)\})/\sim$. Points in $\mathbb{CP}^1$ shall be written as $[\omega]$, $\omega = (\sigma, \tau) \in \mathbb{C}^2 \setminus \{(0,0)\}$. We next identify $[\omega]$ with the orthogonal projection $\Pi_{[\omega]}$ of $\mathbb{C}^2$ onto the line spanned by $\omega$. Thus $\Pi_{[\omega]}$ is the $2 \times 2$ hermitian matrix of rank 1 and trace equal to 1 defined by

$$\Pi_{[\omega]} := \frac{\omega \otimes \omega^*}{|\omega|^2} = \frac{1}{|\sigma|^2 + |\tau|^2} \begin{pmatrix} |\sigma|^2 & \sigma \tau^* \\ \sigma \tau & |\tau|^2 \end{pmatrix}.$$ 

The action of $PSL(2, \mathbb{C})$ on $S^2$ can then be seen as the following action on the (nonlinear) space $H_1$ of $2 \times 2$ hermitian matrices of rank 1 whose trace is equal to 1:

$$\Pi_{[\omega]} \mapsto \mu_M(\Pi_{[\omega]}) := \Pi_{[M\omega]} = \frac{M(\omega \otimes \omega^*)M^*}{|M\omega|^2}, \quad M \in SL(2, \mathbb{C}). \quad (2.3)$$

Observe that $\mu_M = \mu_{(-M)}$ and therefore, the action $\mu$ just defined is indeed an action of $PSL(2, \mathbb{C})$.\[\text{\ }\]
In order to identify a Möbius transformation as a rotation we need to see $S^2$ as a subset of $\mathbb{R}^3$. To this end, we identify $\mathbb{R}^3$ with the space $H_0$ of traceless Hermitian $2 \times 2$ matrices:

$$H_0 := \left\{ \begin{pmatrix} z & x + iy \\ x - iy & -z \end{pmatrix} : x, y, z \in \mathbb{R} \right\}.$$ 

Note that if $H \in H_0$ then $\det H = -(x^2 + y^2 + z^2)$. We may map matrices from $H_1$ to $H_0$ by subtracting $\frac{1}{2}I_2$ and then, for normalisation purposes, multiplying by 2. Thus,

$$S^2 := 2\frac{\omega \otimes \omega^*}{|\omega|^2} - I_2 = \frac{1}{|\sigma|^2 + |\tau|^2}\begin{pmatrix} |\sigma|^2 - |\tau|^2 & 2\sigma \tau \\ 2\bar{\sigma} \bar{\tau} & |\tau|^2 - |\sigma|^2 \end{pmatrix}, \quad [\omega] = [\sigma, \tau] \in \mathbb{C}P^1.$$ 

For $\tau \neq 0$ we have $[(\sigma, \tau)] = [(\xi, 1)]$ where $\xi = \sigma/\tau$ and the corresponding point in $\mathbb{R}^3$ is then

$$\left( \frac{\xi + \bar{\xi}}{|\xi|^2 + 1}, \frac{-i(\xi - \bar{\xi})}{|\xi|^2 + 1}, \frac{|\xi|^2 - 1}{|\xi|^2 + 1} \right)$$ 

which is the image under the inverse of the stereographic projection from the North Pole of the point $\xi \in \mathbb{C}$. In order to preserve the traceless condition on matrices in $H_0$, the action $\mu$ of $PSL(2, \mathbb{C})$ on $H_1$ described in (2.3) can only be transferred to $H_0$ for those matrices $M$ which are unitary and which therefore (since their determinant is 1) must belong to $SU(2)$. Now for $M \in SU(2)$ and $H \in H_0$ we have $\det(\mu_M H) = \det(MHM^*) = \det H$, which shows that $\mu_M$ is an isometry of $\mathbb{R}^3$ for all $M \in SU(2)$. Using the fact that $SU(2)$ is connected and that $\mu_{I_2}$ is the identity on $\mathbb{R}^3$ we see that $\mu_M \in SO(3)$ for all $M \in SU(2)$. As promised, we can now interpret (2.1) as saying that a Möbius transformation can be expressed as a rotation followed by a dilation followed by another rotation. In particular, $M \in PSL(2, \mathbb{C})$ represents a rotation if, and only if, the larger eigenvalue $\lambda$ of $MM^*$ is equal to 1.

Let us go back to our original problem and use the Möbius transformations in the following way: Let $u_\epsilon \in W^{1,2}(S^2, S^2)$ be a degree one critical point of $E_\epsilon$. The idea is to compose $u_\epsilon$ with a Möbius transformation $M$ and show that $(u_\epsilon)_M$ is close to the identity map $Id : S^2 \to S^2$ if $E(u_\epsilon) < 4\pi + 8\pi \epsilon + \mu$ and $\mu > 0$ is sufficiently small. If we are able to show that there exists $M \in PSL(2, \mathbb{C})$ such that $(u_\epsilon)_M$ is actually equal to $Id$, then $u_\epsilon$ itself has to be a Möbius transformation. Moreover, if $M \in SU(2)$ then $u_\epsilon$ is a rotation.

In a first step we investigate how $E_\epsilon$ transforms if we apply $u_M$. To do this we work in stereographic coordinates and consider $u : \hat{\mathbb{C}} \to S^2$. The Riemannian metric on $S^2$ in stereographic coordinates is given by $g_{ij} = \frac{4}{(1 + |\xi|^2)^2} \delta_{ij}$. For $\xi \in \hat{\mathbb{C}}$ we have

$$|\nabla_{S^2} u|^2(\xi) = \frac{(1 + |\xi|^2)^2}{4} |\nabla_{\mathbb{C}} u|^2(\xi) \quad \text{and} \quad |\Delta_{S^2} u|^2(\xi) = \frac{(1 + |\xi|^2)^4}{16} |\Delta_{\mathbb{C}} u|^2(\xi),$$

where $\nabla_\mathbb{C}$ is the gradient on $\mathbb{C}$ and $\Delta_{\mathbb{C}}$ the Laplacian on $\mathbb{C}$ with the flat metric on both the domain and the target. The area element is given by

$$dA_{S^2} = \frac{4}{(1 + |\xi|^2)^2} dA_{\mathbb{C}},$$

with $dA_{\mathbb{C}} = \sqrt{1 + d\xi^2}$ the Euclidean area element on $\mathbb{C}$. We define $u_M$ by

$$u_M(\xi) = u(M \xi) = u\left( \frac{a \xi + b}{c \xi + d} \right).$$

Using the above and the fact that $M \xi$ is harmonic we have

$$|\Delta_{S^2} u_M|^2(\xi) = \frac{(1 + |\xi|^2)^4}{16} |\Delta_{\mathbb{C}} u_M|^2(\xi)$$
Note that the transformation relation depends only on the eigenvalue \( \lambda \). We set
\[
\epsilon, \lambda = \begin{pmatrix} a & b \\ c & d \end{pmatrix},
\]
and therefore
\[
|E_{\epsilon, \lambda}|^2 = (1 + |\xi|^2)^4 |E_{\lambda}|^2(M \xi),
\]
and therefore
\[
|E_{\epsilon, \lambda}|^2 = \frac{\lambda^2(1 + |\xi|^2)^4}{(1 + \lambda^2|\xi|^2)^2} |E_{\lambda}|^2(M \xi).
\]

Analogously we get
\[
|\nabla E_{\epsilon, \lambda}|^2 = \frac{\lambda^2(1 + |\xi|^2)^2}{(1 + \lambda^2|\xi|^2)^2} |\nabla E_{\lambda}|^2(M \xi).
\]

Note that the transformation relation depends only on the eigenvalue \( \lambda \). Hence it is enough to restrict our attention in the following to dilations \( m_{\lambda} \). To show that \( M \in SU(2) \) it suffices to show that \( \lambda = 1 \). We set
\[
u(m_{\lambda}(\xi)) = u(\lambda \xi) =: u_{\lambda}(\xi)
\]
and
\[
\chi_{\lambda}(\xi) = \frac{(1 + \lambda^2|\xi|^2)^2}{\lambda^2(1 + |\xi|^2)^2}.
\]

With (2.4) and (2.5) we have
\[
|\nabla E_{\epsilon, \lambda}|^2(\lambda \xi) = \chi_{\lambda}(\xi)|\nabla E_{\epsilon, \lambda}|^2(\xi)
\]
and
\[
|\Delta E_{\epsilon, \lambda}|^2(\lambda \xi) = \chi_{\lambda}(\xi)|\Delta E_{\epsilon, \lambda}|^2(\xi).
\]

Applying all of this to \( E_{\epsilon} \), we get
\[
E_{\epsilon}(u) = \frac{1}{2} \int_{\mathbb{C}} \left( |\nabla E_{\epsilon, \lambda}|^2(\xi) + \epsilon |\Delta E_{\epsilon, \lambda}|^2(\xi) \right) \frac{4}{(1 + |\xi|^2)^2} dA_{\mathbb{C}}(\xi)
\]
\[
= \frac{1}{2} \int_{\mathbb{C}} \left( |\nabla E_{\epsilon, \lambda}|^2(\lambda \xi) + \epsilon |\Delta E_{\epsilon, \lambda}|^2(\lambda \xi) \right) \frac{4\lambda^2}{(1 + |\lambda \xi|^2)^2} dA_{\mathbb{C}}(\xi)
\]
\[
= \frac{1}{2} \int_{\mathbb{C}} \left( \chi_{\lambda}(\xi)|\nabla E_{\epsilon, \lambda}|^2(\xi) + \epsilon \chi_{\lambda}(\xi)|\Delta E_{\epsilon, \lambda}|^2(\xi) \right) \frac{4\lambda^2}{(1 + |\lambda \xi|^2)^2} dA_{\mathbb{C}}(\xi)
\]
\[
= \frac{1}{2} \int_{\mathbb{C}} \left( |\nabla E_{\epsilon, \lambda}|^2(\xi) + \epsilon \chi_{\lambda}(\xi)|\Delta E_{\epsilon, \lambda}|^2(\xi) \right) \frac{4}{(1 + |\xi|^2)^2} dA_{\mathbb{C}}(\xi)
\]
\[
= \frac{1}{2} \int_{\mathbb{C}} \left( |\nabla E_{\epsilon, \lambda}|^2 + \epsilon \chi_{\lambda}(\xi)|\Delta E_{\epsilon, \lambda}|^2 \right) dA_{\mathbb{C}}(\xi)
\]
\[
=: E_{\epsilon, \lambda}(u_{\lambda}).
\]

Hence \( u \) is a critical point of \( E_{\epsilon} \) if and only if \( u_{\lambda} \) is a critical point of \( E_{\epsilon, \lambda} \). Since \( E_{\epsilon, \lambda}(u_{\lambda}) = E_{\epsilon, \lambda-1}(u_{\lambda-1}) \) we will assume from now on that \( \lambda \geq 1 \). In the following we omit the subscript and write \( \nabla = \nabla_{S^2}, \Delta = \Delta_{S^2} \). An easy calculation (see [3] Proposition 1.1) shows
\[1\]

Note that \( E_{\epsilon, \lambda} \) differs from \( E_{\epsilon} \) only by the factor of \( \chi_{\lambda} \) that multiplies the \( |\Delta E_{\epsilon, \lambda}|^2 \) term. This factor \( \chi_{\lambda} \) is important because it measures the lack of conformal invariance of the integral \( \int_{S^2} |\Delta E_{\epsilon, \lambda}|^2 dA_{S^2} \).
Let \( \mu \) be a constant. We start with a Lemma collecting various properties of critical points of \( E \).

**Proof.**

We start by calculating as in section 5 of \( \tau \),
\[ -\Delta v + \varepsilon \Delta (\chi v) = \tau \left( \nabla v \right)^2 - \varepsilon \Delta (\chi v) = -\varepsilon \Delta (\chi v) + \nabla (\chi v) + \varepsilon \chi v. \]

\( \) \( \) \( \) \( \)

3. Closeness to the Möbius Group

In this section we consider critical points of \( E \) of degree 1 whose \( \varepsilon \)-energy lies below \( 4\pi(1+2\varepsilon) \) small. Our aim is to show that these maps are \( W^{1,2} \)-close to a Möbius transformation. We start with a Lemma collecting various properties of critical points of \( E \).

**Lemma 3.1.** Let \( v \in W^{2,2} \) be a critical point of \( E \), then we have

\[ 0 = \frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(v) = \varepsilon \int_{S^2} \chi \varepsilon |v|^2 dA_{S^2}, \]

where \( \varepsilon (\xi) = \frac{|\xi|^2 - 1}{|\xi|^2 + 1} \). Moreover, we have

\[ E_{\varepsilon, \lambda}(Id) = 4\pi \left( 1 + \frac{2\varepsilon}{3}(\lambda^2 + 1 + \lambda^{-2}) \right), \]

and there exists a constant \( C > 0 \) so that

\[ C\varepsilon (\lambda^2 - 1) \leq \frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(Id) - \frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(v) \leq \sqrt{\varepsilon} \left( \int \Delta v \right)_{L^2(S^2)} \sqrt{\int \lambda \varepsilon |v|} \quad \int \lambda \varepsilon |v| \quad \int \lambda \varepsilon |v| \quad \int \lambda \varepsilon |v|, \]

where \( \varepsilon (\xi) = \frac{|\xi|^2 - 1}{|\xi|^2 + 1} \). Since \( v \) is a critical point of \( E \) we have \( E_{\varepsilon, \lambda}(v) = 0 \). Using \( E_{\varepsilon, \tau}(v) = E_{\varepsilon, \lambda}(v_{\lambda^{-1}}) \) we get

\[ \frac{d}{d \log \tau} E_{\varepsilon, \lambda}(v) = \left( \frac{d}{d \tau} E_{\varepsilon, \lambda}(v_{\lambda^{-1}}) \right)_{\tau = \lambda} = E_{\varepsilon, \lambda}'(v) \left( \frac{d}{d \tau} v_{\lambda^{-1}} \right)_{\tau = \lambda} \]

and thus

\[ \frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(v) = 0, \]

which implies (3.1). Further

\[ E_{\varepsilon, \lambda}(Id) = 4\pi \varepsilon \int_{S^2} \chi \varepsilon dA_{S^2} \]

\[ = 4\pi \varepsilon \int_{\mathbb{C}} \frac{(1 + \lambda^2 |\xi|^2)^2}{\lambda^2 (1 + |\xi|^2)^2} \frac{dA_{C}(\xi)}{r} \]

\[ = 4\pi + 16\pi \varepsilon \int_{0}^{\infty} \frac{(1 + \lambda^2 r^2)^2}{r} \frac{dA_{C}(r)}{r} \]
where we used the substitution \( w = \frac{1 + \lambda^2 x^2}{\lambda (1 + \pi^2)} \). Differentiating this explicit expression for \( E_{\varepsilon, \lambda}(\text{Id}) \) with respect to \( \log \lambda \) yields
\[
\frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(\text{Id}) = \frac{16 \pi \varepsilon}{3} (\lambda^2 - \lambda^{-2}) = \frac{16 \pi \varepsilon}{3} (\lambda^2 - 1) - \frac{1}{\lambda^2}.
\]
\[
\geq \frac{16 \pi \varepsilon}{3} (\lambda^2 - 1).
\]
\[
(3.4)
\]
Since \( ||z||_{L^\infty(S^2)} \leq 1 \), we conclude that
\[
C\varepsilon (\lambda^2 - 1) \leq \frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(\text{Id}) - \frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(v)
\]
\[
= \varepsilon \int_{S^2} \chi \lambda z(\lambda) |(\Delta \text{Id})^2 - |\Delta v|^2| \, dA_{S^2}
\]
\[
\leq \sqrt{\varepsilon} \| \chi \lambda \Delta (v - \text{Id}) \|_{L^2(S^2)} \sqrt{\varepsilon} (\| \chi \lambda \Delta v \|_{L^2(S^2)} + \| \chi \lambda \Delta \text{Id} \|_{L^2(S^2)}).
\]
Inequality (3.3) is our main tool for measuring the deviation of \( \lambda \) from 1. We take the first step in this direction in the next proposition.

**Proposition 3.2.** For any \( \delta > 0 \) there exists \( \mu > 0 \) such that, if \( 0 < \varepsilon \leq 1 \) and if \( E_\varepsilon(u) \leq 4\pi (1 + 2 \varepsilon) + \mu \), where \( u \) is a critical point of \( E_\varepsilon \) of degree 1, then there exists \( M \in \text{PSL}(2, \mathbb{C}) \) such that
\[
\| \nabla (u_M - \text{Id}) \|_{L^2(S^2)} + \sqrt{\varepsilon} \| \chi \lambda \Delta (u_M - \text{Id}) \|_{L^2(S^2)} \leq \delta.
\]
Furthermore, there exists a fixed constant \( C > 0 \) such that if \( \lambda \geq 1 \) is the largest eigenvalue of \( MM^* \) (see (2.2)), then
\[
\varepsilon (\lambda^2 - 1) \leq C\delta.
\]
\[
(3.6)
\]
**Proof.** We prove (3.5) by contradiction using the energy identity in [9]. If (3.5) were not true, then we could find a sequence \( \mu_n \downarrow 0 \), a sequence \( \varepsilon_n \in (0, 1] \), a sequence \( u_n \in W^{2,2}(S^2, S^2) \) of critical points of \( E_{\varepsilon_n} \) of degree one, with \( E_{\varepsilon_n}(u_n) \leq 4\pi (1 + 2 \varepsilon_n) + \mu_n \) and \( \delta > 0 \) such that
\[
\| \nabla (u_n - \text{Id}) \|_{L^2(S^2)} + \sqrt{\varepsilon_n} \| \chi \lambda \Delta (u_n - \text{Id}) \|_{L^2(S^2)} > \delta
\]
for all \( M \in \text{PSL}(2, \mathbb{C}) \). Now we have to consider two cases:

\( \varepsilon_n \to 0 \cdot \)

There exists \( n_0 \in \mathbb{N} \) large enough such that \( \varepsilon_n < \frac{1}{4} \) and \( \mu_n < \frac{1}{2} \) for all \( n \geq n_0 \). Then \( E_{\varepsilon_n}(u_n) \) is uniformly bounded by \( 6\pi + \frac{1}{2} \) for all \( n \geq n_0 \). By Theorem 1.1 in [9] and Theorem 2 in [4], \( (u_n) \) converges to a harmonic map \( u^* \) and finitely many non-trivial harmonic two-spheres \( u^* : S^2 \to S^2 \) with \( \deg(u^*) + \sum_{i=1}^k \deg(u^i) = 1 \). With the result of Lemaire and Wood mentioned in the introduction, \( u^*, u^i \) are rational maps with energy \( E(u^*) = 4\pi |\deg(u^*)|, \ E(u^i) = 4\pi |\deg(u^i)| \). Since
\[
4\pi = \lim_{n \to \infty} E_{\varepsilon_n}(u_n) = E(u^*) + \sum_{i=1}^k E(u^i) = 4\pi (|\deg(u^*)| + \sum_{i=1}^k |\deg(u^i)|) \]
u^* is either a rational map with \( \text{deg}(u^*) = 1 \) and \( k = 0 \), or \( u^* \) is a constant map, \( k = 1 \) and \( u^1 : S^2 \to S^2 \) is a harmonic map of degree one. In the first case, \( u^* = m^* \) with some corresponding \( M^* \in PSL(2, \mathbb{C}) \) which is a contradiction to (3.7).

In the second case, energy concentrates in a small neighborhood of some \( x_0 \in S^2 \) and \( u_n \) converges to a constant map away from \( x_0 \). Without loss of generality let \( x_0 \) be the south pole \( S \). Let \( \sigma_n \downarrow 0 \) and \( D_n \) be a sequence of small disks around \( S \) such that the energy on \( S^2 \setminus D_n \) is smaller than \( \sigma_n \). We project \( D_n \) onto the complex plane. Then \( \Pi(D_n) = B_{r_n}(0) \), the complex ball with radius \( r_n \) and \( r_n \to 0 \). We perform a blow-up as in [9] and define

\[
v_n : \hat{C} \to S^2, \quad v_n(\xi) = u_n \circ \Pi^{-1} \left( \frac{\xi}{r_n} \right).\]

Note that this rescaling corresponds to dilations \( m_{\lambda_n} \) on the sphere with \( \lambda_n = \frac{1}{r_n} \) and \( D_n \) gets mapped to the lower hemisphere. \( v_n \) is a critical point of \( E_{\varepsilon_n} \) with \( \varepsilon_n = \frac{\varepsilon}{r_n} \). By Lemma 3.1 in [9] we have

\[
v_n \to v^* \quad \text{in } C^m_{\text{loc}}(C, S^2) \quad \forall m \in \mathbb{N},\]

where \( v^* : C \to S^2 \) is a non-trivial harmonic map. With the point removability result of Sacks and Uhlenbeck [14] we can lift \( v^* \) to a harmonic map from \( S^2 \) to \( S^2 \) with corresponding \( M^* \in PSL(2, \mathbb{C}) \) such that

\[
0 \leftarrow ||\nabla(v_n - v^*)||_{L^2(S^2)} + \sqrt{\varepsilon_n} \lambda_n ||\Delta(v_n - v^*)||_{L^2(S^2)} \\
\geq ||\nabla(v_n - v^*)||_{L^2(S^2)} + \sqrt{\varepsilon_n} ||\lambda_n \Delta(v_n - v^*)||_{L^2(S^2)} \\
= ||\nabla(u_n - (v^*)_{M_{\lambda_n}^{-1}})||_{L^2(S^2)} + \sqrt{\varepsilon_n} ||\lambda_n \Delta(u_n - (v^*)_{M_{\lambda_n}^{-1}})||_{L^2(S^2)} \\
\geq ||\nabla((u_n)_{M_{\lambda_n}(M^*)^{-1}} - \text{Id})||_{L^2(S^2)} + \sqrt{\varepsilon_n} ||\lambda_n (M_{\lambda_n}(M^*)^{-1} \Delta((u_n)_{M_{\lambda_n}(M^*)^{-1}} - \text{Id})||_{L^2(S^2)}.
\]

\( \varepsilon_n \to \varepsilon_{\infty} \in (0, 1) \):

Here we have, at least for \( n \) large enough, a uniform \( W^{2,2} \)-bound for the sequence \( u_n \). With the regularity results for \( \varepsilon \)-harmonic maps (see [7]), [9], [10]) we conclude that \( u_n \) converges strongly in \( W^{2,2} \) to a limiting map \( u_\infty \) which is a critical point of \( E_{\varepsilon_\infty} \) and which satisfies

\[
E_{\varepsilon_\infty}(u_\infty) = 4\pi(1 + 2\varepsilon_\infty).
\]

By (1.4) this implies that \( u_\infty \) is a rotation, contradicting (3.7).

To establish (3.6) we set \( v := u_M \) and use Lemma 3.1 to get

\[
C\varepsilon(\lambda^2 - 1) \leq \sqrt{\varepsilon} \sqrt{\lambda\Delta|u - \text{Id}|} \|_{L^2(S^2)} \sqrt{\varepsilon} (\|\sqrt{\lambda\Delta v}\|_{L^2(S^2)} + \|\sqrt{\lambda\Delta \text{Id}}\|_{L^2(S^2)}).
\]

By assumption,

\[
4\pi(1 + 2\varepsilon) + \mu \geq E_{\varepsilon}(u) = E_{\varepsilon, \lambda}(u_M) = E_{\varepsilon, \lambda}(v) \geq 4\pi + \frac{\varepsilon}{2} \int_{S^2} \lambda |\Delta v|^2 dA_{S^2},
\]

where we used

\[
\frac{1}{2} \int_{S^2} |\nabla v|^2 dA_{S^2} \geq 4\pi
\]

in the second inequality, which holds because \( \text{deg}(v) = 1 \). Thus

\[
\varepsilon \|\sqrt{\lambda\Delta v}\|_{L^2(S^2)}^2 \leq \sqrt{\varepsilon}(\|\sqrt{\lambda\Delta \text{Id} - v}\|_{L^2(S^2)} + \|\sqrt{\lambda\Delta v}\|_{L^2(S^2)}).
\]

By the triangle inequality

\[
\sqrt{\varepsilon} \|\sqrt{\lambda\Delta \text{Id}}\|_{L^2(S^2)} \leq \sqrt{\varepsilon}(\|\sqrt{\lambda\Delta (\text{Id} - v)}\|_{L^2(S^2)} + \|\sqrt{\lambda\Delta v}\|_{L^2(S^2)}).
\]
Proof.

We use (3.5), (3.8) and (3.9) in (3.3), we get
\[ \varepsilon(\lambda^2 - 1) \leq C\delta. \]

\[ \square \]

4. Improved bounds on \( \lambda \)

Next we want to improve the \( W^{1,2} \)-closeness result from the previous section and get a better bound on the eigenvalue \( \lambda \).

Proposition 4.1. Suppose \( v \) is a critical point of \( E_{\varepsilon, \lambda} \). Setting \( \psi := v - \text{Id} \) we have
\[ \Delta \psi + \psi|\nabla\psi|^2 + 2\psi(\nabla \psi, \nabla \text{Id}) + 2\psi + \text{Id}|\nabla\psi|^2 + 2\text{Id}(\nabla \psi, \nabla \text{Id}) = \varepsilon \sum_{j=1}^{3} \Psi_j(\psi, \text{Id}) \] (4.1)

with\[ \Psi_1(\psi, \text{Id}) = \chi_{\lambda} \left[ \Delta^2 \psi - 4\psi + (\psi + \text{Id}) \left( 4(\nabla \Delta \psi, \nabla \psi) + 4(\nabla \Delta \psi, \nabla \text{Id}) + |\Delta \psi|^2 + 2|\nabla^2 \psi|^2 \ight. \\
+ 4(\nabla^2 \psi, \nabla^2 \text{Id}) - 4(\Delta \psi, \text{Id}) - 8(\nabla \psi, \nabla \text{Id}) \right] \],
\[ \Psi_2(\psi, \text{Id}) = \nabla_i \chi_{\lambda} \left[ 2\nabla_i \Delta \psi - 4\nabla_i \text{Id} + (\psi + \text{Id}) \left( 4(\nabla_i \nabla \psi, \nabla \psi) + 2(\Delta \psi, \nabla_i \psi) + 4(\nabla_i \nabla \psi, \nabla \text{Id}) \ight. \\
+ 2(\Delta \psi, \nabla_i \text{Id}) + 4(\nabla \psi, \nabla_i \text{Id}) - 4(\nabla_i \psi, \text{Id}) \right] \],
\[ \Psi_3(\psi, \text{Id}) = \Delta \chi_{\lambda} \left[ \Delta \psi + 2\psi + (\psi + \text{Id}) \left( |\nabla \psi|^2 + 2(\nabla \psi, \nabla \text{Id}) \right) \right] . \]

Proof. We use (2.7) and replace \( v \) with \( \psi + \text{Id} \). Note that \( \Delta \text{Id} = -2 \text{Id} \), \( |\text{Id}|^2 = 1 \) and \( |\nabla \text{Id}|^2 = 2 \).

Then we have
\[ -\Delta \psi - \Delta \text{Id} - (\psi + \text{Id})|\nabla \psi + \nabla \text{Id}|^2 = -\varepsilon \Delta (\chi_{\lambda}(\Delta \psi + \Delta \text{Id})) \\
+ \varepsilon(\psi + \text{Id}) \chi_{\lambda}|\Delta \psi + \Delta \text{Id}|^2 \\
- \varepsilon(\psi + \text{Id}) \Delta (\chi_{\lambda}|\nabla \psi + \nabla \text{Id}|^2) \\
- 2\varepsilon(\psi + \text{Id}) \text{div}(\chi_{\lambda}(\Delta \psi + \Delta \text{Id}), \nabla \psi + \nabla \text{Id}) \]
\[ \Leftrightarrow \Delta \psi + \psi|\nabla \psi|^2 + 2\psi + 2\psi(\nabla \psi, \nabla \text{Id}) + \text{Id}|\nabla \psi|^2 + 2\text{Id}(\nabla \psi, \nabla \text{Id}) \\
= \varepsilon \Delta (\chi_{\lambda}(\Delta \psi + \Delta \text{Id})) \\
- \varepsilon(\psi + \text{Id}) \chi_{\lambda}|\Delta \psi + \Delta \text{Id}|^2 \\
+ \varepsilon(\psi + \text{Id}) \Delta (\chi_{\lambda}|\nabla \psi + \nabla \text{Id}|^2) \\
+ 2\varepsilon(\psi + \text{Id}) \text{div}(\chi_{\lambda}(\Delta \psi + \Delta \text{Id}), \nabla \psi + \nabla \text{Id}) \] (4.2)

The claim now follows from a direct computation (see [7], Proposition 4.3.1 for details). \( \square \)

Before we get to the next lemma note that we can estimate \( \chi_{\lambda} \) and its derivatives in terms of \( \lambda \). Note that we assumed \( \lambda \geq 1 \). It is easy to see that \( |\chi_{\lambda}| \leq \lambda^2 \). Additionally
\[ \partial_1 \chi_{\lambda}(\xi) = \frac{4\xi(1 + \lambda^2|\xi|^2)(\lambda^2 - 1)}{\lambda^4(1 + |\xi|^2)^3} \]
Lemma 4.2. To estimate the Laplacian of $\chi_0$ with $v^||S_2 = |\nabla_2 \chi_0| = \frac{2|\xi(1 + \lambda^2|\xi|^2)(\lambda^2 - 1)}{\lambda^2(1 + |\xi|^2)^4}$, we calculate:

$$\nabla_2 \chi_0 = 4|\xi|^2(1 + \lambda^2|\xi|^2)^2(\lambda^2 - 1)^2 \leq c(\lambda^2 - 1).$$

Further

$$\nabla_2 \chi_0^2 = \frac{2|\xi|^2(1 + \lambda^2|\xi|^2)^2(\lambda^2 - 1)^2}{\lambda^2(1 + |\xi|^2)^4} \leq c(\lambda^2 - 1).$$

To estimate the Laplacian of $\chi_\lambda$ we calculate

$$\sum_{i=1}^2 g_i^2 \chi_\lambda(\xi) = \frac{8(\lambda^2 - 1)(1 + 2\lambda^2|\xi|^2)}{\lambda^2(1 + |\xi|^2)^4} - \frac{24(\lambda^2 - 1)|\xi|^2(1 + \lambda^2|\xi|^2)}{\lambda^2(1 + |\xi|^2)^4} \leq c(\lambda^2 - 1).$$

With this we show

**Lemma 4.2.** There exist $0 < \varepsilon_0, \delta_0 < 1$ and a constant $C > 0$ depending only on $\varepsilon_0$ and $\delta_0$ such that for every $0 < \varepsilon < \varepsilon_0$, every $0 < \delta < \delta_0$ and every critical point $v \in W^{2,2}(S^2, S^2)$ of $E_{\varepsilon,\lambda}$ satisfying (3.5) and (3.6) we have

$$||\nabla_2 \chi_\lambda^2 \psi||_{L^2(S^2)} + \sqrt{\varepsilon}||\chi_\lambda \nabla^3 \psi||_{L^2(S^2)} \leq C(\delta + \varepsilon)\lambda,$$

with $\psi = v - \text{Id}$.

**Proof.** Note that $||\psi||_{L^\infty(S^2)} \leq 2$. We start by estimating the mean value of $\psi$ with (3.5), (4.2) and integration by parts. Note that $\int_{S^2} \Delta (\chi_\lambda(\Delta \psi + \Delta \text{Id})) = 0$.

$$\begin{align*}
2 \int_{S^2} \psi dA_{S^2} &= \left| - \int_{S^2} \left( \psi |\nabla \psi|^2 + 2 \psi (\nabla \psi, \nabla \text{Id}) + \text{Id} |\nabla \psi|^2 + 2 \text{Id}(\nabla \psi, \nabla \text{Id}) \right) dA_{S^2} \\
&+ \varepsilon \int_{S^2} \left( \Delta (\chi_\lambda(\Delta \psi + \Delta \text{Id})) - (\psi + \text{Id})\chi_\lambda |\Delta \psi + \Delta \text{Id}|^2 \\
&+ \Delta (\chi_\lambda |\nabla \psi + \nabla \text{Id}|^2) \\
&+ 2(\psi + \text{Id}) \text{div}(\chi_\lambda(\Delta \psi + \Delta \text{Id}), \nabla \psi + \nabla \text{Id}) \right) dA_{S^2} \right| \\
&\leq c \int_{S^2} |\nabla \psi|^2 dA_{S^2} + c \left( \int_{S^2} |\nabla \psi|^2 dA_{S^2} \right)^\frac{1}{2} \\
&+ \varepsilon \left| \int_{S^2} \left( - (\psi + \text{Id})\chi_\lambda |\Delta \psi + \Delta \text{Id}|^2 + (\psi + \text{Id})\Delta (\chi_\lambda |\nabla \psi + \nabla \text{Id}|^2) \\
+ 2(\psi + \text{Id}) \text{div}(\chi_\lambda(\Delta \psi + \Delta \text{Id}), \nabla \psi + \nabla \text{Id}) \right) dA_{S^2} \right| \\
&\leq c\delta \varepsilon + \varepsilon \left| \int_{S^2} \left( - (\psi + \text{Id})\chi_\lambda |\Delta \psi + \Delta \text{Id}|^2 + (\Delta \psi + \Delta \text{Id}) (\chi_\lambda |\nabla \psi + \nabla \text{Id}|^2) \\
- 2(\nabla \psi + \nabla \text{Id})(\chi_\lambda(\Delta \psi + \Delta \text{Id}), \nabla \psi + \nabla \text{Id}) \right) dA_{S^2} \right|.}
\end{align*}$$

(4.6)
We estimate the remaining terms using Young’s inequality, (3.5) and (3.6)
\[
\varepsilon \left| \int_{S^2} (\psi + \text{Id})\chi_\lambda |\Delta \psi + \Delta \text{Id}|^2 dA_{S^2} \right| \leq c\varepsilon \int_{S^2} \chi_\lambda (|\Delta \psi|^2 + 1) dA_{S^2} \\
\leq c(\delta^2 + \varepsilon) + \varepsilon\varepsilon(\lambda^2 - 1) \\
\leq c(\delta + \varepsilon)
\]
and
\[
\varepsilon \left| \int_{S^2} (\Delta \psi + \Delta \text{Id}) (\chi_\lambda |\nabla \psi + \nabla \text{Id}|^2) dA_{S^2} \right| \\
\leq c\varepsilon \int_{S^2} \chi_\lambda (|\nabla \psi| + 1) (|\nabla \psi|^2 + 1) dA_{S^2} \\
\leq c\varepsilon \int_{S^2} \chi_\lambda |\Delta \psi|^2 dA_{S^2} + c\varepsilon \int_{S^2} \chi_\lambda |\nabla \psi|^4 dA_{S^2} + c\varepsilon \int_{S^2} \chi_\lambda |\nabla \psi|^2 dA_{S^2} + c\varepsilon \lambda^2 \\
\leq c(\delta + \varepsilon) + c\varepsilon \int_{S^2} \chi_\lambda |\nabla \psi|^4 dA_{S^2}.
\]
(4.7)
For the last term we use the Sobolev embedding \(W^{1,1} \hookrightarrow L^2(S^2)\), (3.5) and (4.4)
\[
\int_{S^2} \chi_\lambda |\nabla \psi|^4 dA_{S^2} \leq c \left( \int_{S^2} \frac{\nabla \chi_\lambda}{\sqrt{\lambda}} |\nabla \psi|^2 dA_{S^2} \right)^2 + c \left( \int_{S^2} \chi_\lambda |\nabla^2 \psi|^2 dA_{S^2} \right) \left( \int_{S^2} |\nabla \psi|^2 dA_{S^2} \right) \\
+ \left( \int_{S^2} \sqrt{\lambda} |\nabla \psi|^2 dA_{S^2} \right)^2 \\
\leq c\delta^4 \lambda^2 + c\delta^2 \left( \int_{S^2} \chi_\lambda |\nabla^2 \psi|^2 dA_{S^2} \right).
\]
To get an estimate on the full second derivative we integrate by parts and exchange derivatives.
By Lemma 2.1.2 in [8] we have \(|\nabla \Delta \psi - \Delta \nabla \psi| \leq c(|\nabla \psi|^2 + |\nabla \psi|)\) and therefore
\[
\int_{S^2} \chi_\lambda |\nabla \psi|^2 dA_{S^2} \leq c \int_{S^2} |\nabla \chi_\lambda| |\nabla \psi|^2 |\nabla^2 \psi| dA_{S^2} + c \int_{S^2} \chi_\lambda |\Delta \psi|^2 dA_{S^2} \\
+ c \int_{S^2} \chi_\lambda (|\nabla \psi|^4 + |\nabla \psi|^2) dA_{S^2} \\
\leq (c\delta^2 + \eta) \int_{S^2} \chi_\lambda |\nabla^2 \psi|^2 dA_{S^2} + c\eta \int_{S^2} \frac{|\nabla \chi_\lambda|^2}{\chi_\lambda} |\nabla \psi|^2 dA_{S^2} \\
+ c \int_{S^2} \chi_\lambda |\Delta \psi|^2 dA_{S^2} + c\delta^2 \lambda^2.
\]
For \(\delta, \eta > 0\) small we absorb the first term to the left-hand side and with (4.4) we have
\[
\int_{S^2} \chi_\lambda |\nabla^2 \psi|^2 dA_{S^2} \leq c\delta^2 \lambda^2 + c \int_{S^2} \chi_\lambda |\Delta \psi|^2 dA_{S^2}
\]
(4.8)
and thus
\[
\int_{S^2} \chi_\lambda |\nabla \psi|^4 dA_{S^2} \leq c\delta^2 \lambda^2 + c\delta^2 \int_{S^2} \chi_\lambda |\Delta \psi|^2 dA_{S^2}.
\]
(4.9)
Going back to (4.7) and using the above estimates we get
\[
\varepsilon \left| \int_{S^2} (\Delta \psi + \Delta \text{Id}) (\chi_\lambda |\nabla \psi + \nabla \text{Id}|^2) dA_{S^2} \right| \leq c(\delta + \varepsilon).
\]
Analogously we estimate
\[ \varepsilon \left| \int_{S^2} -2(\nabla \psi + \nabla \operatorname{Id}) \langle \chi \lambda (\Delta \psi + \Delta \operatorname{Id}), \nabla \psi + \nabla \operatorname{Id} \rangle dA_{S^2} \right| \leq c(\delta + \varepsilon). \]
Combining all these estimates in (4.6) we obtain
\[ \left| \int_{S^2} \psi dA_{S^2} \right| \leq c(\delta + \varepsilon). \tag{4.10} \]
Further we have with $W^{1,1} \hookrightarrow L^2(S^2)$, (4.4), (4.8), (4.9) and Proposition 3.2
\[ \varepsilon \int_{S^2} \chi_3^2 |\nabla \psi|^6 dA_{S^2} \leq c \varepsilon \left( \int_{S^2} \chi \lambda |\nabla \psi|^3 dA_{S^2} \right)^2 + c \varepsilon \left( \int_{S^2} \chi |\nabla \psi|^2 dA_{S^2} \right)^2 \]
and similarly
\[ \varepsilon^2 \int_{S^2} \chi_3^2 |\nabla \psi|^4 dA_{S^2} \leq c \left( \varepsilon \int_{S^2} \sqrt{\chi} \lambda |\nabla \psi|^2 dA_{S^2} \right)^2 + c \left( \varepsilon \int_{S^2} \chi^2 |\nabla \psi|^2 dA_{S^2} \right)^2 \tag{4.11} \]
With this we can estimate the $L^2$-norm of $\chi \lambda |\nabla \psi|$. As above we integrate by parts and exchange derivatives. By Lemma 2.1.2 in [8] we have $|\nabla^2 \Delta \psi - \Delta \nabla^2 \psi| \leq c(|\nabla^2 \psi| |\nabla \psi| + |\nabla^2 \psi| + |\nabla^3 \psi| + |\nabla \psi|)$. With Proposition 3.2 and the estimates above we get
\[ \varepsilon \int_{S^2} \chi_3^2 |\nabla \psi|^2 dA_{S^2} \leq c \varepsilon \int_{S^2} \chi \lambda |\nabla \chi \lambda |\nabla^2 \psi| |\nabla ^3 \psi| dA_{S^2} + c \varepsilon \int_{S^2} \chi_3^2 |\nabla \Delta \psi|^2 dA_{S^2} \]
and similarly
\[ \varepsilon^2 \int_{S^2} \chi_3^2 |\nabla \psi|^4 dA_{S^2} \leq c \left( \varepsilon \int_{S^2} \sqrt{\chi} \lambda |\nabla \psi|^2 dA_{S^2} \right)^2 + c \left( \varepsilon \int_{S^2} \chi^2 |\nabla \psi|^2 dA_{S^2} \right)^2 \]
With Proposition 3.2 and the estimates above we get
Together with (4.14) we have
\[\int_{S^2} \delta \lambda^2 \psi dS + \epsilon \int_{S^2} \lambda^2 |\nabla \Delta \psi|^2 dS \leq c \delta^2 \lambda^2 + (c \delta^2 + \eta) \int_{S^2} \lambda |\nabla^2 \psi|^2 dS + c \epsilon \int_{S^2} \lambda^2 |\nabla \psi|^2 dS + c \epsilon \int_{S^2} \lambda^2 |\nabla \Delta \psi|^2 dS. \]

Now we multiply (4.1) with \(\lambda \Delta \psi\) and integrate over \(S^2\). After rearranging we get
\[\int_{S^2} \langle (\Delta - \epsilon \lambda \Delta^2) \psi, \lambda \Delta \psi \rangle dA_{S^2} = \int_{S^2} \left( - |\nabla \psi|^2 - 2 \psi \langle \nabla \psi, \nabla \text{Id} \rangle - 2 \psi - 2 \text{Id} |\nabla \psi|^2 - 2 \text{Id} \langle \nabla \psi, \nabla \text{Id} \rangle \right) + \epsilon \left( \Psi_1 - \lambda \Delta^2 \psi + \Psi_2 + \Psi_3 \right) R_{\lambda} \Delta \psi \rangle dA_{S^2}. \]

We estimate the left-hand side further
\[\int_{S^2} \langle (\Delta - \epsilon \lambda \Delta^2) \psi, \lambda \Delta \psi \rangle dA_{S^2} = \int_{S^2} \lambda |\nabla^2 \psi|^2 dA_{S^2} + \epsilon \int_{S^2} \lambda^2 |\nabla \Delta \psi|^2 dA_{S^2} + 2 \epsilon \int_{S^2} \lambda \nabla \lambda \nabla \Delta \psi dA_{S^2} \geq \int_{S^2} \lambda |\nabla \psi|^2 dA_{S^2} + \frac{3}{4} \epsilon \int_{S^2} \lambda^2 |\nabla \Delta \psi|^2 dA_{S^2} - c \epsilon \int_{S^2} |\nabla \lambda|^2 |\Delta \psi|^2 dA_{S^2}. \]

Together with (3.5), (4.4), (4.8) and (4.12) we have
\[\int_{S^2} \lambda |\nabla^2 \psi|^2 dA_{S^2} + \epsilon \int_{S^2} \lambda^2 |\nabla \psi|^2 dA_{S^2} \leq c \int_{S^2} \langle (\Delta - \epsilon \lambda \Delta^2) \psi, \lambda \Delta \psi \rangle dA_{S^2} + c \epsilon \int_{S^2} \lambda |\nabla \lambda|^2 |\Delta \psi|^2 dA_{S^2} + c \epsilon \int_{S^2} \lambda^2 |\nabla \psi|^2 dA_{S^2} \leq c \int_{S^2} \langle (\Delta - \epsilon \lambda \Delta^2) \psi, \lambda \Delta \psi \rangle dA_{S^2} + c \epsilon \int_{S^2} \lambda^2 |\nabla \psi|^2 dA_{S^2} + c \epsilon \int_{S^2} \lambda^2 |\nabla \Delta \psi|^2 dA_{S^2}. \]

On the right-hand side of (4.13) we have
\[\int_{S^2} \langle (\Delta - \epsilon \lambda \Delta^2) \psi, \lambda \Delta \psi \rangle dA_{S^2} = I + II + III + IV. \]

We estimate each term separately. With Young’s inequality, (3.5) and (4.9)
\[I = \int_{S^2} \langle (\Delta - \epsilon \lambda \Delta^2) \psi, \lambda \Delta \psi \rangle dA_{S^2} \leq \eta \int_{S^2} \lambda |\Delta \psi|^2 dA_{S^2} + c \eta \int_{S^2} \lambda \left( |\nabla \psi|^4 + |\nabla \psi|^2 \right) dA_{S^2} - 2 \int_{S^2} \langle \psi, \lambda \Delta \psi \rangle dA_{S^2} \]
Similarly we get for the next term with \((\psi - \tilde{\psi}) + \tilde{\psi}\)
\[
\leq \eta \int_{S^2} |\nabla \psi|^2 dA_{S^2} + c_\eta \lambda^2 \int_{S^2} |\nabla \psi|^2 dA_{S^2}
\]
\[
+ \eta \int_{S^2} \chi \lambda |\Delta \psi|^2 dA_{S^2}
\]
\[
+ c\lambda \int_{S^2} \frac{|\nabla \psi|^2}{\lambda} dA_{S^2}
\]
\[
+ \frac{c}{2} \int_{S^2} \frac{|\nabla \psi|^2}{\lambda} dA_{S^2}
\]
\[
+ (\lambda^2 - 1) |\tilde{\psi}|^2 \left( \int_{S^2} |\nabla \psi|^2 dA_{S^2} \right)^{\frac{1}{2}}
\]
\[
\leq \eta \int_{S^2} \chi \lambda |\Delta \psi|^2 dA_{S^2} + c_\eta \lambda^2
\]
All in all we have
\[
I \leq \eta \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2} + c_\eta \delta (\delta + \epsilon) \lambda^2.
\]
To estimate the second term we use (3.5), (3.6), (4.8), (4.9), (4.11) and (4.12)
\[
II = \epsilon \int_{S^2} \chi \lambda \left[ -4\psi + (\psi + \text{Id}) \left( 4(\nabla \Delta \psi, \nabla \psi) + 4(\nabla \Delta \psi, \nabla \text{Id}) + |\Delta \psi|^2 + 2|\nabla \psi|^2 \right) + 4(\nabla \psi, \nabla \text{Id}) - 4\Delta \psi \text{Id} - 8(\nabla \psi, \nabla \text{Id}) \right] dA_{S^2}
\]
\[
\leq c \epsilon \int_{S^2} \chi \lambda \left[ \nabla \Delta \psi ||\nabla \psi| + |\nabla \psi||\Delta \psi| + |\nabla \psi||\nabla \psi| + |\nabla \psi||\Delta \psi| + |\Delta \psi| \right] dA_{S^2}
\]
\[
\leq \eta \epsilon \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2} + \eta \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2} + c_\eta \epsilon^2 \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2}
\]
\[
+ c_\eta \epsilon \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2} + c_\eta \epsilon^2 \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2}
\]
\[
\leq (\eta + c_\eta \delta) \epsilon \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2} + c_\eta \epsilon^2 \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2}
\]
Similarly we get for the next term with (3.5), (3.6), (4.4), (4.8), (4.11) and (4.12)
\[
III = \epsilon \int_{S^2} \chi \lambda \left[ \nabla \Delta \psi \right. \left. + 4\nabla \psi, \nabla \text{Id} + (\psi + \text{Id}) \left( 4(\nabla \psi, \nabla \psi) + 2|\Delta \psi, \nabla \psi) \right) \right] dA_{S^2}
\]
\[
\leq c \epsilon \int_{S^2} \chi \lambda |\nabla \Delta \psi| \left( |\nabla \Delta \psi| |\Delta \psi| + |\nabla \psi|^2 |\nabla \psi| + |\nabla \psi|^2 |\Delta \psi| + |\Delta \psi| |\nabla \psi| + |\Delta \psi| \right) dA_{S^2}
\]
\[
\leq \eta \epsilon \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2} + \eta \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2} + c_\eta \epsilon \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2}
\]
\[
+ c_\eta \epsilon \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2} + c_\eta \epsilon^2 \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2}
\]
\[
+ c_\eta \epsilon \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2} + c_\eta \epsilon^2 \int_{S^2} \chi \lambda |\nabla \psi|^2 dA_{S^2}
\]
we have \(4.15\) and finally we have with (4.5)

\[
IV = \varepsilon \int_{S^2} \left\langle \delta \chi \left[ \Delta \psi + 2 \psi + (\psi + \text{Id}) \left( |\nabla \psi|^2 + 2 \langle \nabla \psi, \nabla \text{Id} \rangle \right) \right] \cdot \chi \Delta \psi \right\rangle dA_{S^2}
\]

\[
\leq c\varepsilon \int_{S^2} \chi \Delta \chi \left( |\Delta \psi|^2 + |\Delta \psi| |\nabla \psi|^2 + |\Delta \psi| |\nabla \psi| + |\Delta \psi| \right) dA_{S^2}
\]

\[
\leq c\varepsilon \int_{S^2} \chi \Delta \chi |\Delta \psi|^2 dA_{S^2} + c\varepsilon \int_{S^2} \chi \Delta \chi |\nabla \psi|^4 dA_{S^2} + c\varepsilon \int_{S^2} \chi \Delta \chi |\nabla \psi|^2 dA_{S^2}
\]

\[
+ \eta \int_{S^2} \chi \Delta \psi |\Delta \psi|^2 dA_{S^2} + c\eta^2 \int_{S^2} \chi \Delta \chi |\nabla \psi|^2 dA_{S^2}
\]

\[
\leq \eta \int_{S^2} \chi \Delta \psi |\Delta \psi|^2 dA_{S^2} + c(\delta + \varepsilon)^2 \lambda^2.
\]

Now we put (4.14), (4.15) and the above estimates together. Choosing \(\delta\) and \(\eta\) small enough so that we can absorb these terms to the left-hand side we arrive at

\[
\int_{S^2} \chi |\nabla \psi|^2 dA_{S^2} + \varepsilon \int_{S^2} \chi^2 |\nabla \psi|^2 dA_{S^2} \leq c(\delta + \varepsilon)^2 \lambda^2.
\]

\[\square\]

**Corollary 4.3.** There exist \(\varepsilon_0 > 0\) and \(\delta_0 > 0\), possibly smaller than those in Lemma 4.2, such that for every \(0 < \varepsilon \leq \varepsilon_0, 0 < \delta \leq \delta_0\) and every critical point \(v \in W^{2,2}(S^2, S^2)\) of \(E_{\varepsilon, \lambda}\) satisfying (3.5) and (3.6), we have

\[
\lambda^2 - 1 \leq c(\delta + \varepsilon).
\]

Moreover, the following estimate holds

\[
||\psi||_{L^\infty(S^2)} + ||\psi||_{W^{2,2}(S^2)} + \sqrt{2} ||\nabla \psi||_{L^2(S^2)} \leq c(\delta + \varepsilon). \quad (4.16)
\]

**Proof.** With (3.3) and Lemma 4.2 we have

\[
C\varepsilon(\lambda^2 - 1) \leq \frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(\text{Id}) - \frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(v)
\]

\[
= \varepsilon \int_{S^2} \chi \lambda z(\lambda)(|\Delta \text{Id}|^2 - |\Delta v|^2) dA_{S^2}
\]

\[
\leq \sqrt{\lambda} \sqrt{\lambda \Delta v}(v - \text{Id}) ||\nabla \Delta v||_{L^2(S^2)} \sqrt{\lambda} \sqrt{\lambda \Delta v} ||\nabla \text{Id}||_{L^2(S^2)}
\]

\[
\leq c\varepsilon(\delta + \varepsilon) \lambda^2 = c\varepsilon(\delta + \varepsilon)(\lambda^2 - 1) + c\varepsilon(\delta + \varepsilon).
\]

For \(\varepsilon + \delta\) small enough

\[
\lambda^2 - 1 \leq c(\delta + \varepsilon)
\]

and \(\lambda^2 \leq 2\). But then

\[
\frac{1}{2} \leq \frac{1}{\lambda^2} \leq |\chi \lambda| \leq \lambda^2 \leq 2.
\]

With this and Lemma 4.2 we get

\[
\int_{S^2} |\nabla \psi|^2 dA_{S^2} + \varepsilon \int_{S^2} |\nabla^3 \psi|^2 dA_{S^2} \leq c(\delta + \varepsilon)^2.
\]

By the Sobolev embedding \(W^{2,2} \hookrightarrow L^\infty(S^2)\), the Poincaré inequality and (4.10) it follows that

\[
||v||_{L^\infty(S^2)} \leq c ||v||_{W^{2,2}(S^2)} \leq c(\delta + \varepsilon) + c ||v - \bar{v}||_{L^2(S^2)} + c||v||
\]
\[ |\chi \lambda - 1| \leq \begin{cases} \lambda^2 - 1, & \text{if } \chi \lambda \geq 1 \\ 1 - \frac{1}{\lambda^2}, & \text{if } \chi \lambda < 1 \end{cases} \leq \lambda^2 - 1 \leq c(\delta + \epsilon). \] (4.17)

**Remark 4.4.** Note that

\[ |\chi \lambda - 1| \leq \left\{ \begin{array}{ll} \lambda^2 - 1, & \text{if } \chi \lambda \geq 1 \\ 1 - \frac{1}{\lambda^2}, & \text{if } \chi \lambda < 1 \end{array} \right\} \leq \lambda^2 - 1 \leq c(\delta + \epsilon). \]

5. Optimal Möbius Transformation

This section follows in parts chapter 6 in [11]. For a better comprehension of the arguments we repeat some of the calculations here.

So far our results suggest that there exists \( M \in PSL(2, \mathbb{C}) \) such that \( u_M \) is close to the identity, however there is still some freedom in the choice of \( M \). To show that \( u_M = \text{Id} \) and \( \lambda \) is equal to one, we have to choose the optimal Möbius transformation \( M \) with corresponding eigenvalue \( \lambda \) which minimizes \( \|\nabla(u_M - \text{Id})\|_{L^2(S^2)} \). This can be done as in section 6 of [11].

From now on we choose the optimal \( M \in PSL(2, \mathbb{C}) \) that minimizes \( \|\nabla(u_M - \text{Id})\|_{L^2(S^2)} \). Let \( \psi := u_M \) satisfy the assumptions of Lemma 4.2 and Corollary 4.3. Our goal is to improve the bound in (4.16) to \( \sqrt{2}(\lambda^2 - 1) \). In (4.1), a problematic term to estimate is \( \text{Id}(\nabla \psi, \nabla \text{Id}) \), because it involves \( \nabla \psi \) of order one. To eliminate this term we exploit that it is an element of the normal space at the identity.

By (4.16) \( \psi \) converges pointwise to the identity map as \( \delta \) and \( \varepsilon \) tend to zero. Thus we can write \( v \) in terms of the tangential component of \( \psi \) at the identity

\[ v = \text{Id} + \psi = \text{exp}_{\text{Id}} \hat{\psi} \quad (= \text{Id} + \hat{\psi} + O(|\hat{\psi}|^2)), \quad \hat{\psi} \in T_{\text{Id}}W^{3,2}(S^2, S^2). \]

In the following we want to work with \( \hat{\psi} \) instead of \( \psi \). To do this we need formulas to express \( \hat{\psi} \) in terms of \( v \) and \( \psi \). Let \( x = (x, y, z) \in S^2 \subset \mathbb{R}^3 \), then

\[ v(x) = x\sqrt{1 - |\hat{\psi}(x)|^2} + \hat{\psi}(x), \quad \hat{\psi}(x) \cdot x = 0, \]

\[ \hat{\psi}(x) = \psi(x) + \frac{1}{2}|\psi(x)|^2 x, \quad \psi(x) = \hat{\psi}(x) - \left( 1 - \sqrt{1 - |\hat{\psi}(x)|^2} \right)x, \] (5.1)

With this we get for the error terms of higher order

\[ |\nabla \psi - \nabla \hat{\psi}| = O(|\hat{\psi}| |\nabla \hat{\psi}|) + O(|\hat{\psi}|^2) = O(|\psi| |\nabla \psi|) + O(|\psi|^2), \]

\[ |\nabla^2 \psi - \nabla^2 \hat{\psi}| = O(|\hat{\psi}| |\nabla^2 \hat{\psi}|) + O(|\hat{\psi}|^2) = O(|\psi| |\nabla^2 \psi|) + O(|\psi|^2), \]

\[ |\nabla^3 \psi - \nabla^3 \hat{\psi}| = O(|\hat{\psi}| |\nabla^3 \hat{\psi}|) + O(|\hat{\psi}|^2) = O(|\psi| |\nabla^3 \psi|) + O(|\psi|^2). \] (5.2)

Let \( x^T \) be the orthogonal projection of \( x \in S^2 \) onto the tangent space \( T_xS^2 \). The tangential component of (4.1) is given by

\[ \left[ \varepsilon \chi \lambda \Delta^2 \psi - \Delta \psi - 2\psi - 4\varepsilon \chi \lambda \psi \right]^T = \\
\left[ (\text{Id} + \psi) |\nabla \psi|^2 + 2(\text{Id} + \psi) (\nabla \psi, \nabla \text{Id}) \right.
\]

\[ - \varepsilon \left( \Psi_1(\psi, \text{Id}) - \chi \lambda (\Delta^2 \psi - 4\psi) + \Psi_2(\psi, \text{Id}) + \Psi_3(\psi, \text{Id}) \right) \]

\[ \Leftrightarrow -\varepsilon (\Delta (\Delta \psi)^T)^T + (\Delta \psi)^T + 2\psi + 4\varepsilon \psi \]

\[ = -2\hat{\psi}(\nabla \psi, \nabla \text{Id}) + O(|\hat{\psi}| |\nabla^2 \hat{\psi}|) + O(|\hat{\psi}|^2) + O(|\hat{\psi}| |\nabla \hat{\psi}|) + O(|\hat{\psi}|^2) \]
With this we get

\[
\begin{align*}
&\varepsilon \left( (\chi_\lambda - 1)(\Delta(\Delta \hat{\psi})^T)^T + \chi_\lambda (\Delta(\Delta (\psi - \hat{\psi}))^T)^T + \chi_\lambda (\Delta(\Delta \psi)^N)^T \right) \\
&\quad - \varepsilon \left[ \Psi_1 - \chi_\lambda \Delta^2 \psi + 4 \chi_\lambda \psi + \Psi_2 + \Psi_3 \right]^T.
\end{align*}
\]

We shall denote by \( J_\varepsilon \) the operator on the left hand side of (5.3) and note that it can be written as

\[ J_\varepsilon = (1 - \varepsilon ((\Delta)^T - 2)) ((\Delta)^T + 2). \]

Observe that \((1 - \varepsilon ((\Delta)^T - 2)) \) is a positive operator and therefore \( J_\varepsilon \) has the same kernel as \( J := ((\Delta)^T + 2) \).

Let \( \psi = \tilde{\psi}_0 + \hat{\psi}_1 \), where \( \tilde{\psi}_0 \in \ker J_\varepsilon \) and \( \hat{\psi}_1 \in \ker J_\varepsilon \) with respect to the inner product in \( L^2 \). \( \Delta^T \tilde{\psi}_0 = -2 \tilde{\psi}_0 \) since \( J_\varepsilon \) and \( J \) have the same kernel. Further note that \( J_\varepsilon \) is self-adjoint with respect to the inner product in \( L^2(S^2, TS^2) \) and

\[
\int_{S^2} (J_\varepsilon \psi_1, (\Delta \hat{\psi})^T) dA_{S^2} = -2 \int_{S^2} (J_\varepsilon \hat{\psi}_1, \tilde{\psi}_0) dA_{S^2} = -2 \int_{S^2} (\hat{\psi}_1, J_\varepsilon \tilde{\psi}_0) dA_{S^2} = 0.
\]

With this we get

\[
\int_{S^2} (J_\varepsilon \psi, (\Delta \hat{\psi})^T) dA_{S^2} = \int_{S^2} (J_\varepsilon \hat{\psi}_1, (\Delta \hat{\psi}_1)^T) dA_{S^2} = \varepsilon \int_{S^2} |\nabla (\Delta \hat{\psi}_1)^T|^2 dA_{S^2} + \int_{S^2} |(\Delta \hat{\psi}_1)^T|^2 dA_{S^2}
\]

\[
+ (2 + 4\varepsilon) \int_{S^2} (\hat{\psi}_1, (\Delta \hat{\psi}_1)^T) dA_{S^2}.
\]

We want to control the \( \sqrt{\varepsilon} W^{3,2} \)-norm of \( \hat{\psi} \) by the left-hand side of (5.4). The first two terms on the right-hand side are positive, which leaves us with the last term. To get a control on this term we decompose the vectorfield \( \hat{\psi}_1 \) into eigenvectorfields of \( (\Delta)^T \). In [11], these eigenvectorfields were obtained by comparing \( (\Delta)^T \) with the (rough) connection Laplacian \( \Delta_{TS^2} \) and with the Hodge Laplacian, but here we shall proceed more directly and derive them from gradients of the eigenfunctions of \( \Delta \) on \( L^2(S^2) \).

In the computations that follow, it will be convenient to denote the coordinates of \( x \in \mathbb{R}^3 \) by \( (x_1, x_2, x_3) \), the corresponding partial derivatives by \( \partial_1, \partial_2, \partial_3 \), the \( \mathbb{R}^3 \) gradient by \( \partial = (\partial_1, \partial_2, \partial_3) \) and the \( \mathbb{R}^3 \) laplacian \( \partial_1^2 + \partial_2^2 + \partial_3^2 \) by \( \partial^2 \). If \( f \) is the restriction to \( S^2 \) of \( F \in C^\infty(\mathbb{R}^3) \) and \( \nabla f \) is its \( S^2 \)-gradient then we have:

\[
\nabla f(x) = \partial F(x) - (x \cdot \partial F(x)) x
\]

where \( x \cdot \partial F(x) = \sum_{i=1}^3 x_i \partial_i F(x) \). The relation between the spherical Laplacian \( \Delta f \) and \( \partial^2 F \) is given by

\[
\Delta f(x) = \partial^2 F(x) - 2x \cdot \partial F(x) - \sum_{i,j=1}^3 x_i x_j \partial_i \partial_j F(x).
\]

Let \( F = P_k \), a homogeneous polynomial of degree \( k \) that is harmonic on \( \mathbb{R}^3 \) and let \( p_k \) be the restriction of \( P_k \) to \( S^2 \). We see from (5.6) that \( p_k \) is an eigenfunction of \( \Delta \) with eigenvalue equal to \(-k(k + 1)\). All eigenfunctions of \( \Delta \) on \( S^2 \) are of this form. We claim that

\[
(\Delta (\nabla p_k))^T = -k(k + 1) \nabla p_k.
\]

To prove this claim, observe from (5.5) that, as an \( \mathbb{R}^3 \)-valued function,

\[
\nabla p_k(x) = \partial P_k(x) - k(P_k(x)) x.
\]
The components $\partial_i P_k(x)$ of $\partial P_k(x)$ are harmonic homogeneous polynomials of degree $(k - 1)$ and therefore

$$\Delta(\partial P_k(x)) = -k(k - 1)\partial P_k(x).$$

In the next calculation we let $e_1, e_2$ be an orthonormal basis of $T_x S^2$ so that $D_{e_i} e_j(x) = 0$, where $D$ is the Levi-Civita covariant derivative on $\mathbb{R}$. We compute:

$$\Delta((P_k(x))\mathbf{x}) = (\Delta P_k(x))\mathbf{x} + 2 \sum_{i=1}^2 e_i(P_k(x))e_i(\mathbf{x}) + (P_k(x))(\Delta \mathbf{x})$$

$$= -k(k + 1)(P_k(x))\mathbf{x} + 2\nabla p_k(\mathbf{x}) - 2(P_k(x))\mathbf{x}$$

where we have used $\nabla P_k(x) = (\partial P_k(x))^T = \nabla p_k(x)$ and $\Delta(x) = -2x$. We can now complete the proof of (5.7):

$$(\Delta(\nabla p_k))^T = (\Delta(\partial P_k(x)))^T - k(\Delta((P_k(x))\mathbf{x}))^T$$

$$= -k(k - 1)\nabla p_k - 2k\nabla p_k$$

$$= -k(k + 1)\nabla p_k.$$

To proceed further with the spectral analysis of $(\Delta)^T$ we recall the Helmholtz-Hodge decomposition of vector fields on $S^2$. Let $\star$ denote the anticlockwise rotation by $90^\circ$ in $TS^2$. The curl operator $\nabla \wedge$ is then defined by

$$\nabla \wedge \xi := -\nabla \cdot (\mathbf{x} \xi) \quad \text{for all } C^1 \text{-vector fields } \xi \text{ on } S^2.$$

It is easy to check that $\nabla \wedge (\nabla f) = 0 \forall f \in C^2(S^2)$. Since $S^2$ is simply connected, if $\nabla \wedge \xi = 0$ then there exists $f: S^2 \to \mathbb{R}$ (unique up to a constant) such that $\xi = \nabla f$. Similarly, if $\nabla \cdot \xi = 0$ then there exists $g: S^2 \to \mathbb{R}$ (unique up to a constant) such that $\xi = \mathbf{x} \nabla g$. We claim that if $\nabla \wedge \xi = 0$ and $\nabla \cdot \xi = 0$ then $\xi = 0$. This is because, writing $\xi$ as $\nabla f$ we get that

$$0 = \int_{S^2} f \nabla \cdot \xi dA_{S^2} = \int_{S^2} f(\Delta f) dA_{S^2} = -\int_{S^2} |\nabla f|^2 dA_{S^2}.$$
and
\[ \Delta(a \cdot (\xi x)) = (\Delta \xi) \cdot (a \times x) + 2 \sum_{i=1}^{2} e_i(\xi) \cdot (a \times e_i) - 2\xi \cdot (a \times x). \]

We now fix \( x \in S^2 \) and let \( a \) run over the orthonormal basis \( e_1(x), e_2(x) \) oriented so that \( x \times e_1 = e_2 \) to get
\[ \Delta(\xi x) = (-\Delta \xi - 2\xi \cdot e_2) e_1 + ((\Delta \xi - 2\xi \cdot e_1 - 2e_1(\xi) \cdot x) e_2 + 2(e_1(\xi) \cdot e_2 - e_2(\xi) \cdot e_1) x. \]

By differentiating the relation \( \xi(x) \cdot x = 0 \) we see that \( e_1(\xi) \cdot x + \xi \cdot e_1 = 0 \) and \( e_2(\xi) \cdot x + \xi \cdot e_2 = 0 \). Therefore
\[ (\Delta(\xi x))^T = (\Delta \xi \cdot e_2) e_1 + (\Delta \xi \cdot e_1) e_2 = * (\Delta \xi)^T. \]

(5.9) follows immediately from (5.7) and (5.10).

Let \( \varphi_0, \varphi_1, \varphi_2, \ldots \) be a complete orthonormal set of eigenfunctions of \( \Delta \) on \( S^2 \). (We have remarked that these eigenfunctions are the restrictions of harmonic homogeneous polynomials on \( \mathbb{R}^3 \) to \( S^2 \).) Then, because of the decomposition (5.8), the collection \( \nabla \varphi_1, * \nabla \varphi_1, \nabla \varphi_2, * \nabla \varphi_2, \ldots \) is a complete orthogonal set of eigenvector fields of \( (\Delta)^T \). \( (\varphi_0 \equiv 0 \) and therefore \( \nabla \varphi_0 = 0 \).

Furthermore, by (5.7) and (5.9), the eigenvalues corresponding to \( \nabla \varphi_j \) and \( * \nabla \varphi_j \) are the same as the eigenvalue (which is of the form \( -k(k+1), \ k \in \mathbb{N} \)) of \( \varphi_j \). In particular, the lowest eigenvalue of \( -(\Delta)^T \) is 2 and it has multiplicity equal to 6. The corresponding eigenvector fields are \( \nabla x_1, \nabla x_2, \nabla x_3 \) (whose flows are dilatations) and \( * \nabla x_1, * \nabla x_2, * \nabla x_3 \) (whose flows are rotations).

We can now proceed with obtaining a lower bound for the right hand side of (5.4) by decomposing \( W^{3,2}(S^2, T S^2) \) as \( \bigoplus_{j=1}^{\infty} E_{\lambda_j} \), where \( E_{\lambda_j} \) is the eigenspace of \( (\Delta)^T \) corresponding to the eigenvalue \( \lambda_j = -j(j + 1), \ j \in \mathbb{N} \). This decomposition enables us to express \( \hat{\psi} \) in (5.4) as
\[ \hat{\psi} = \sum_{j=1}^{\infty} \hat{\psi}_{\lambda_j}, \quad \hat{\psi}_{\lambda_j} \in E_{\lambda_j}. \]

As already remarked, the first eigenvalue is \( \lambda_1 = -2 \) and the eigenvectorfield \( \hat{\psi}_{\lambda_1} \) lies in the kernel of \( J_y \). Therefore
\[ \hat{\psi}_1 = \sum_{j=2}^{\infty} \hat{\psi}_{\lambda_j}. \]

Note that \( \int_{S^2} \langle \hat{\psi}_{\lambda_j}, \hat{\psi}_{\lambda_j} \rangle = 0 \) if \( i \neq j \). Then
\[ 2 \int_{S^2} \langle \hat{\psi}_1, (\Delta \hat{\psi}_1)^T \rangle dA_{S^2} = \sum_{j=2}^{\infty} \int_{S^2} \left\langle \frac{2}{\lambda_j} (\Delta \hat{\psi}_{\lambda_j})^T, (\Delta \hat{\psi}_{\lambda_j})^T \right\rangle dA_{S^2}, \]
\[ = \sum_{j=2}^{\infty} \frac{2}{\lambda_j} \int_{S^2} |(\Delta \hat{\psi}_{\lambda_j})^T|^2 dA_{S^2}. \]

(5.11)

Analogously we have
\[ 4\varepsilon \int_{S^2} \langle \hat{\psi}_1, (\Delta \hat{\psi}_1)^T \rangle dA_{S^2} = -4\varepsilon \sum_{j=2}^{\infty} \int_{S^2} \langle \nabla \hat{\psi}_{\lambda_j}, \nabla \hat{\psi}_{\lambda_j} \rangle dA_{S^2}, \]
\[ = -4\varepsilon \sum_{j=2}^{\infty} \int_{S^2} \frac{1}{\lambda_j} \langle \nabla (\Delta \hat{\psi}_{\lambda_j})^T, \nabla (\Delta \hat{\psi}_{\lambda_j})^T \rangle dA_{S^2}. \]
\[ = -4\varepsilon \sum_{j=2}^{\infty} \int_{S^2} \frac{1}{\lambda_j^2} |\nabla (\Delta \hat{\psi}_{\lambda_j})|^2 \, dA_{S^2} \]

Inserting this in (5.4) yields
\[
\int_{S^2} \left< J_{x} \hat{\psi}, (\Delta \hat{\psi})^T \right> \, dA_{S^2} = \varepsilon \int_{S^2} |\nabla (\Delta \hat{\psi})^T|^2 \, dA_{S^2} + \int_{S^2} |(\Delta \hat{\psi})^T|^2 \, dA_{S^2} + \int_{S^2} |(\Delta \hat{\psi})^T|^2 \, dA_{S^2} \]
\[ + \sum_{j=2}^{\infty} \frac{2}{\lambda_j} \int_{S^2} |(\Delta \hat{\psi}_{\lambda_j})^T|^2 \, dA_{S^2} - 4\varepsilon \sum_{j=2}^{\infty} \int_{S^2} \frac{1}{\lambda_j^2} |\nabla (\Delta \hat{\psi}_{\lambda_j})|^2 \, dA_{S^2} \]
\[ = \sum_{j=2}^{\infty} \frac{\lambda_j + 2}{\lambda_j} \int_{S^2} |(\Delta \hat{\psi}_{\lambda_j})^T|^2 \, dA_{S^2} + \varepsilon \sum_{j=2}^{\infty} \int_{S^2} \frac{\lambda_j^2 - 4}{\lambda_j^2} |\nabla (\Delta \hat{\psi}_{\lambda_j})|^2 \, dA_{S^2} \]

Note that \( \lambda_2 = -6 > \lambda_3 > \lambda_4 \ldots \) Therefore
\[ \frac{\lambda_j^2 - 4}{\lambda_j} \geq \frac{8}{9} \quad \text{and} \quad \frac{\lambda_j + 2}{\lambda_j} \geq \frac{2}{3} \quad \forall j \geq 2 \]
and so,
\[
\int_{S^2} \left< J_{x} \hat{\psi}, (\Delta \hat{\psi})^T \right> \, dA_{S^2} \geq \varepsilon \frac{8}{9} \|\nabla (\Delta \hat{\psi})^T\|_{L^2(S^2)}^2 + \frac{2}{3} \|\nabla (\Delta \hat{\psi})^T\|_{L^2(S^2)}. \tag{5.12} \]

We want to bound the full second and third derivative of \( \hat{\psi}_1 \) in terms of \( (\Delta \hat{\psi}_1)^T \) and \( \nabla (\Delta \hat{\psi}_1)^T \). To do this we take a closer look at the normal part \( (\Delta \hat{\psi}_1)^N = (\Delta \hat{\psi}_1 \cdot x) \). Note that with \( e_1, e_2 \) the orthonormal basis for \( T_x \mathbb{S}^2 \) as before we have
\[
\Delta \hat{\psi}_1 \cdot x = \sum_{i=1}^{2} e_i(\hat{\psi}_1) \cdot x = \sum_{i=1}^{2} e_i \left( (\hat{\psi}_1) \cdot x - (\hat{\psi}_1) \cdot e_i \right) (x) \\
= \sum_{i=1}^{2} e_i \left( (\hat{\psi}_1) \cdot x \right) = -2 \sum_{i=1}^{2} (\hat{\psi}_1) \cdot e_i (x) = -2 \text{ div } \hat{\psi}_1, \tag{5.13} \]

where we used that \( \hat{\psi}_1 \) is tangential in the second line and \( (\hat{\psi}_1) \cdot e_i (e_i) = \hat{\psi}_1 \cdot D_i e_i, e_i = 0 \) in the last line. Then
\[
(\Delta \hat{\psi}_1)^T = \Delta \hat{\psi}_1 + 2 \text{ div } \hat{\psi}_1 \cdot x \quad \text{and} \quad \varepsilon \nabla (\Delta \hat{\psi}_1)^T = \varepsilon \nabla (\Delta \hat{\psi}_1) + 2\varepsilon \nabla (\text{div } \hat{\psi}_1 \cdot x) \]
and therefore
\[
\|\Delta \hat{\psi}_1\|_{L^2(S^2)} \leq \|\nabla \hat{\psi}_1\|_{L^2(S^2)} + c \|\nabla \hat{\psi}_1\|_{L^2(S^2)} + c \varepsilon \left( \|\nabla \hat{\psi}_1\|_{L^2(S^2)} + \|\nabla \hat{\psi}_1\|_{L^2(S^2)} \right). 
\]

Integrating by parts we get
\[
\int_{S^2} |\nabla \hat{\psi}_1|^2 \, dA_{S^2} = -\int_{S^2} \hat{\psi}_1 \Delta \hat{\psi}_1 \, dA_{S^2} = \sum_{j=2}^{\infty} \frac{1}{(\lambda_j)} \int_{S^2} |(\Delta \hat{\psi}_{\lambda_j})|^2 \, dA_{S^2} \leq \frac{1}{6} \|\nabla \hat{\psi}_1\|_{L^2(S^2)}^2 
\]
where we have used (5.11) and \( -\lambda_j \geq 6 \quad \forall j \geq 2 \). We also have
\[
\|\hat{\psi}_1\|_{L^2(S^2)}^2 = \sum_{j=2}^{\infty} \int_{S^2} \frac{1}{\lambda_j^2} \left( (\Delta \hat{\psi}_{\lambda_j})^T, (\Delta \hat{\psi}_{\lambda_j})^T \right) \, dA_{S^2} \leq \frac{1}{36} \|\nabla \hat{\psi}_1\|_{L^2(S^2)}^2. 
\]
Therefore, going back to (5.12) we see that
\[
\sqrt{\varepsilon} \| \nabla \Delta \hat{\psi}_1 \|_{L^2(S^2)} + \| \Delta \hat{\psi}_1 \|_{L^2(S^2)} + \| \nabla \hat{\psi}_1 \|_{L^2(S^2)} + \| \hat{\psi}_1 \|_{L^2(S^2)} \\
\leq c \left( \int_{S^2} \langle J_\varepsilon \hat{\psi}, (\Delta \hat{\psi})^T \rangle dA_{S^2} \right)^{\frac{1}{2}} + c \sqrt{\varepsilon} \| \nabla^2 \hat{\psi}_1 \|_{L^2(S^2)}.
\]

To get an estimate for the full second and third derivative we integrate by parts and exchange derivatives as in the proof of Lemma 4.2. All in all we arrive at
\[
\sqrt{\varepsilon} \| \nabla^3 \hat{\psi}_1 \|_{L^2(S^2)} + \| \hat{\psi}_1 \|_{W^{2,2}(S^2)} \leq c \left( \int_{S^2} \langle J_\varepsilon \hat{\psi}, (\Delta \hat{\psi})^T \rangle dA_{S^2} \right)^{\frac{1}{2}}.
\tag{5.14}
\]

We have seen that the kernel of \( J \), and therefore the kernel of \( J_\varepsilon \), is finite (in fact 6) dimensional and all norms on it are equivalent. We estimate
\[
\sqrt{\varepsilon} \| \nabla^3 \hat{\psi}_0 \|_{L^2(S^2)} + \| \hat{\psi}_0 \|_{W^{2,2}(S^2)} \leq c \| \hat{\psi}_0 \|_{L^2(S^2)}.
\]

Together with (5.14)
\[
\sqrt{\varepsilon} \| \nabla^3 \hat{\psi} \|_{L^2(S^2)} + \| \hat{\psi} \|_{W^{2,2}(S^2)} \leq \left( \int_{S^2} \langle J_\varepsilon \hat{\psi}, (\Delta \hat{\psi})^T \rangle dA_{S^2} \right)^{\frac{1}{2}} + \| \hat{\psi} \|_{L^2(S^2)}.
\tag{5.15}
\]

At this point we use the fact that \( M \in PSL(2, \mathbb{C}) \) has been selected so as to minimize \( \| \nabla (u_M - Id) \|_{L^2(S^2)} \). This implies that
\[
- \int_{S^2} \nabla v \cdot \nabla \xi dA_{S^2} + \int_{S^2} \nabla Id \cdot \nabla \xi dA_{S^2} = 0 \quad \forall \xi \in Z,
\tag{5.16}
\]

where \( Z \) is the tangent space of the Möbius group at the identity and \( v \) is related to \( \hat{\psi} \) by (5.1). But \( Z = \ker J = \ker J_\varepsilon \), and, as shown in §6 of [11], this makes it possible to use (5.16) to estimate \( \hat{\psi}_0 \) as follows
\[
\| \hat{\psi}_0 \|_{L^2(S^2)} \leq c \| \hat{\psi} \|_{L^\infty(S^2)} \leq c \| \hat{\psi} \|_{L^2(S^2)} \| \hat{\psi} \|_{W^{2,2}(S^2)}.
\]

Going back to (5.15) and choosing \( \delta + \varepsilon \) in Corollary 4.3 small enough gives
\[
\sqrt{\varepsilon} \| \nabla^3 \hat{\psi} \|_{L^2(S^2)} + \| \hat{\psi} \|_{W^{2,2}(S^2)} \leq \left( \int_{S^2} \langle J_\varepsilon \hat{\psi}, (\Delta \hat{\psi})^T \rangle dA_{S^2} \right)^{\frac{1}{2}}.
\tag{5.17}
\]

Now we estimate the right hand side further. By (5.3)
\[
\int_{S^2} \langle J_\varepsilon \hat{\psi}, (\Delta \hat{\psi})^T \rangle dA_{S^2}
\]
\[
= \int_{S^2} \left[ -2 \hat{\psi} (\nabla \hat{\psi}, \nabla Id) + O(\| \hat{\psi} \| \nabla^2 \hat{\psi} \|) + O(\| \nabla \hat{\psi} \|^2) + O(\| \hat{\psi} \| \nabla \hat{\psi} \|) + O(\| \hat{\psi} \|^2) \right. \\
\left. + \varepsilon \left( (\chi^2 - 1)(\Delta (\Delta \hat{\psi})^T) + \chi (\Delta (\Delta (\hat{\psi} - \hat{\psi}))^T + \chi \lambda (\Delta (\nabla \hat{\psi})^T)^T \lambda \right) \\
- \varepsilon \left[ \Psi_1 - \chi \lambda (\nabla \hat{\psi} + \Psi_2 + \Psi_3) \right] (\Delta (\Delta \hat{\psi})^T) dA_{S^2}. \right. \tag{5.18}
\]

We estimate each part separately. For the first part we use \( W^{1,1} \hookrightarrow L^2(S^2) \), Hölder’s inequality and Corollary 4.3
\[
\int_{S^2} \langle 2 \hat{\psi} (\nabla \hat{\psi}, \nabla Id) + O(\| \hat{\psi} \| \nabla^2 \hat{\psi} \|) + O(\| \nabla \hat{\psi} \|^2) + O(\| \hat{\psi} \| \nabla \hat{\psi} \|) + O(\| \hat{\psi} \|^2), (\Delta \hat{\psi})^T \rangle dA_{S^2}
\]
\[
\leq \int_{S^2} (| \nabla^2 \hat{\psi} | (\| \hat{\psi} \| \nabla^2 \hat{\psi} | + | \nabla \hat{\psi} |^2 + | \hat{\psi} |^2) dA_{S^2}
\]
\[
\varepsilon \int_{S^2} \left\langle (\chi_\lambda - 1)(\Delta (\Delta \hat{\psi})^T)^T + \chi_\lambda (\Delta (\Delta (\psi - \hat{\psi}))^T + \chi_\lambda (\Delta (\Delta (\psi - \hat{\psi}))^T, (\Delta \hat{\psi})^T \right) dA_{S^2}
\]
\[
= -\varepsilon \int_{S^2} (\chi_\lambda - 1)|\nabla (\Delta \hat{\psi})|^2 + \chi_\lambda \left\langle \nabla \left( (\Delta (\psi - \hat{\psi})^T, (\Delta \hat{\psi})^T \right) dA_{S^2}
\]
\[
- \varepsilon \int_{S^2} \chi_\lambda \left\langle \nabla (\nabla \psi, \nabla \psi), (\Delta \hat{\psi})^T \right) dA_{S^2}
\]
\[
- \varepsilon \int_{S^2} \chi_\lambda \left\langle \nabla (\Delta \hat{\psi})^T, (\Delta \hat{\psi})^T \right) + \nabla \chi_\lambda \left\langle \nabla \left( (\Delta (\psi - \hat{\psi})^T, (\Delta \hat{\psi})^T \right) dA_{S^2}
\]
\[
- \varepsilon \int_{S^2} \chi_\lambda \left\langle \nabla (\nabla \psi, \nabla \psi), (\Delta \hat{\psi})^T \right) dA_{S^2}
\]
\[
\leq \varepsilon((\lambda^2 - 1) + \eta + c_\eta(\delta + \varepsilon)) \int_{S^2} |\nabla^3 \hat{\psi}|^2 dA_{S^2}
\]
\[
+ (c_\eta + (\lambda^2 - 1))\varepsilon \int_{S^2} |\nabla^2 \hat{\psi}|^4 + |\nabla \hat{\psi}|^4 + |\hat{\psi}|^4 dA_{S^2} + c_\varepsilon(\lambda^2 - 1) \int_{S^2} |\nabla^2 \hat{\psi}|^2 dA_{S^2}
\]
\[
+ (c_\eta + (\lambda^2 - 1))\varepsilon \int_{S^2} |\nabla \psi|^6 + |\nabla^2 \psi|^2|\nabla \psi|^2 |\hat{\psi}|^2 dA_{S^2}
\]
\[
\leq \varepsilon(\eta + c_\eta(\delta + \varepsilon)) \int_{S^2} |\nabla^2 \hat{\psi}|^2 dA_{S^2} + c_\eta \varepsilon \|\hat{\psi}\|_{H^2,2(S^2)}^2.
\]

In the same way we estimate
\[
\varepsilon \int_{S^2} \left\langle (\Psi_1 - \chi_\lambda (\Delta^2 \psi + 4\chi_\lambda \psi)^T, (\Delta \hat{\psi})^T \right) dA_{S^2}
\]
\[
= \varepsilon \int_{S^2} \chi_\lambda \left[ 4(\nabla \Delta \psi, \nabla \psi) + 2|\nabla \psi|^2 + |\Delta \psi|^2 + 6(\nabla \Delta \psi, \nabla \psi) + 4(\nabla \Delta \psi, \nabla \psi, \nabla \psi, \nabla \psi) - 4(\Delta \psi, \nabla \psi, \nabla \psi, \nabla \psi) - 8(\nabla \psi, \nabla \psi, \nabla \psi) \right]^T, (\Delta \hat{\psi})^T \right) dA_{S^2}
\]
\[
\leq \varepsilon(\eta + c_\eta(\delta + \varepsilon)) \int_{S^2} |\nabla^4 \hat{\psi}|^2 dA_{S^2} + c_\eta \varepsilon \|\hat{\psi}\|_{H^2,2(S^2)}^2.
\]

For the third term we use (4.3), \(W^{1,1} \hookrightarrow L^2(S^2)\) and Corollary 4.3
\[
\varepsilon \int_{S^2} \left\langle (\Psi_2^T, (\Delta \hat{\psi})^T \right) dA_{S^2} = \varepsilon \int_{S^2} \left\langle \nabla \chi_\lambda (\nabla \psi + \nabla \psi, \nabla \psi, \nabla \psi, \nabla \psi + \nabla \psi, \nabla \psi, \nabla \psi, \nabla \psi) + 2(\Delta \psi, \nabla \psi) + 4(\nabla \psi, \nabla \psi, \nabla \psi, \nabla \psi) - 4(\nabla \psi, \nabla \psi, \nabla \psi) \right)
\]
and with (4.5)

\[ \varepsilon \int_{S^2} \left\langle \Psi_3^T, (\Delta 2 \hat{\psi})^T \right\rangle dA_{S^2} = \varepsilon \int_{S^2} \left\langle \Delta \chi \left( (1d + \psi) (|\nabla \psi|^2 + 2(\nabla \psi, \nabla 2 \psi)) + \Delta 2 \hat{\psi} + 2 \psi \right)^T, (\Delta 2 \hat{\psi})^T \right\rangle dA_{S^2} \leq c_\varepsilon \varepsilon^2 (\lambda^2 - 1)^2 \int_{S^2} (|\nabla 2 \hat{\psi}|^2 + |\nabla \psi|^4 + |\nabla \psi|^2 + |\psi|^2) dA_{S^2} + \eta \int_{S^2} |\nabla 2 \hat{\psi}|^2 dA_{S^2} \leq c_\eta (\delta + \varepsilon) \varepsilon (\lambda^2 - 1)^2 + \eta \int_{S^2} |\nabla 2 \hat{\psi}|^2 dA_{S^2}. \]

Thus we have in (5.18)

\[ \int_{S^2} \left\langle J 2 \hat{\psi}, (\Delta 2 \hat{\psi})^T \right\rangle dA_{S^2} \leq c(\eta + \varepsilon + \delta) \left( \varepsilon |\nabla 3 2 \hat{\psi}|_{L^2(S^2)}^2 + |\nabla 2 \hat{\psi}|_{W^2,2(S^2)}^2 \right) + c(\delta + \varepsilon) \varepsilon (\lambda^2 - 1)^2. \]

We apply this to (5.17) and choose \( \eta, \varepsilon, \delta > 0 \) small enough so that we can absorb the higher order terms to the left-hand side.

\[
\sqrt{\varepsilon} |\nabla 3 2 \hat{\psi}|_{L^2(S^2)} + |\nabla \psi|_{W^2,2(S^2)} \leq \left< \int_{S^2} \left< J 2 \hat{\psi}, (\Delta 2 \hat{\psi})^T \right> dA_{S^2} \right>^{\frac{1}{2}} \leq c(\delta + \varepsilon) \sqrt{\varepsilon} (\lambda^2 - 1). \tag{5.19}
\]

To get the same bound for \( \psi \) we estimate with (5.2), \( W^{1, 1} \rightarrow L^2(S^2) \) and Corollary 4.3

\[ \varepsilon \int_{S^2} |\nabla 3 \psi|^2 dA_{S^2} \leq \varepsilon \int_{S^2} |\nabla 3 \psi|^2 dA_{S^2} + c \int_{S^2} (|\psi|^2 |\nabla 3 \psi|^2 + |\nabla 2 \psi|^4 + |\nabla \psi|^4 + |\psi|^4) dA_{S^2} \leq \varepsilon \int_{S^2} |\nabla 3 \psi|^2 dA_{S^2} + c(\delta + \varepsilon)^2 \int_{S^2} (|\nabla 3 \psi|^2 + |\nabla 2 \psi|^2 + |\nabla \psi|^2 + |\psi|^2) dA_{S^2} \]

Analogously we get

\[ \int_{S^2} |\nabla 2 \hat{\psi}|^2 dA_{S^2} \leq \int_{S^2} |\nabla 2 \hat{\psi}|^2 dA_{S^2} + c \int_{S^2} (|\psi|^2 |\nabla 2 \psi|^2 + |\nabla \psi|^4 + |\psi|^4) dA_{S^2} \leq \int_{S^2} |\nabla 2 \hat{\psi}|^2 dA_{S^2} + c(\delta + \varepsilon)^2 \int_{S^2} (|\nabla 2 \psi|^2 + |\nabla \psi|^2 + |\psi|^2) dA_{S^2} \]

and

\[ \int_{S^2} |\nabla \psi|^2 dA_{S^2} \leq \int_{S^2} |\nabla \psi|^2 dA_{S^2} + c(\delta + \varepsilon)^2 \int_{S^2} (|\nabla \psi|^2 + |\psi|^2) dA_{S^2}, \]

\[ \int_{S^2} |\psi|^2 dA_{S^2} \leq \int_{S^2} |\psi|^2 dA_{S^2} + c(\delta + \varepsilon)^2 \int_{S^2} |\psi|^2 dA_{S^2}. \]
Putting everything together and absorbing terms yields
\[
\sqrt{\varepsilon} \| \nabla^3 \psi \|_{L^2(S^2)} + \| \psi \|_{W^{2,2}(S^2)} \leq \left( \int_{S^2} (J_{\varepsilon} \psi, (\Delta \psi)^T) dA_{S^2} \right)^{\frac{1}{2}} \leq c(\delta + \varepsilon) \frac{1}{2} \sqrt{\lambda^2 - 1}. \tag{5.20}
\]

With this we show the following

**Theorem 5.1.** There exist \( \delta > 0 \) and \( \varepsilon > 0 \) small such that the only critical points \( u_\varepsilon \) of \( E_\varepsilon \) of degree \( \pm 1 \) with \( E_\varepsilon(u_\varepsilon) \leq 4 \pi (1 + 2\varepsilon + \delta) \) and \( \varepsilon \leq \varepsilon \) are maps of the form \( u^R(x) = Rx, \ R \in O(3) \).

**Proof.** Since a map of degree \(-1\) only differs from a map of degree \(1\) by a reflection, we can assume without loss of generality that \( u_\varepsilon \) is a critical point of degree one. Let \( M \) be the Möbius transformation that minimizes \( \| (u_\varepsilon)_M - \text{Id} \|_{L^2(S^2)} \) and let \( v = (u_\varepsilon)_M \). We use (5.20) to estimate (3.3) further

\[
C\varepsilon(\lambda^2 - 1) \leq \frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(\text{Id}) - \frac{d}{d \log \lambda} E_{\varepsilon, \lambda}(v) \leq 2\sqrt{\varepsilon} \sqrt{\lambda} \Delta(v - \text{Id})_{L^2(S^2)} \lambda \varepsilon \left( (\| \sqrt{\lambda} \Delta v \|_{L^2(S^2)} + \| \sqrt{\lambda} \Delta \text{Id} \|_{L^2(S^2)}) \right) \leq c\varepsilon^{\frac{1}{2}} (\delta + \varepsilon)^{\frac{1}{2}} (\lambda^2 - 1).
\]

Choosing \( \varepsilon > 0 \) small enough yields \( \lambda = 1 \). By (5.20) \( \psi \) must vanish and therefore \( v = \text{Id} \) and the Möbius transformation \( M \) must be a rotation. Hence \( u \) is a rotation.

Now we prove the Main Theorem 1.2

**Proof of Theorem 1.2.** If the statement is not true, there exist sequences \( \varepsilon_k \searrow 0 \) and critical points \( u_{\varepsilon_k} \) with \( E_{\varepsilon_k}(u_{\varepsilon_k}) \leq 12 \pi - \mu \) and \( \deg(u_{\varepsilon_k}) = 1 \) but \( u_{\varepsilon_k} \) is not of the form \( u^R(x) = Rx, \ R \in SO(3) \).

With Theorem 1.1 in [9] and Theorem 2 in [4] we get (up to the choice of a subsequence)

\[
12\pi > \lim_{k \to \infty} E_{\varepsilon_k}(u_{\varepsilon_k}) = \sum_{k=1}^{m} E(\omega^i) = 4\pi \sum_{i=1}^{m} |\deg(\omega^i)| \quad \text{and} \quad \deg(u_{\varepsilon_k}) = \sum_{i=1}^{m} \deg(\omega^i) = 1,
\]

with \( \omega^i \) non-trivial harmonic maps. Therefore \( m = 1 \) and \( \deg(\omega^1) = 1 \). Since \( \omega^1 \) is harmonic, \( E(\omega^1) = 4\pi \). Thus for every \( \delta > 0 \) there exists a \( k \) large enough so that

\[
E_{\varepsilon_k}(u_{\varepsilon_k}) \leq 4\pi + \delta \leq 4\pi (1 + 2\varepsilon) + \delta
\]

and Theorem 5.1 implies that \( u_{\varepsilon_k} \) is a rotation which, is a contradiction to our assumption. The proof follows analogously for maps of degree \(-1\).

\[ \square \]

6. Gap Theorem for \( \varepsilon \)-harmonic maps of degree zero

Now we turn our attention to \( \varepsilon \)-harmonic maps of degree zero. Theorem 1.1 follows analogously to [12]. Before we get to the proof, we need a \( \varepsilon \)-version of the \( \alpha \)-harmonic gap theorem of Sacks and Uhlenbeck ([14] Theorem 3.3).

**Lemma 6.1.** There exists \( \delta, \varepsilon > 0 \) such that if \( u_\varepsilon \in W^{2,2}(S^2, S^2) \) is a critical point of \( E_\varepsilon \) and \( E_\varepsilon(u_\varepsilon) < \delta \), then \( E_\varepsilon(u_\varepsilon) = 0 \) and \( u_\varepsilon \) is a constant map.

**Proof.** Let \( u_\varepsilon \) be a critical point of \( E_\varepsilon \). We multiply the Euler Lagrange equation with \( \Delta u_\varepsilon \) and integrate by parts.

\[
\int_{S^2} |\Delta u_\varepsilon|^2 + \varepsilon |\nabla \Delta u_\varepsilon|^2 dA_{S^2} = \int_{S^2} (\Delta u_\varepsilon - \varepsilon \Delta^2 u_\varepsilon, \Delta u_\varepsilon) dA_{S^2}
\]
we obtain for

By the Sobolev embedding $W^{1,1}$.

Arguing as in the proof of Lemma 4.2 we obtain for $\delta, \varepsilon, \eta > 0$ small enough

Hence $E_\varepsilon(u_\varepsilon) = 0$ and $u_\varepsilon$ is constant.

**Proof of Theorem 1.1.** We assume there exists a sequence $(u_\varepsilon)_{\varepsilon \in \mathbb{N}}$ of non-constant critical points of $E_\varepsilon$ with $E_\varepsilon(u_\varepsilon) \leq 8\pi - \delta$. The energy identity for $\varepsilon$-harmonic maps yields

$$8\pi > \lim_{k \to \infty} E_{\varepsilon_k}(u_{\varepsilon_k}) = \sum_{i=1}^{N} E(u_i) = \sum_{i=1}^{N} 4\pi |\deg(u_i)|,$$
where \( u^i : S^2 \to S^2, \ i = 1, \ldots, N \) are harmonic maps, which are non-trivial for \( i \geq 2 \). With the results of Duzaar and Kuwert \cite{DK} we have

\[
0 = \deg(u_{\epsilon_k}) = \sum_{i=1}^{N} \deg(u^i).
\]

Thus \( N = 1 \) and \( u^1 \) is a constant harmonic map. Then \( \lim_{k \to \infty} E_{\epsilon_k}(u_{\epsilon_k}) = 0 \) and with Lemma 6.1 it follows that \( u_{\epsilon_k} = \text{const} \).

Next we construct explicit examples of \( \epsilon \)-harmonic maps of degree zero with \( E_{\epsilon}(u_{\epsilon}) \geq 8\pi \) which are not constant. This shows that the bound in Theorem 1.1 is optimal. We follow \cite{Ku} and start by defining a class of rotationally symmetric maps. Let \( n \in \mathbb{N} \) and

\[
[n\pi, (n+1)\pi] \times [0, 2\pi] \to \mathbb{R}^3
\]

be a parametrization of \( S^2 \). For \( n \) even, this parametrization is orientation preserving and for \( n \) odd orientation reversing. Further let \( f \in C([0, \pi], \mathbb{R}) \) with

\[
f(0) = 0 \quad \text{and} \quad f(\pi) = n\pi.
\]

Then we define \( u_f : S^2 \to S^2 \) by

\[
u_f : [0, \pi] \times [0, 2\pi] \to \mathbb{R}^3
\]

\[
(r, \theta) \mapsto (\sin r \cos \theta, \sin r \sin \theta, \cos f(r)).
\]

\( u_f \) is rotationally symmetric and wraps \( n \) times around \( S^2 \), reversing orientation after each round. Hence \( u_f \) has degree zero if \( n \) is even and degree one if \( n \) is odd.

Let \( n = 2 \),

\[
X = \{ f : [0, \pi] \to \mathbb{R} : u_f \in W^{2,2}(S^2, \mathbb{R}^3), \ f(0) = 0, \ f(\pi) = 2\pi \}
\]

and \( M^* = \inf_{f \in X} I(f) \), where

\[
I(f) := E_{\epsilon}(u_f).
\]

\( E_{\epsilon}(u_f) \) is invariant under rotations about the \( z \)-axis and reflections in planes containing the line \((0, 0, z)\). Thus, by the principle of symmetric criticality of Palais (see \cite{Pa} or Remark 11.4(a) in \cite{F}), \( f \) is a critical point of \( I \) if and only if \( u_f \) is a critical point of \( E_{\epsilon} \). We show that there exists \( f^* \in X \) with \( I(f^*) = M^* \). Let \( (f_j) \) be a sequence in \( X \) with corresponding sequence \( (u_{f_j}) \in W^{2,2}(S^2, \mathbb{R}^3) \) and \( I(f_j) \searrow M^* \) and \( (u_{f_j}) \) is bounded in \( W^{2,2}(S^2, \mathbb{R}^3) \) and hence contains a subsequence (again denoted \( u_{f_j} \)) with \( u_{f_j} \to u_{f^*} \) weakly in \( W^{2,2}(S^2, \mathbb{R}^3) \) and uniformly in \( C^0(S^2, \mathbb{R}^3) \), with \( f^* \in X \). By the lower semi-continuity of \( E_{\epsilon} \) with respect to weak convergence in \( W^{2,2}(S^2, \mathbb{R}^3) \) we have \( I(f^*) = E_{\epsilon}(u_{f^*}) = M^* \).

Now we want to express \( E_{\epsilon}(u_f) \) in terms of \( f \) and compute

\[
\frac{\partial u_f}{\partial r} = f'(r) \left( \cos(f(r)) \cos(\theta), \cos(f(r)) \sin(\theta), -\sin(f(r)) \right)
\]

\[
\frac{\partial u_f}{\partial \theta} = \left( -\sin(f(r)) \sin(\theta), \sin(f(r)) \cos(\theta), 0 \right).
\]
Thus we have
\[
\frac{1}{2} |\nabla u_f|^2 = \frac{1}{2} \left( (f')^2 + \frac{(\sin(f))^2}{(\sin(r))^2} \right) .
\]
Noting again that
\[|\Delta u_f|^2 = |(\Delta u_f)^T|^2 + |\nabla u_f|^4 \geq |\nabla u_f|^4 ,\]
we get a lower bound on \(E_\varepsilon(u_{f_\ast})\)
\[
E_\varepsilon(u_{f_\ast}) \geq \pi \int_0^\pi \left( (f')^2 + \frac{(\sin(f))^2}{(\sin(r))^2} \right) \sin r \, dr + \frac{\varepsilon}{2} \int_{S^2} |\nabla u_{f_\ast}|^4 dA_{S^2} \\
> 2\pi \int_0^\pi |f''(\sin f^\ast)|dr + \frac{\varepsilon}{2} \left( 2\pi \int_0^\pi |f''(\sin f^\ast)|dr \right)^2 .
\]
There exist \(r_1 \in (0, \pi)\) such that \(f''(r_1) = \pi\) and
\[
\int_0^\pi |f''(\sin f^\ast)|dr \geq \int_{r_1}^\pi f''(\sin f^\ast)dr - \int_{r_1}^\pi f''(\sin f^\ast)dr \\
= - \cos f'(r)|_{r_1}^\pi + \cos f'(r)|_{r_1}^\pi \\
= 4 .
\]
Hence
\[
E_\varepsilon(u_{f_\ast}) \geq 8\pi + 32\varepsilon\pi^2
\]
and \(u_{f_\ast}\) is a non-constant \(\varepsilon\)-harmonic map of degree zero. To complete the proof of Theorem 1.3 we show the following

**Proposition 6.2.** There exists a universal constant \(c > 0\), such that for any \(0 < \varepsilon < \frac{1}{4}\)
\[
E_\varepsilon(u_\varepsilon) < 8\pi + 3\varepsilon \frac{\pi^2}{4} ,
\]
where \(u_\varepsilon\) is the minimizer of \(E_\varepsilon\) among all \(u_f\).

**Proof.** Let \(\Lambda > 1\) and
\[
f(r) = \begin{cases} 2 \arctan(\Lambda \tan(r)), & 0 \leq r \leq \frac{\pi}{2} , \\
2 \arctan(\Lambda \tan(r)) + 2\pi, & \frac{\pi}{2} < r \leq \pi .
\end{cases}
\]
We consider the corresponding map \(u_f \in X\). As \(r\) increases from 0 to \(\frac{\pi}{2}\), \(f(r)\) increases from 0 to \(\pi\), which means that \(u_f\) maps the upper hemisphere to the full sphere with the equator being mapped to the south pole \((0, 0, -1)\). As \(r\) increases from \(\frac{\pi}{2}\) to \(\pi\), \(f(r)\) increases from \(\pi\) to \(2\pi\), which means that \(u_f\) maps the lower hemisphere to the full sphere but with opposite orientation. Thus \(u_f\) has degree zero. We want to estimate \(E_\varepsilon(u_f)\). An explicit calculation (which can be found in section 4.5 of [7]) gives
\[
\frac{1}{2} \int_{S^2} |\Delta u_f|^2 dA_{S^2} = 32\pi \int_0^1 \Lambda^{-6} \frac{t^2(1-t^2)}{(1-a^2t^2)^4} dt + 32\pi \int_0^1 \Lambda^{-4} \frac{(1+t^2)^2}{(1-a^2t^2)^4} dt,
\]
where \(a^2 = 1 - \Lambda^{-2}\). We now estimate each term separately. First we note that \(t \in [0, 1]\) and \(0 \leq a \leq 1\)
\[
\frac{t^2(1-t^2)}{(1-a^2t^2)^4} \leq \frac{(a^2 - a^2t^2)}{a^2(1-a^2t^2)^4} \leq \frac{1}{a^2(1-a^2t^2)^3} = \frac{1}{a^2(1+at)^3} \leq \frac{1}{a^2(1+at)}.
\]
Then
\[
32\pi \int_0^1 \Lambda^{-6} \frac{t^2(1-t^2)}{(1-a^2t^2)^4} dt \leq 32\pi \Lambda^{-6} \int_0^1 \frac{1}{a^2(1-at)^3} dt = 32\pi \Lambda^{-6} \left( \frac{1}{2a^3(1-a)^2} - \frac{1}{2a^3} \right) .
\]
\[
\leq 32\pi\Lambda^{-6} \frac{(1 + a)^2}{a^2(1 - a^2)^2} \leq 128\pi\Lambda^{-2} \frac{1}{a^2} = 128\pi \frac{\Lambda^{-2}}{1 - \Lambda^{-2}} = 128\pi \frac{1}{\Lambda^2 - 1},
\]
where we used that $\Lambda^{-2} = 1 - a^2$. Analogously we get for the second term
\[
\frac{(1 + t^2)^2}{(1 - a^2t^2)^2} \leq \frac{(1 + t)^4}{(1 + at)^4(1 - at)^4} = \frac{(a + at)^4}{a^4(1 + at)^4(1 - at)^4} \leq \frac{1}{a^4(1 - at)^4}
\]
and therefore
\[
32\pi \int_0^1 \Lambda^{-4} \frac{(1 + t^2)^2}{(1 - a^2t^2)^2} dt \leq 32\pi\Lambda^{-4} \int_0^1 \frac{1}{a^4(1 - at)^4} dt = 32\pi\Lambda^{-4} \left( \frac{1}{3a^5(1 - a)^3} - \frac{1}{3a^5} \right)
\leq 32\pi\Lambda^{-4} \frac{(1 + a)^3}{a^4(1 - a^2)^3} \leq 256\pi\Lambda^2 \frac{1}{a^4} = 256\pi \frac{\Lambda^2}{(1 - \Lambda^{-2})^2}
\]
\[
= 256\pi \frac{\Lambda^6}{(\Lambda^2 - 1)^2}.
\]
All in all we get
\[
\frac{1}{2} \int_{S^2} |\Delta u_f|^2 dA_{S^2} \leq 128\pi \frac{1}{\Lambda^2 - 1} + 256\pi \frac{\Lambda^6}{(\Lambda^2 - 1)^2}.
\]
Analogously we estimate the first part of $E_\varepsilon(u_f)$ (see [12] Proposition 3.2)
\[
\frac{1}{2} \int_{S^2} |\nabla u_f|^2 dA_{S^2} = 8\pi \int_{S^2} \Lambda^{-2} \frac{1 + \cos^2 r}{(\cos^2 r + \Lambda^2 \sin^2 r)^2} \sin r dr
\leq 8\pi \int_0^1 \Lambda^{-2} \frac{(1 + t^2)}{(1 - a^2t^2)^2} dt
\leq 8\pi \int_0^1 \frac{\Lambda^{-2}}{2a^2} \left( \frac{1}{(1 - at)^2} + \frac{1}{(1 + at)^2} \right) dt
= 8\pi\Lambda^{-2} \frac{1}{a^2(1 - a^2)} = 8\pi \frac{1}{a^2} = 8\pi \frac{\Lambda^2}{\Lambda^2 - 1}.
\]
Together with the above we get
\[
\frac{1}{2} \int_{S^2} \left( |\nabla u_f|^2 + \varepsilon|\Delta u_f|^2 \right) dA_{S^2} \leq 8\pi \frac{\Lambda^2}{\Lambda^2 - 1} + 128\pi \varepsilon \frac{1}{\Lambda^2 - 1} + 256\pi \varepsilon \frac{\Lambda^6}{(\Lambda^2 - 1)^2}.
\]
We choose $\Lambda^2 > 2$, then $\frac{\Lambda^2}{\Lambda^2 - 1} < 1 + 2\Lambda^{-2} < 2$ and thus
\[
E_\varepsilon(u_f) < 8\pi \left( 1 + \frac{2}{\Lambda^2} \right) + 128\pi \varepsilon + 1024\pi \varepsilon \Lambda^2.
\]
We set $\Lambda := \varepsilon^{-\frac{1}{2}}$. Note that for $\varepsilon \in (0, \frac{1}{4})$, $\Lambda^2 > 2$ still holds. Then we get
\[
E_\varepsilon(u_f) < 8\pi \left( 1 + 2\varepsilon^{\frac{1}{2}} \right) + 1152\pi \varepsilon^{\frac{1}{2}} = 8\pi + 1168\pi \varepsilon^{\frac{1}{2}}.
\]
Since $u_\varepsilon$ minimizes $E_\varepsilon$ among all maps in $X$, we have
\[
E_\varepsilon(u_\varepsilon) \leq E_\varepsilon(u_f) < 8\pi + c\varepsilon^{\frac{1}{2}},
\]
with $c = 1168\pi$. \qed
Proof of Theorem 1.3. Given $\delta > 0$ we choose $\varepsilon \in (0, \frac{1}{4})$ such that $\varepsilon < (\frac{2}{c})^2$, where $c$ is the constant in (6.3). With Proposition 6.2 and Theorem 1.1 we get

$$8\pi + 32\varepsilon \pi^2 \leq E_\varepsilon (u_\varepsilon) < 8\pi + c\varepsilon \frac{1}{2} < 8\pi + \delta.$$ 

□

References

[1] Antonio Ambrosetti and Andrea Malchiodi. Nonlinear analysis and semilinear elliptic problems, volume 104 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2007. doi:10.1017/CBO9780511618260.

[2] Jason Cantarella, Dennis DeTurck, and Herman Gluck. Vector calculus and the topology of domains in 3-space. Amer. Math. Monthly, 109(5):409–442, 2002. doi:10.2307/2695643.

[3] Sun-Yung A. Chang, Liihe Wang, and Paul C. Yang. A regularity theory of biharmonic maps. Comm. Pure Appl. Math., 52(6):1113–1137, 1999. doi:10.1002/(SICI)1097-0325(199906)52:6<1113::AID-CPA4>3.0.CO;2-7.

[4] Frank Duzaar and Ernst Kuwert. Minimization of conformally invariant energies in homotopy classes. Calc. Var. Partial Differential Equations, 6(4):285–313, 1998. doi:10.1007/s005260050092.

[5] J. Eells and L. Lemaire. A report on harmonic maps. Bull. London Math. Soc., 10(1):1–68, 1978. doi:10.1112/blms/10.1.1.

[6] Robert Hardt and Michael Wolf, editors. Nonlinear partial differential equations in differential geometry, volume 2 of IAS/Park City Mathematics Series. American Mathematical Society, Providence, RI; Institute for Advanced Study (IAS), Princeton, NJ, 1996. Lecture notes from the Summer School held in Park City, Utah, June 20–July 12, 1992. doi:10.1090/pcms/002.

[7] Jasmin Hörter. Polyharmonic maps: conservation laws and approximations. PhD thesis, Karlsruher Institut für Technologie (KIT), 2021. doi:10.5445/IR/1000136367.

[8] Tobias Lamm. Biharmonic maps. Dissertation Albert-Ludwig-Universität Freiburg, 2005. URL: https://freidok.uni-freiburg.de/data/1755.

[9] Tobias Lamm. Fourth order approximation of harmonic maps from surfaces. Calc. Var. Partial Differential Equations, 27(2):125–157, 2006. doi:10.1007/s00526-005-0001-1.

[10] Tobias Lamm. Energy identity for approximations of harmonic maps from surfaces. Trans. Amer. Math. Soc., 362(8):4077–4097, 2010. doi:10.1090/S0002-9947-10-04912-3.

[11] Tobias Lamm, Andrea Malchiodi, and Mario Micallef. Limits of $\alpha$-harmonic maps. J. Differential Geom., 116(2):321–348, 2020. doi:10.4310/jdg/160396814.

[12] Tobias Lamm, Andrea Malchiodi, and Mario Micallef. A gap theorem for $\alpha$-harmonic maps between two-spheres. Anal. PDE, 14(3):881–889, 2021. doi:10.2140/apde.2021.14.881.

[13] Richard S. Palais. The principle of symmetric criticality. Comm. Math. Phys., 69(1):19–30, 1979. URL: http://projecteuclid.org/euclid.cmp/1103905401.

[14] J. Sacks and K. Uhlenbeck. The existence of minimal immersions of 2-spheres. Ann. of Math. (2), 113(1):1–24, 1981. doi:10.2307/1971313.

(J. Höter) Department of Mathematics, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany

Email address: jasmin.hoerter@kit.edu

(T. Lamm) Department of Mathematics, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany

Email address: tobias.lamm@kit.edu

(M. Micallef) Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK

Email address: M.J.Micallef@warwick.ac.uk