Annals of Mathematics, **159** (2004), 465–534

Repulsion and quantization in almost-harmonic maps, and asymptotics of the harmonic map flow

By Peter Topping*

Abstract

We present an analysis of bounded-energy low-tension maps between 2-spheres. By deriving sharp estimates for the ratio of length scales on which bubbles of opposite orientation develop, we show that we can establish a 'quantization estimate' which constrains the energy of the map to lie near to a discrete energy spectrum. One application is to the asymptotics of the harmonic map flow; we find uniform exponential convergence in time, in the case under consideration.

Contents

*Partly supported by an EPSRC Advanced Research Fellowship.

1. Introduction

1.1. *Overview.* To a sufficiently regular map $u : S^2 \to S^2 \hookrightarrow \mathbb{R}^3$ we may assign an energy

$$E(u) = \frac{1}{2}\int_{S^2} |\nabla u|^2, \tag{1.1}$$

and a tension field

$$\mathcal{T}(u) = \Delta u + u|\nabla u|^2, \tag{1.2}$$

orthogonal to u, which is the negation of the L^2-gradient of the energy E at u. Critical points of the energy — i.e. maps u for which $\mathcal{T}(u) \equiv 0$ — are called 'harmonic maps.' In this situation, the harmonic maps are precisely the rational maps and their complex conjugates (see [2, (11.5)]). In particular, being conformal maps from a surface, their energy is precisely the area of their image, and thus

$$E(u) = 4\pi|\deg(u)| \in 4\pi\mathbb{Z},$$

for any harmonic u.

In this work, we shall study 'almost-harmonic' maps $u : S^2 \to S^2$ which are maps whose tension field is small in $L^2(S^2)$ rather than being identically zero. One may ask whether such a map u must be close to some harmonic map; the answer depends on the notion of closeness. Indeed, it is known that u will resemble a harmonic 'body' map $h : S^2 \to S^2$ with a finite number of harmonic bubbles attached. Therefore, since the L^2 norm is too weak to detect these bubbles, u will be close to h in L^2. In contrast, when we use the natural energy norm $W^{1,2}$, there are a limited number of situations in which bubbles may be 'glued' to h to create a new harmonic map. In particular, if h is nonconstant and holomorphic, and one or more of the bubbles is antiholomorphic, then u cannot be $W^{1,2}$-close to any harmonic map. Nevertheless, by exploiting the bubble tree structure of u, it is possible to show that $E(u)$ must be close to an integer multiple of 4π.

One of the goals of this paper is to control just how close $E(u)$ must be to $4\pi k$, for some $k \in \mathbb{Z}$, in terms of the tension. More precisely, we are able to establish a 'quantization' estimate of the form

$$|E(u) - 4\pi k| \leq C\|\mathcal{T}(u)\|^2_{L^2(S^2)},$$

neglecting some exceptional special cases. Aside from the intrinsic interest of such a nondegeneracy estimate, control of this form turns out to be the key to an understanding of the asymptotic properties of the harmonic map heat flow (L^2-gradient flow on E) of Eells and Sampson. Indeed, we establish uniform exponential convergence in time and uniqueness of the positions of bubbles, in the situation under consideration, extending our work in [15].

A further goal of this paper, which turns out to be a key ingredient in the development of the quantization estimate, is a sharp bound for the length scale λ of any bubbles which develop with opposite orientation to the body map, given by

$$\lambda \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u)\|^2_{L^2(S^2)}}\right],$$

which we establish using an analysis of the Hopf differential and theory of the Hardy-Littlewood maximal function. The estimate asserts a repulsive effect between holomorphic and antiholomorphic components of a bubble tree, and could never hold for components of like orientation. (Indeed in general, bubbling may occur within sequences of harmonic maps.) From here, we proceed with a careful analysis of energy decay along necks, inspired by recent work of Qing-Tian and others, and a programme of 'analytic surgery,' which enables us to quantize the energy on each component of some partition of a bubble tree.

Our heat flow results, and our attempt to control energy in terms of tension, have precedent in the seminal work of Leon Simon [11]. However, our analysis is mainly concerned with the fine structure of bubble trees, and the only prior work of this nature which could handle bubbling in any form is our previous work [15]. The foundations of bubbling in almost-harmonic maps, on which this work rests, have been laid over many years by Struwe, Qing, Tian and others as we describe below.

1.2. *Statement of results.*

1.2.1. *Almost-harmonic map results.* It will be easier to state the results of this section in terms of *sequences* of maps $u_n : S^2 \to S^2$ with uniformly bounded energy, and tension decreasing to zero in L^2.

The following result represents the current state of knowledge of the bubbling phenomenon in almost-harmonic maps, and includes results of Struwe [13], Qing [7], Ding-Tian [1], Wang [17] and Qing-Tian [8].

THEOREM 1.1. *Suppose that* $u_n : S^2 \to S^2 \hookrightarrow \mathbb{R}^3$ $(n \in \mathbb{N})$ *is a sequence of smooth maps which satisfy*

$$E(u_n) < M,$$

for some constant M, and all $n \in \mathbb{N}$, and

$$\mathcal{T}(u_n) \to 0$$

in $L^2(S^2)$ as $n \to \infty$.

Then we may pass to a subsequence in n, and find a harmonic map $u_\infty : S^2 \to S^2$, and a (minimal) set $\{x^1, \ldots, x^m\} \subset S^2$ (with $m \leq \frac{M}{4\pi}$) such that

(a) *$u_n \rightharpoonup u_\infty$ weakly in $W^{1,2}(S^2)$,*

(b) *$u_n \to u_\infty$ strongly in $W^{2,2}_{\text{loc}}(S^2 \backslash \{x^1, \ldots, x^m\})$.*

Moreover, for each x^j, if we precompose each u_n and u_∞ with an inverse stereographic projection sending $0 \in \mathbb{R}^2$ to $x^j \in S^2$ (and continue to denote these compositions by u_n and u_∞ respectively) then for $i \in \{1, \ldots, k\}$ (for some $k \leq \frac{M}{4\pi}$ depending on x^j) there exist sequences $a^i_n \to 0 \in \mathbb{R}^2$ and $\lambda^i_n \downarrow 0$ as $n \to \infty$, and nonconstant harmonic maps $\omega^i : S^2 \to S^2$ (which we precompose with the same inverse stereographic projection to view them also as maps $\mathbb{R}^2 \cup \{\infty\} \to S^2$) such that:

(i)
$$\frac{\lambda^i_n}{\lambda^j_n} + \frac{\lambda^j_n}{\lambda^i_n} + \frac{|a^i_n - a^j_n|^2}{\lambda^i_n \lambda^j_n} \to \infty,$$
as $n \to \infty$, for each unequal $i, j \in \{1, \ldots, k\}$.

(ii)
$$\lim_{\mu \downarrow 0} \lim_{n \to \infty} E(u_n, D_\mu) = \sum_{i=1}^{k} E(\omega^i).$$

(iii)
$$u_n(x) - \sum_{i=1}^{k} \left(\omega^i \left(\frac{x - a^i_n}{\lambda^i_n} \right) - \omega^i(\infty) \right) \to u_\infty(x),$$
as functions of x from D_μ to $S^2 \hookrightarrow \mathbb{R}^3$ (for sufficiently small $\mu > 0$) both in $W^{1,2}$ and L^∞.

(iv) *For each $i \in \{1, \ldots, k\}$ there exists a finite set of points $\mathcal{S} \subset \mathbb{R}^2$ (which may be empty, but could contain up to $k - 1$ points) with the property that*
$$u_n(a^i_n + \lambda^i_n x) \to \omega^i(x),$$
in $W^{2,2}_{\text{loc}}(\mathbb{R}^2 \backslash \mathcal{S})$ as $n \to \infty$.

We refer to the map $u_\infty : S^2 \to S^2$ as a 'body' map, and the maps $\omega^i : S^2 \to S^2$ as 'bubble' maps. The points $\{x^1, \ldots, x^m\}$ will be called 'bubble points.' Since each ω^i is a nonconstant harmonic map between 2-spheres, the energy of each must be at least 4π.

When we say above that $\{x^1, \ldots, x^m\}$ is a 'minimal' set, we mean that we cannot remove any one point x^j without (b) failing to hold.

We have used the notation D_μ to refer to the open disc of radius μ centred at the origin in the stereographic coordinate chart $\mathbb{R}^2$.

Let us now state our main result for almost-harmonic maps. As we mentioned in Section 1.1 (see also Lemma 2.6) any harmonic map between 2-spheres is either holomorphic or antiholomorphic, and in particular, we may assume, without loss of generality, that the body map u_∞ is holomorphic (by composing each map with a reflection).

THEOREM 1.2. *Suppose we have a sequence $u_n : S^2 \to S^2$ satisfying the hypotheses of Theorem* 1.1, *and that we pass to a subsequence and find a limit u_∞, bubble points $\{x^j\}$ and bubble data ω^i, λ_n^i, a_n^i at each bubble point — as we know we can from Theorem* 1.1.

Suppose that u_∞ is holomorphic, and that at each x^j (separately) either

- *each ω^i is holomorphic, or*
- *each ω^i is antiholomorphic and $|\nabla u_\infty| \neq 0$ at that x^j.*

Then there exist constants $C > 0$ and $k \in \mathbb{N} \cup \{0\}$ such that after passing to a subsequence, the energy is quantized according to

$$|E(u_n) - 4\pi k| \leq C \|\mathcal{T}(u_n)\|^2_{L^2(S^2)}, \tag{1.3}$$

and at each x^j where an antiholomorphic bubble is developing, the bubble concentration is controlled by

$$\lambda_n^i \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right], \tag{1.4}$$

for each bubble ω^i.

By virtue of the hypotheses above, we are able to talk of a 'holomorphic' or 'antiholomorphic' bubble point x^j depending on the orientation of the bubbles at that point.

Remark 1.3. In particular, in the case that u_n is a holomorphic u_∞ with antiholomorphic bubbles attached, in the limit of large n, this result bounds the area A of the set on which u_n may deviate from u_∞ substantially in 'energy' (i.e. in $W^{1,2}$) by

$$A \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right],$$

for some $C > 0$.

In the light of [15], it is mixtures of holomorphic and antiholomorphic components in a bubble tree which complicate the bubbling analysis. However, in this theorem it is precisely the mix of orientations which leads to the repulsion estimate (1.4), forcing the bubble to concentrate as the tension decays. This repulsion is then crucial during our bubble tree decomposition, as we seek to squeeze the energy into neighbourhoods of integer multiples of 4π, according to (1.3). We stress that it is impossible to establish a repulsion estimate for holomorphic bubbles developing on a holomorphic body. Indeed, working in stereographic complex coordinates on the domain and target, the homotheties

$$u_n(z) = nz$$

are *harmonic* for each n, but still undergo bubbling.

The theorem applies to bubble trees which do not have holomorphic *and* antiholomorphic bubbles developing at the same point. Note that our previous work [15] required the stronger hypothesis that all bubbles (even those developing at different points) shared a common orientation, which permitted an entirely global approach. The restriction that $|\nabla u_\infty| \neq 0$ at antiholomorphic bubble points ensures the repulsive effect described above.[1]

Note that the hypotheses on the bubble tree in Theorem 1.2 will certainly be satisfied if only one bubble develops at any one point, and at each bubble point we have $|\nabla u_\infty| \neq 0$. In particular, given a nonconstant body map, our theorem applies to a 'generic' bubble tree in which bubble points are chosen at random, since $|\nabla u_\infty| = 0$ is only possible at finitely many points for a nonconstant rational map u_∞.

Remark 1.4. We should say that it is indeed possible to have an antiholomorphic bubble developing on a holomorphic body map u_∞ at a point where $|\nabla u_\infty| \neq 0$. For example, working in stereographic complex coordinates on the domain (z) and target, we could take the sequence

$$u_n(z) = |z|^{\frac{1}{n}} z - \frac{n^{-n}}{\bar{z}}$$

as a prototype, which converges to the identity map whilst developing an antiholomorphic bubble. However, we record that our methods force any further

[1] *Note added in proof.* The hypothesis $|\nabla u_\infty| \neq 0$ has since been justified; in [16] we find that the nature of bubbles at points where u_∞ has zero energy density can be quite different, and both the quantization (1.3) and the repulsion (1.4) may fail.

restrictions on the tension such as $\mathcal{T}(u_n) \to 0$ in the Lorentz space $L^{2,1}$ (a space marginally smaller than L^2) to impose profound restrictions on the type of bubbling which may occur. In particular, an antiholomorphic bubble could only occur at a point where $|\nabla u_\infty| = 0$ on a holomorphic body map.

We do not claim that the constant C from Theorem 1.2 is universal. We are concerned only with its independence of n.

1.2.2. *Heat flow results.* As promised earlier, Theorem 1.2 may be applied to the problem of convergence of the harmonic map heat flow of Eells and Sampson [3]. We recall that this flow is L^2-gradient descent for the energy E, and is a solution $u : S^2 \times [0, \infty) \to S^2$ of the heat equation

$$\frac{\partial u}{\partial t} = \mathcal{T}(u(t)), \tag{1.5}$$

with prescribed initial map $u(0) = u_0$. Here we are using the shorthand notation $u(t) = u(\cdot, t)$. Clearly, (1.5) is a nonlinear parabolic equation, whose critical points are precisely the harmonic maps. For any flow u which is regular at time t, a simple calculation shows that

$$\frac{d}{dt} E(u(t)) = -\|\mathcal{T}(u(t))\|^2_{L^2(S^2)}. \tag{1.6}$$

The following existence theorem is due to Struwe [13] and holds for any compact Riemannian domain surface, and any compact Riemannian target manifold without boundary.

THEOREM 1.5. *Given an initial map $u_0 \in W^{1,2}(S^2, S^2)$, there exists a solution $u \in W^{1,2}_{\mathrm{loc}}(S^2 \times [0, \infty), S^2)$ of the heat equation* (1.5) *which is smooth in $S^2 \times (0, \infty)$ except possibly at finitely many points, and for which $E(u(t))$ is decreasing in t.*

We note that the energy $E(u(t))$ is a smoothly decaying function of time, except at singular times when it jumps to a lower value. At the singular points of the flow, bubbling occurs and the flow may jump homotopy class; see [13] or [14].

Throughout this paper, when we talk about a solution of the heat equation (1.5), we mean a solution of the form proved to exist in Theorem 1.5 — for some initial map u_0.

Remark 1.6. Integrating (1.6) over time yields

$$\int_0^\infty \|\mathcal{T}(u(t))\|^2_{L^2(S^2)} dt = E(u_0) - \lim_{t\to\infty} E(u(t)) < \infty.$$

Therefore, we can select a sequence of times $t_n \to \infty$ for which $\mathcal{T}(u(t_n)) \to 0$ in $L^2(S^2)$, and $E(u(t_n)) \leq E(u_0)$. From here, we can apply Theorem 1.1 to find bubbling at a subsequence of this particular sequence of times.

In particular, we find the convergence

(a) $u(t_n) \rightharpoonup u_\infty$ weakly in $W^{1,2}(S^2)$,

(b) $u(t_n) \to u_\infty$ strongly in $W^{2,2}_{\mathrm{loc}}(S^2\backslash\{x^1,\ldots,x^m\})$,

as $n \to \infty$, for some limiting harmonic map u_∞, and points $x^1,\ldots,x^m \in S^2$. Unfortunately, this tells us nothing about what happens for intermediate times $t \in (t_i, t_{i+1})$, and having passed to a subsequence, we have no control of how much time elapses between successive t_i. Our main heat flow result will address precisely this question; our goal is uniform convergence in time. Let us note that in the case of no 'infinite time blow-up' (i.e. the convergence in (b) above is strong in $W^{2,2}(S^2)$) the work of Leon Simon [11] may be applied to give the desired uniform convergence, and if all bubbles share a common orientation with the body map, then we solved the problem with a global approach in [15]. On the other hand, if we drop the constraint that the target manifold is S^2, we may construct examples of nonuniform flows for which $u(s_i) \to u'_\infty \neq u_\infty$ for some new sequence $s_i \to \infty$, or even for which the bubbling is entirely different at the new sequence s_i; see [14] and [15].

We should point out that many examples of finite time and infinite time blow-up are known to exist for flows between 2-spheres — see [14] for a survey — beginning with the works of Chang, Ding and Ye. In fact, singularities are forced to exist for topological reasons, since if there were none, then the flow would provide a deformation retract of the space of smooth maps $S^2 \to S^2$ of degree k onto the space of *rational* maps of degree k, which is known to be impossible. Indeed, we can think of the bubbling of the flow as measuring the discrepancy between the topology of these mapping spaces. Note that here we are implicitly using the uniform convergence (in time) of the flow in the absence of blow-up, in order to define the deformation retract. Indeed, if we hope to draw topological conclusions from the properties of the heat flow in general (for example in the spirit of [10]) then results of the form of our next theorem are essential.

We now state our main uniformity result for the harmonic map heat flow. We adopt notation from Theorem 1.1.

THEOREM 1.7. *Suppose $u : S^2 \times [0,\infty) \to S^2$ is a solution of* (1.5) *from Theorem* 1.5, *and let us define*

$$\overline{E} := \lim_{t\to\infty} E(u(t)) \in 4\pi\mathbb{Z}.$$

By Remark 1.6 *above, we know that we can find a sequence of times $t_n \to \infty$ such that $\mathcal{T}(u(t_n)) \to 0$ in $L^2(S^2)$ as $n \to \infty$. Therefore, the sequence $u(t_n)$ satisfies the hypotheses of Theorem* 1.1 *and a subsequence will undergo bubbling as described in that theorem. Let us suppose that this bubbling satisfies the* *hypotheses of Theorem* 1.2. *Then there exists a constant* C_0 *such that for* $t \geq 0$,

$$|E(u(t)) - \overline{E}| \leq C_0 \exp\left[-\frac{t}{C_0}\right]. \tag{1.7}$$

Moreover, for all $k \in \mathbb{N}$ *and* $\Omega \subset\subset S^2 \backslash \{x^1, \ldots, x^m\}$ *— i.e. any compact set not containing any bubble points — and any closed geodesic ball* $B \subset S^2$ *centred at a bubble point which contains no other bubble point, there exist a constant* C_1 *and a time* t_0 *such that*

(i) $\|u(t) - u_\infty\|_{L^2(S^2)} \leq C_1 |E(u(t)) - \overline{E}|^{\frac{1}{2}}$ *for* $t \geq 0$,

(ii) $\|u(t) - u_\infty\|_{C^k(\Omega)} \leq C_1 |E(u(t)) - \overline{E}|^{\frac{1}{4}}$ *for* $t > t_0$,

(iii) $|E(u(t), B) - \limsup_{s \to \infty} E(u(s), B)| \leq C_1 |E(u(t)) - \overline{E}|^{\frac{1}{4}}$ *for* $t \geq 0$.

In particular, the left-hand sides of (i) *to* (iii) *above decay to zero exponentially, and we have the uniform convergence*

(a) $u(t) \rightharpoonup u_\infty$ *weakly in* $W^{1,2}(S^2)$ *as* $t \to \infty$,

(b) $u(t) \to u_\infty$ *strongly in* $C^k_{\mathrm{loc}}(S^2 \backslash \{x^1, \ldots, x^m\})$ *as* $t \to \infty$.

The fact that $\overline{E}$ is an integer multiple of 4π will follow from Theorem 1.1 (see part (i) of Lemma 2.15) but may be considered as part of the theorem if desired. The constants C_i above may have various dependencies; we are concerned only with their independence of t. The time t_0 could be chosen to be any time beyond which there are no more finite time singularities in the flow u.

Given our discussion in Remark 1.4, if we improved the strategy of Remark 1.6 to obtain a sequence of times at which the convergence $\mathcal{T}(u(t_n)) \to 0$ extended to a topology slightly stronger than L^2, then we could deduce substantial restrictions on the bubbling configurations which are possible in the harmonic map flow at infinite time.

Note added in proof. By requiring the hypotheses of Theorem 1.2 in Theorem 1.7, we are restricting the type of bubbles allowed at points where $|\nabla u_\infty| = 0$. Without this restriction, we now know the flow's convergence to be *nonexponential* in general; see [16].

1.3. *Heuristics of the proof of Theorem* 1.2. This section will provide a rough guide to the proof of Theorem 1.2, in which we extract some key ideas at the expense of full generality and full accuracy. Where possible, we refer to the lemmata in Section 2 in which we pin down the details.

We begin with some definitions of ∂ and $\bar{\partial}$-energies which will serve us throughout this paper. We work in terms of local isothermal coordinates x and y on the domain, and calculate ∇ and Δ with respect to these, as if we were working on a portion of $\mathbb{R}^2$ (in contrast to (1.1) and (1.2)).

In this way, if we define an energy density

$$e(u) := \frac{1}{2\sigma^2}|\nabla u|^2,$$

where σ is the scaling factor which makes $\sigma^2 dx \wedge dy$ the volume form on the domain S^2, then

$$E(u) = \int_{S^2} e(u).$$

Similarly, we have the ∂-energy, and $\bar{\partial}$-energy defined by

$$E_\partial(u) := \int_{S^2} e_\partial(u) \qquad \text{and} \qquad E_{\bar{\partial}}(u) := \int_{S^2} e_{\bar{\partial}}(u),$$

respectively, where

$$e_\partial(u) := \frac{1}{4\sigma^2}|u_x - u \times u_y|^2 = \frac{1}{4\sigma^2}|u \times u_x + u_y|^2,$$

and

$$e_{\bar{\partial}}(u) := \frac{1}{4\sigma^2}|u_x + u \times u_y|^2 = \frac{1}{4\sigma^2}|u \times u_x - u_y|^2,$$

are the ∂ and $\bar{\partial}$-energy densities. Of course, $E_{\bar{\partial}}(u) = 0$ or $E_\partial(u) = 0$ are equivalent to u being holomorphic or antiholomorphic respectively.

These 'vector calculus' definitions are a little unconventional, but will simplify various calculations in the sequel, when we derive and apply integral formulae for e_∂. Note that "$u\times$" has the effect of rotating a tangent vector by a right-angle.

Typically, the coordinates x and y will arise as stereographic coordinates, and thus

$$\sigma(x, y) := \frac{2}{1 + x^2 + y^2}. \tag{1.8}$$

We will repeatedly use the fact that $\sigma \leq 2$, and that $\sigma \geq 1$ for $(x, y) \in D := D_1$ the unit disc.

We also have the local energies $E(u, \Omega)$, $E_\partial(u, \Omega)$ and $E_{\bar{\partial}}(u, \Omega)$ where the integral is performed over some subset $\Omega \subset S^2$ rather than the whole of S^2, or equivalently over some subset Ω of an isothermal coordinate patch.

Note that all these energies are conformally invariant since our domain is of dimension two, a crucial fact which we use implicitly throughout this work.

A short calculation reveals the fundamental formulae

$$E(u) = E_\partial(u) + E_{\bar{\partial}}(u), \tag{1.9}$$

and

$$4\pi \deg(u) = E_\partial(u) - E_{\bar\partial}(u). \tag{1.10}$$

In particular, we have $E_\partial(u) \leq E(u)$ and $E_{\bar\partial}(u) \leq E(u)$. The identity (1.10) arises since $e_\partial(u) - e_{\bar\partial}(u) = u.(u_x \times u_y)$ is the Jacobian of u.

We now proceed to sketch the proof of Theorem 1.2. In order to simplify the discussion, we assume that the limiting body map u_∞ is simply the identity map. In particular, this ensures that $|\nabla u_\infty| \neq 0$ everywhere. We also assume that all bubbles are antiholomorphic rather than holomorphic. In some sense this is the difficult case, in the light of [15]. Here, and throughout this work, C will denote a constant whose value is liable to change with every use. During later sections — but not here — we will occasionally have cause to keep careful track of the dependencies of C.

Step 1. Since u_∞ is the identity map, we have $e_\partial(u_\infty) \equiv 1$ throughout the domain. We might then reasonably expect that $e_\partial(u_n) \sim 1$ for large n, since u_n is 'close' to u_∞. The first step of the proof is to quantify this precisely. We find that

$$\mathrm{Area}_{S^2}\left\{e_\partial(u_n) < \tfrac{1}{2}\right\} \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right],$$

for sufficiently large n. In the proof of the general case, this will be a *local* estimate; see Lemma 2.16.

The proof of this step will involve deriving an integral expression for $e_\partial(u_n) - e_\partial(u_\infty)$ using a Cauchy-type formula (see Lemma 2.1). A careful analysis will then control most of the terms of this expression in L^∞, and we will be left with an inequality of the form

$$|e_\partial(u_n) - 1| \leq \frac{1}{4} + |\mathcal{T}| * \frac{C}{|z|}$$

for sufficiently large n (cf. (2.51) in the proof of Lemma 2.16). We are therefore reduced to estimating the area of the set on which the convolution term in this expression is greater than $\frac{1}{4}$. In fact, this term is almost controllable in L^∞. Certainly we can control it in any L^p space for $p < \infty$, and the control disintegrates sufficiently slowly as $p \to \infty$ that this term is exponentially integrable; this is where the exponential in our estimate arises.

Step 2. The next step is where we capture much of the global information we require in the proof; here we use the fact that the domain is S^2. Using no special properties of the map u_n (other than some basic regularity) we find that for any $\eta > 0$, we have the estimate

$$\mathrm{Area}_{S^2}\left\{e_\partial(u_n)e_{\bar\partial}(u_n) > \eta\right\} \leq \frac{C}{\eta}E(u_n)\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}.$$

This estimate — which we phrase in a slightly different, but equivalent form in part(a) of Lemma 2.5 — follows via an analysis of the Hopf differential $\varphi\, dz^2$ (see §2.1.4) which would like to become holomorphic as the tension $\mathcal{T}$ becomes small. Note that the square of the magnitude of the Hopf differential is a measure of the product $e_\partial(u_n)e_{\bar\partial}(u_n)$. We prove a pointwise estimate for $|\varphi|^2$ in terms of the Hardy-Littlewood maximal function of $\varphi_{\bar z}$ which constitutes a sharp extension of the fact that there are no nontrivial holomorphic quadratic differentials on S^2. The desired estimate then follows upon applying maximal function theory.

Step 3. Steps 1 and 2 are sufficient to control the length scale λ_n of any antiholomorphic bubble according to

$$\lambda_n \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right],$$

for sufficiently large n.

Roughly speaking, any antiholomorphic bubble must lie within the small set where $e_\partial(u_n)$ is small. If instead the bubble — which carries a nontrivial amount of $\bar\partial$-energy — overlapped significantly with a region where $e_\partial(u_n)$ was of order one, then the product $e_\partial(u_n)e_{\bar\partial}(u_n)$ would have to be larger than is permitted by Step 2. The borderline nature of this contradiction is what prevents us from phrasing Step 2 in terms of integral estimates for $e_\partial(u_n)e_{\bar\partial}(u_n)$.

Step 4. A combination of Step 3 and a neck analysis in the spirit of Parker [6], Qing-Tian [8] and Lin-Wang [5], allows us to isolate (for each n) a dyadic annulus $\Omega = D_{2r}\backslash D_r$ around each antiholomorphic bubble (or groups of them) with energy bound according to

$$E(u_n, \Omega) \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right].$$

In essence, we can enclose the bubbles within annuli $\Sigma = D_1\backslash D_{r^2}$, with — according to Step 3 — r extremely small. In Lemma 2.9 of Section 2.2, we see that by viewing Σ conformally as a very long cylinder (of length $-2\ln r$) we can force an 'angular' energy to decay exponentially as we move along the cylinder from each end. By the centre of the cylinder, the energy over a fixed length portion — which corresponds to the energy over Ω — must have decayed to become extremely small.

Step 5. Our Step 3 is not unique in combining Step 1 with a Hopf differential argument. Part (b) of Lemma 2.5 also uses the Hopf differential, this time to establish that for $q \in [1, 2)$, we have

$$\|(e_\partial(u_n)e_{\bar\partial}(u_n))^{\frac{1}{2}}\|_{L^q(S^2)} \leq C\|\mathcal{T}(u_n)\|_{L^2(S^2)}.$$

Since $e_{\partial}(u_n)$ is small only on a very small set — according to Step 1 — this estimate can be improved to

$$\|(e_{\bar{\partial}}(u_n))^{\frac{1}{2}}\|_{L^q(S^2)} \leq C\|\mathcal{T}(u_n)\|_{L^2(S^2)},$$

for sufficiently large n; see Lemma 2.18 of Section 2.5.1. After a bootstrapping process, this may be improved to an estimate

$$E_{\bar{\partial}}(u_n, \Omega) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)},$$

for any compact Ω which contains no antiholomorphic bubble points (see Lemma 2.58). Crucially, this estimate contains no boundary term; one might expect a term involving the $\bar{\partial}$-energy of u_n over a region around the boundary of Ω. Indeed, here, as in Step 2, we are injecting global information using the Hopf differential and the fact that the domain is S^2.

Step 6. Armed with the energy estimates on dyadic annuli surrounding clusters of antiholomorphic bubbles, from Step 4, we can now carry out a programme of surgery on the map u_n to isolate the body, and bubble clusters. (See §§2.5.3, 2.5.2 and 2.5.4.) For example, we can find a new smooth map $w_n^1 : S^2 \to S^2$, for each n, which agrees with u_n outside the dyadic annuli (i.e. on most of the domain S^2) but which is constant within the annuli, and which retains the energy estimates of u_n on the annuli themselves. By invoking Step 5, and developing local $\bar{\partial}$-energy estimates for u_n in the regions just outside the dyadic annuli, we find that

$$E_{\bar{\partial}}(w_n^1) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)},$$

for sufficiently large n, which coupled with (1.9) and (1.10) gives the partial quantization estimate

$$|E(w_n^1) - 4\pi \deg(w_n^1)| \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}.$$

A similar procedure carried out for the *interior* of each annulus, yields maps which isolate the bubble clusters (i.e. which are equal to u_n within the annulus but are constant outside) and which also have quantized energy. Finally, $E(u_n)$ is well approximated by the sum of the energies of all these isolated maps, each of which has quantized energy, and we conclude that

$$|E(u_n) - 4\pi k| \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)},$$

for some integer $k \geq 0$ and sufficiently large n.

2. Almost-harmonic maps — the proof of Theorem 1.2

The goal of this section is to understand the structure of maps whose tension field is small when measured in L^2, and prove the bubble concentration estimates and energy quantization estimates of Theorem 1.2.

Before we begin, we outline some conventions which will be adopted throughout this section. Since the domain is S^2 in our results, we may stereographically project about any point in the domain to obtain isothermal coordinates x and y. It is within such a stereographic coordinate chart that we shall normally meet the notation D_μ which represents the open disc in $\mathbb{R}^2$ centred at the origin and of radius $\mu \in (0,\infty)$. We also abbreviate $D := D_1$ for the unit disc, which corresponds to an open hemisphere under (inverse) stereographic projection. An extension of this is the notation $D_{b,\nu}$ which corresponds to a disc of radius $\nu \in (0,\infty)$ centred at $b \in \mathbb{R}^2$ (and thus $D_\mu = D_{0,\mu}$).

When these discs are within a stereographic coordinate chart, we use the same notation for the corresponding discs in S^2. By the conformality of stereographic projection, and the conformal invariance of the energy functionals, we can talk about energies over discs (or the whole chart $\mathbb{R}^2$) without caring whether we calculate with respect to the flat metric or the spherical metric. In contrast, when we talk about function spaces such as $L^p(D_\mu)$ over these discs, we are using the standard measure from $\mathbb{R}^2$ rather than S^2. Moreover, the gradient ∇ and Laplacian Δ on one of these discs, will be calculated with respect to the $\mathbb{R}^2$ metric.

Given these remarks, the tension field of a smooth map $u : D_\mu \to S^2 \hookrightarrow \mathbb{R}^3$ from a stereographic coordinate chart D_μ is given by

$$\mathcal{T} = \mathcal{T}(u) := \frac{1}{\sigma^2}\left(\Delta u + u|\nabla u|^2\right), \tag{2.1}$$

with σ defined as in (1.8). Note that with our conventions, we may write $\|\mathcal{T}\sigma\|_{L^2(\mathbb{R}^2)} = \|\mathcal{T}(u)\|_{L^2(S^2)}$.

By default, when we consider a map $u : D_\mu \to S^2$, we imagine it to be a map from a stereographic coordinate chart, and (2.1) will be assumed. However, in Section 2.2, we will consider $\mathcal{T}$ with respect to a metric other than $\sigma^2(dx^2+dy^2)$; see Remark 2.10. The definition (2.1) will be reasserted in Lemma 2.16 where u is not the only map under consideration.

2.1. *Basic technology.* In this section we develop a number of basic estimates for the ∂ and $\bar{\partial}$-energies, and for the Hopf differential, which we shall require throughout this work. Most of these estimates are original, or represent new variations on known results. However, the reader may reasonably opt to extract results from this section only when they are required.

2.1.1. *An integral representation for e_∂.* The following lemma is a real casting of Cauchy's integral formula.

LEMMA 2.1. *Suppose that $u : S^2 \cong \mathbb{R}^2 \cup \{\infty\} \to S^2 \hookrightarrow \mathbb{R}^3$ is smooth and recall the definition of $\mathcal{T}$ from* (2.1). *Then*

$$2\pi(u \times u_x + u_y)(0,0)$$
$$= \int_{\mathbb{R}^2} \left(\frac{-1}{x^2+y^2}(y\,\mathcal{T} + x\,u \times \mathcal{T})\sigma^2 + \frac{y}{x^2+y^2}\,u|u_x - u \times u_y|^2 \right) dx \wedge dy.$$

More generally, if $\varphi : \mathbb{R}^2 \to \mathbb{R}$ *is smooth with compact support, and* $(a,b) \in \mathbb{R}^2$, *then*

$$2\pi(u \times u_x + u_y)\varphi(a,b)$$
$$= -\int_{\mathbb{R}^2} \frac{1}{(x-a)^2+(y-b)^2}\left((y-b)\,\mathcal{T} + (x-a)\,u \times \mathcal{T}\right)\sigma^2\varphi\, dx \wedge dy$$
$$+ \int_{\mathbb{R}^2} \frac{(y-b)}{(x-a)^2+(y-b)^2}\,(u|u_x - u \times u_y|^2)\varphi\, dx \wedge dy$$
$$- \int_{\mathbb{R}^2} \frac{1}{(x-a)^2+(y-b)^2}\Big(((x-a)\varphi_x + (y-b)\varphi_y)(u \times u_x + u_y)$$
$$-((x-a)\varphi_y - (y-b)\varphi_x)(u_x - u \times u_y)\Big) dx \wedge dy.$$

Proof of Lemma 2.1. Let us define the one-forms

$$\beta = u \times (*du) + du = (u \times u_x + u_y)dy + (u_x - u \times u_y)dx,$$

and

$$\alpha = \frac{1}{x^2+y^2}(x\beta - y\,u \times \beta).$$

Then

$$d\alpha = \frac{1}{x^2+y^2}(x - y\,u\times)d\beta + d\left(\frac{1}{x^2+y^2}(x - y\,u\times)\right) \wedge \beta \tag{2.2}$$
$$= I + II,$$

where I and II represent the two terms on the right-hand side of (2.2). Notice that

$$d\beta = (u \times \Delta u)dx \wedge dy = u \times \mathcal{T}\,\sigma^2 dx \wedge dy,$$

and therefore, since $u \times (u \times v) = -v$ for any vector v perpendicular to u, we have

$$I = \frac{1}{x^2+y^2}(x\,u \times \mathcal{T} + y\,\mathcal{T})\sigma^2 dx \wedge dy.$$

Meanwhile,

$$
\begin{aligned}
II &= \frac{1}{x^2+y^2}(dx - dy\, u\times -y\, dx\, u_x \times -y\, dy\, u_y\times)\wedge\beta \\
&\quad -\frac{1}{(x^2+y^2)^2}(2xdx+2ydy)(x-y\,u\times)\wedge\beta \\
&= \frac{1}{x^2+y^2}\left((1-y\,u_x\times)(u\times u_x+u_y)+(u\times +y\,u_y\times)(u_x-u\times u_y)\right)dx\wedge dy \\
&\quad -\frac{2}{(x^2+y^2)^2}(x-y\,u\times)\left(x(u\times u_x+u_y)-y(u_x-u\times u_y)\right)dx\wedge dy \\
&= \frac{1}{x^2+y^2}\left(2u\times u_x+2u_y-y\,u|u_x|^2-y\,u|u_y|^2-2y\,u_x\times u_y\right)dx\wedge dy \\
&\quad -\frac{2}{(x^2+y^2)^2}(x^2+y^2)(u\times u_x+u_y)dx\wedge dy \\
&= -\frac{y}{x^2+y^2}\left(u(|u_x|^2+|u_y|^2)+2\,u_x\times u_y\right)dx\wedge dy.
\end{aligned}
$$

Here we are using the fact that u is orthogonal to u_x and u_y, and hence that $u_x\times(u\times u_x)=u|u_x|^2$, and likewise for u_y. Observing that $u|u_x-u\times u_y|^2=u(|u_x|^2+|u_y|^2)+2\,u_x\times u_y$, we may assemble the expression for $d\alpha$

$$
d\alpha=\frac{1}{x^2+y^2}(x\,u\times\mathcal{T}+y\,\mathcal{T})\sigma^2dx\wedge dy-\frac{y}{x^2+y^2}u|u_x-u\times u_y|^2dx\wedge dy.
$$

From here, we may integrate $d\alpha$ over the annulus $D_R\backslash D_\varepsilon$ for $R>\varepsilon>0$. Writing C_r for the circle centred at the origin of radius r with an anticlockwise orientation, we apply Stokes' theorem to find that

$$
\int_{D_R\backslash D_\varepsilon} d\alpha=\int_{C_R}\alpha-\int_{C_\varepsilon}\alpha. \tag{2.3}
$$

Let us consider this expression in the limits $R\to\infty$ and $\varepsilon\to 0$. We find that

$$
\begin{aligned}
\int_{C_r}\alpha &= \int_{C_r}\frac{1}{x^2+y^2}\left((xdy-ydx)(u\times u_x+u_y)+(xdx+ydy)(u_x-u\times u_y)\right) \\
&= \int_{C_r}\frac{1}{r}(u\times u_x+u_y)ds,
\end{aligned}
$$

where ds is the normal length form on C_r. Therefore

$$
\left|\int_{C_R}\alpha\right|\leq 2\pi\max_{C_R}|u\times u_x+u_y|\to 0, \tag{2.4}
$$

as $R\to\infty$ because u was originally a smooth function on S^2 and so $|u_x|^2+|u_y|^2\to 0$ as $x^2+y^2\to\infty$. Moreover,

$$
\int_{C_\varepsilon}\alpha\to 2\pi(u\times u_x+u_y)(0,0) \tag{2.5}
$$

as $\varepsilon \to 0$ since u is C^1 at the origin $(0,0)$. Combining (2.3), (2.4) and (2.5), we conclude the first part of the lemma. The second part follows in the same way, only now we replace α by the form

$$\frac{1}{(x-a)^2+(y-b)^2}((x-a)\beta-(y-b)\,u\times\beta)\varphi,$$

and work with circles C_r centred at (a,b) rather than the origin. □

2.1.2. *Riesz potential estimates.* Riesz potentials will arise many times during the proof of Theorem 1.2 — especially when we prove L^p estimates for e_∂ and $e_{\bar\partial}$ in Section 2.1.3, when we look at the Hopf differential in Section 2.1.4, when we control the size of antiholomorphic bubbles in Section 2.4 and when we analyse necks in Section 2.2.

LEMMA 2.2. *Suppose $f \in L^1(\mathbb{R}^2,\mathbb{R})$, and $g:\mathbb{R}^2\to\mathbb{R}$ is defined by*

$$g(\mathbf{a}) = \int_{\mathbb{R}^2}\frac{f(\mathbf{x})}{|\mathbf{x}-\mathbf{a}|}d\mathbf{x}.$$

(i) *If $q\in(1,2)$ and $f\in L^q(\mathbb{R}^2)$ then there exists $C=C(q)$ such that*

$$\|g\|_{L^{\frac{2q}{2-q}}(\mathbb{R}^2)} \le C\|f\|_{L^q(\mathbb{R}^2)}.$$

(ii) *For each $q\in[1,2)$ there exists $C=C(q)$ such that*

$$\|g\|_{L^q(D)} \le C\|f\|_{L^1(\mathbb{R}^2)}.$$

(iii) *There exist positive universal constants C_1 and C_2 such that if $f(\mathbf{x})=0$ for $\mathbf{x}\in\mathbb{R}^2\backslash D_2$, and $f\in L^2(D_2)$, then*

$$\int_{D_2}\exp\left[\frac{g(\mathbf{x})}{C_1\|f\|_{L^2(D_2)}}\right]^2 d\mathbf{x} \le C_2.$$

(iv) *If $\Omega\subset\mathbb{R}^2$ is a measurable set, of finite measure $|\Omega|$, then*

$$\|g\|_{L^1(\Omega)} \le 2(\pi|\Omega|)^{\frac{1}{2}}\|f\|_{L^1(\mathbb{R}^2)}.$$

(v) *There exists a universal constant C such that*

$$|g(\mathbf{a})|^2 \le C\|f\|_{L^1(\mathbb{R}^2)}\mathcal{M}f(\mathbf{a}),$$

for each $\mathbf{a}\in\mathbb{R}^2$, where $\mathcal{M}f$ represents the Hardy-Littlewood maximal function corresponding to f (see [12, §1.1]).

We remark that the well-known analytic fact that part (iii) cannot be improved to an L^∞ bound for g, will later manifest itself in the geometric fact that antiholomorphic bubbles may occur attached anywhere on a holomorphic body map, in an almost harmonic map.

Proof of Lemma 2.2. For parts (i) and (iii) we direct the reader to [18, Th. 2.8.4] and [4, Lemma 7.13] respectively. The latter proof involves controlling the blow-up of the L^n norms of g in terms of n (as $n \to \infty$) sufficiently well that the exponential sum converges.

Part (ii). We observe that

$$|g(\mathbf{a})| \leq \int_{\mathbb{R}^2} \frac{|f(\mathbf{x})|^{\frac{1}{q}}}{|\mathbf{x}-\mathbf{a}|} |f(\mathbf{x})|^{1-\frac{1}{q}} d\mathbf{x} \leq \left(\int_{\mathbb{R}^2} \frac{|f(\mathbf{x})|}{|\mathbf{x}-\mathbf{a}|^q} d\mathbf{x} \right)^{\frac{1}{q}} \|f\|_{L^1(\mathbb{R}^2)}^{1-\frac{1}{q}},$$

using Hölder's inequality, and therefore

$$\|g\|_{L^q(D)} \leq \|f\|_{L^1(\mathbb{R}^2)}^{1-\frac{1}{q}} \left(\sup_{\mathbf{x}\in\mathbb{R}^2} \int_D \frac{1}{|\mathbf{x}-\mathbf{a}|^q} d\mathbf{a} \right)^{\frac{1}{q}} \|f\|_{L^1(\mathbb{R}^2)}^{\frac{1}{q}} \leq C(q) \|f\|_{L^1(\mathbb{R}^2)}.$$

Part (iv). We need merely to combine the observation

$$\|g\|_{L^1(\Omega)} \leq \|f\|_{L^1(\mathbb{R}^2)} \left(\int_\Omega \frac{1}{|\mathbf{x}-\mathbf{a}|} d\mathbf{a} \right),$$

with the 'symmetrisation' estimate

$$\int_\Omega \frac{1}{|\mathbf{x}-\mathbf{a}|} d\mathbf{a} \leq \int_{\hat{\Omega}} \frac{1}{|\mathbf{x}-\mathbf{a}|} d\mathbf{a} = 2(\pi|\Omega|)^{\frac{1}{2}},$$

where $\hat{\Omega}$ is the disc centred at $\mathbf{x}$, having the same measure as Ω.

Part (v). We begin with the change of variables

$$g(\mathbf{a}) = \int_{\mathbb{R}^2} \frac{f(\mathbf{x}+\mathbf{a})}{|\mathbf{x}|} d\mathbf{x}. \tag{2.6}$$

Now for any $R > 0$, we have

$$\int_{\{|\mathbf{x}|\geq R\}} \frac{|f(\mathbf{x}+\mathbf{a})|}{|\mathbf{x}|} d\mathbf{x} \leq \frac{1}{R} \|f\|_{L^1(\mathbb{R}^2)},$$

and the complementary estimate

$$\begin{aligned} \int_{\{|\mathbf{x}|<R\}} \frac{|f(\mathbf{x}+\mathbf{a})|}{|\mathbf{x}|} d\mathbf{x} &\leq \sum_{k=0}^{\infty} \int_{2^{-k-1}R\leq|\mathbf{x}|<2^{-k}R} \frac{|f(\mathbf{x}+\mathbf{a})|}{|\mathbf{x}|} d\mathbf{x} \\ &\leq \sum_{k=0}^{\infty} \frac{1}{2^{-k-1}R} \int_{|\mathbf{x}|<2^{-k}R} |f(\mathbf{x}+\mathbf{a})| d\mathbf{x} \\ &\leq \sum_{k=0}^{\infty} \frac{1}{2^{-k-1}R} (2^{-k}R)^2 \mathcal{M}f(\mathbf{a}) = 4R\mathcal{M}f(\mathbf{a}). \end{aligned}$$

Combining these two estimates with (2.6) and setting

$$R = \left(\frac{\|f\|_{L^1(\mathbb{R}^2)}}{\mathcal{M}f(\mathbf{a})} \right)^{\frac{1}{2}}$$

(or taking a limit if $R = 0$ or $R = \infty$) we conclude that

$$|g(\mathbf{a})| \leq \frac{1}{R}\|f\|_{L^1(\mathbb{R}^2)} + 4R\mathcal{M}f(\mathbf{a}) = 5\left(\|f\|_{L^1(\mathbb{R}^2)}\mathcal{M}f(\mathbf{a})\right)^{\frac{1}{2}}. \qquad \square$$

2.1.3. *L^p estimates for e_∂ and $e_{\bar\partial}$.* The following lemma provides control of ∂-energies and $\bar\partial$-energies (and their higher p-energies) which we shall require on numerous occasions in this work. The estimates are variations and extensions of the global key lemma from [15]. In practice, the disc $D_{2\mu}$ will always arise as a disc in a stereographic coordinate chart.

LEMMA 2.3. *Suppose that $\mu \in (0,1]$ and that $u : D_{2\mu} \to S^2 \hookrightarrow \mathbb{R}^3$ is smooth, and recall the definition of $\mathcal{T}$ from (2.1). Then we have the following estimates for $e_\partial(u)$ and $e_{\bar\partial}(u)$:*

(a) *Given $p \in [1,\infty)$, there exist $\varepsilon_0 = \varepsilon_0(p) \in (0,1]$ and $C = C(\mu,p)$ such that if $E_\partial(u, D_{2\mu}) < \varepsilon_0$ and $\|\mathcal{T}\sigma\|_{L^2(D_{2\mu})} \leq 1$ then*

$$\|u_x - u \times u_y\|_{L^p(D_{\frac{3\mu}{2}})} < C.$$

(b) *There exist universal constants $\varepsilon_1 \in (0,1]$ and C such that whenever $E_\partial(u, D_{2\mu}) < \varepsilon_1$ and $b \in D_{2\mu}$, $\nu \in (0,1)$ satisfy $D_{b,e\nu} \subset D_{2\mu}$, we have the estimate*

$$E_\partial(u, D_{b,\nu}) \leq C\left(\|\mathcal{T}\sigma\|^2_{L^2(D_{b,e\nu})} + E(u, D_{b,e\nu}\backslash D_{b,\nu})\right).$$

(c) *Given $q \in (1,2)$, there exist $\varepsilon_2 = \varepsilon_2(q) \in (0,1]$ and $C = C(\mu,q)$ such that if $E_{\bar\partial}(u, D_{2\mu}) < \varepsilon_2$ then*

$$E_{\bar\partial}(u, D_\mu) \leq C\left(\|\mathcal{T}\sigma\|^2_{L^2(D_{2\mu})} + \|u_x + u \times u_y\|^2_{L^q(D_{2\mu}\backslash D_\mu)}\right).$$

(d) *Given $q \in (1,2)$ and $l \in \mathbb{N}$, there exist $\varepsilon_3 = \varepsilon_3(q) \in (0,1]$ and $C = C(\mu,l,q)$ achieving the following. Suppose that $\nu > 0$ and that for each $i \in \{1,\ldots,l\}$, we have points $b^i \in D_\mu$ with $D_{b^i,e\nu} \subset D_\mu$ disjoint discs. Then writing*

$$\Lambda = D_\mu\backslash(\textstyle\bigcup_i D_{b^i,e\nu}), \qquad \text{and} \qquad \hat\Lambda = D_{2\mu}\backslash(\textstyle\bigcup_i D_{b^i,\nu}),$$

whenever $E_{\bar\partial}(u,\hat\Lambda) < \varepsilon_3$, we have

$$E_{\bar\partial}(u,\Lambda) \leq C\left(\|\mathcal{T}\sigma\|^2_{L^2(D_{2\mu})} + E_{\bar\partial}(u, D_{2\mu}\backslash D_\mu) + \nu^{-\frac{4(q-1)}{q}} E(u, \bigcup_i (D_{b^i,e\nu}\backslash D_{b^i,\nu}))\right).$$

Remark 2.4. Each ∂-energy estimate in Lemma 2.3 has a $\bar{\partial}$-energy equivalent — and vice-versa — which arises by composing u with a reflection in the target S^2. Reflections in the target are orientation reversing isometries. Therefore we need only prove ∂-energy estimates.

Proof of Lemma 2.3. Our starting point is the second integral formula of Lemma 2.1. Let us adopt the shorthand $\mathbf{a} = (a, b)$ and $\mathbf{x} = (x, y)$, and assume that φ has compact support $\Omega \subset D_2$ and range in $[0, 1]$. Then we have

$$\begin{aligned}2\pi|u_x - u \times u_y|\varphi(a,b) \leq &\int_{\mathbb{R}^2} \frac{1}{|\mathbf{x}-\mathbf{a}|}\left(|\mathcal{T}|\sigma^2 + |u_x - u\times u_y|^2\right)\varphi\, dx\wedge dy \\ &+ \int_{\mathbb{R}^2} \frac{1}{|\mathbf{x}-\mathbf{a}|}|\nabla\varphi|.|u_x - u\times u_y| dx\wedge dy.\end{aligned}$$

For any $q \in (1, 2)$, we may now take the $L^{\frac{2q}{2-q}}$ norm, and apply part (i) of Lemma 2.2, to give

$$\begin{aligned}&\|(u_x - u\times u_y)\varphi\|_{L^{\frac{2q}{2-q}}} \\ &\quad\leq C\left(\|\mathcal{T}\sigma^2\varphi\|_{L^q} + \||u_x - u\times u_y|^2\varphi\|_{L^q} + \||\nabla\varphi|.|u_x - u\times u_y|\|_{L^q}\right),\end{aligned} \tag{2.7}$$

where C depends on q. Let us take a closer look at the individual terms on the right-hand side of (2.7). Using Hölder's inequality, the bound on the range and the support Ω of φ, and the fact that $\sigma \leq 2$, we see that

$$\|\mathcal{T}\sigma^2\varphi\|_{L^q(\mathbb{R}^2)} \leq C\|\mathcal{T}\sigma\|_{L^2(\Omega)},$$

for some universal C. Meanwhile, an alternative application of Hölder's inequality yields

$$\||u_x - u\times u_y|^2\varphi\|_{L^q(\mathbb{R}^2)} \leq \|u_x - u\times u_y\|_{L^2(\Omega)}\|(u_x - u\times u_y)\varphi\|_{L^{\frac{2q}{2-q}}(\mathbb{R}^2)}.$$

Returning to (2.7) this means that

$$\begin{aligned}&\|(u_x - u\times u_y)\varphi\|_{L^{\frac{2q}{2-q}}} \leq C_0\Big(\|\mathcal{T}\sigma\|_{L^2(\Omega)} + E_\partial(u,\Omega)^{\frac{1}{2}}\|(u_x - u\times u_y)\varphi\|_{L^{\frac{2q}{2-q}}} \\ &\qquad\qquad + \||\nabla\varphi|.|u_x - u\times u_y|\|_{L^q}\Big),\end{aligned} \tag{2.8}$$

with C_0 dependent only on q. If we now choose any $\varepsilon \in (0, (2C_0)^{-2})$ then whenever $E_\partial(u, \Omega) < \varepsilon$, we may absorb a term on the right-hand side of (2.8) into the left-hand side, and deduce that

$$\|(u_x - u\times u_y)\varphi\|_{L^{\frac{2q}{2-q}}} \leq C_1\left(\|\mathcal{T}\sigma\|_{L^2(\Omega)} + \||\nabla\varphi|.|u_x - u\times u_y|\|_{L^q}\right), \tag{2.9}$$

where $C_1 = 2C_0$ is dependent only on q. This is the estimate which we refine in different directions to yield the four parts of Lemma 2.3.

Part (a). Let us choose φ to satisfy $\varphi \equiv 1$ on $D_{\frac{3\mu}{2}}$, and support$(\varphi) \subset\subset D_{2\mu}$ with $|\nabla\varphi| \leq \frac{4}{\mu}$ at each point. We retain the restriction that the range of φ lies within $[0,1]$. In this case, (2.9) tells us that

$$\|u_x - u \times u_y\|_{L^{\frac{2q}{2-q}}(D_{\frac{3\mu}{2}})} \leq C_1 \left(\|\mathcal{T}\sigma\|_{L^2(D_{2\mu})} + \frac{4}{\mu}\|u_x - u \times u_y\|_{L^q(D_{2\mu})} \right),$$

whenever $E_\partial(u, D_{2\mu}) < \varepsilon$. Now Hölder's inequality tells us that

$$\|u_x - u \times u_y\|_{L^q(D_{2\mu})} \leq C\|u_x - u \times u_y\|_{L^2(D_{2\mu})} = E_\partial(u, D_{2\mu})^{\frac{1}{2}},$$

for some C which may be considered universal since $\mu \leq 1$. Therefore, with the hypotheses of part (a), we find that

$$\|u_x - u \times u_y\|_{L^{\frac{2q}{2-q}}(D_{\frac{3\mu}{2}})} \leq C,$$

for $C = C(\mu, q)$. This establishes part (a) for $p \in (2, \infty)$. The case $p \in [1,2]$ follows simply from the Hölder estimate

$$\|u_x - u \times u_y\|_{L^p(D_{\frac{3\mu}{2}})} \leq C\|u_x - u \times u_y\|_{L^2(D_{\frac{3\mu}{2}})} \leq CE_\partial(u, D_{2\mu})^{\frac{1}{2}},$$

where C may be considered universal since $\mu \leq 1$.

Part (b). Now we redefine φ to satisfy $\varphi \equiv 1$ on $D_{b,\nu}$, and support$(\varphi) \subset\subset D_{b,e\nu}$ with $|\nabla\varphi| \leq \frac{1}{\nu}$ at each point. As always, we retain the restriction that the range of φ lies within $[0,1]$. Then (2.9) and Hölder's inequality tell us that

$$\begin{aligned}\|u_x - u \times u_y\|_{L^2(D_{b,\nu})} &\leq (\pi\nu^2)^{\frac{q-1}{q}}\|u_x - u \times u_y\|_{L^{\frac{2q}{2-q}}(D_{b,\nu})} \\ &\leq C\nu^{\frac{2(q-1)}{q}}\left(\|\mathcal{T}\sigma\|_{L^2(D_{b,e\nu})} + \frac{1}{\nu}\|u_x - u \times u_y\|_{L^q(D_{b,e\nu}\backslash D_{b,\nu})}\right),\end{aligned}$$

where $C = C(q)$, but then since

$$\|u_x - u \times u_y\|_{L^q(D_{b,e\nu}\backslash D_{b,\nu})} \leq C\nu^{\frac{2-q}{q}}\|u_x - u \times u_y\|_{L^2(D_{b,e\nu}\backslash D_{b,\nu})},$$

for some universal C, and $\nu \leq 1$, we find that

$$E_\partial(u, D_{b,\nu})^{\frac{1}{2}} \leq C\left(\|\mathcal{T}\sigma\|_{L^2(D_{b,e\nu})} + \nu^{\frac{2(q-1)}{q}}\nu^{-1}\nu^{\frac{2-q}{q}}E_\partial(u, D_{b,e\nu}\backslash D_{b,\nu})^{\frac{1}{2}}\right),$$

with $C = C(q)$. This clearly implies

$$E_\partial(u, D_{b,\nu}) \leq C\left(\|\mathcal{T}\sigma\|^2_{L^2(D_{b,e\nu})} + E_\partial(u, D_{b,e\nu}\backslash D_{b,\nu})\right), \tag{2.10}$$

for some C which we may consider universal by fixing $q \in (1,2)$ at $q = \frac{3}{2}$ say, and (2.10) is just a slightly stronger version of part (b) of Lemma 2.3 since the ∂-energy cannot exceed the ordinary energy.

Part (c). This part is little different from part (b). We require φ to satisfy $\varphi \equiv 1$ on D_μ, and support$(\varphi) \subset\subset D_{2\mu}$ with $|\nabla\varphi| \leq \frac{2}{\mu}$ at each point. Tracking the proof of part (b) leads us easily to

$$E_\partial(u, D_\mu)^{\frac{1}{2}} \leq C\left(\|\mathcal{T}\sigma\|_{L^2(D_{2\mu})} + \mu^{\frac{2(q-1)}{q}}\mu^{-1}\|u_x - u \times u_y\|_{L^q(D_{2\mu}\setminus D_\mu)}\right),$$

where $C = C(q)$, but then by allowing C to depend on μ, we deduce that

$$E_\partial(u, D_\mu) \leq C\left(\|\mathcal{T}\sigma\|^2_{L^2(D_{2\mu})} + \|u_x - u \times u_y\|^2_{L^q(D_{2\mu}\setminus D_\mu)}\right),$$

which is part (c) modulo a change of orientation as discussed in Remark 2.4.

Part (d). Again, we see this part as a variation on part (b) — and part (c). Now φ should satisfy $\varphi \equiv 1$ on Λ, and support$(\varphi) \subset\subset \hat{\Lambda}$. The gradient restriction splits into $|\nabla\varphi| \leq \frac{2}{\mu}$ on the external collar $D_{2\mu}\setminus D_\mu$ and $|\nabla\varphi| \leq \frac{1}{\nu}$ on the small collars $D_{b^i,e\nu}\setminus D_{b^i,\nu}$ for each i. Note that in applications, we will have $\nu \ll \mu$. Invoking (2.9) as usual gives us

$$\begin{aligned}\|u_x - u \times u_y\|_{L^2(\Lambda)} &\leq C\mu^{\frac{2(q-1)}{q}}\|u_x - u \times u_y\|_{L^{\frac{2q}{2-q}}(\Lambda)} \\ &\leq C\mu^{\frac{2(q-1)}{q}}\Bigg(\|\mathcal{T}\sigma\|_{L^2(\hat{\Lambda})} + \frac{2}{\mu}\|u_x - u \times u_y\|_{L^q(D_{2\mu}\setminus D_\mu)} \\ &\qquad + \frac{1}{\nu}\sum_{i=1}^{l}\|u_x - u \times u_y\|_{L^q(D_{b^i,e\nu}\setminus D_{b^i,\nu})}\Bigg),\end{aligned}$$

where C is dependent only on q. If we now allow dependence of C on μ, and apply Hölder's inequality, we see that

$$\begin{aligned}E_\partial(u, \Lambda)^{\frac{1}{2}} \leq C\Bigg(&\|\mathcal{T}\sigma\|_{L^2(\hat{\Lambda})} + \|u_x - u \times u_y\|_{L^2(D_{2\mu}\setminus D_\mu)} \\ &+ \nu^{-1}\sum_{i=1}^{l}\nu^{\frac{2-q}{q}}\|u_x - u \times u_y\|_{L^2(D_{b^i,e\nu}\setminus D_{b^i,\nu})}\Bigg),\end{aligned}$$

and therefore

$$\begin{aligned}&E_\partial(u, \Lambda) \\ &\quad\leq C\left(\|\mathcal{T}\sigma\|^2_{L^2(\hat{\Lambda})} + E_\partial(u, D_{2\mu}\setminus D_\mu) + \nu^{\frac{-4(q-1)}{q}}\sum_{i=1}^{l}E_\partial(u, D_{b^i,e\nu}\setminus D_{b^i,\nu})\right),\end{aligned}$$

where $C = C(q, \mu, l)$ at this stage. This may then be weakened to

$$\begin{aligned}&E_\partial(u, \Lambda) \\ &\quad\leq C\left(\|\mathcal{T}\sigma\|^2_{L^2(D_{2\mu})} + E_\partial(u, D_{2\mu}\setminus D_\mu) + \nu^{\frac{-4(q-1)}{q}}E(u, \bigcup_i(D_{b^i,e\nu}\setminus D_{b^i,\nu}))\right),\end{aligned}$$

which is part (d) modulo a change of orientation (see Remark 2.4). $\square$

2.1.4. *Hopf differential estimates.* Given a sufficiently regular map u from a surface into $S^2 \hookrightarrow \mathbb{R}^3$, we may choose a local complex coordinate $z = x + iy$ on the domain, and define the *Hopf differential* to be the quadratic differential $\varphi(z)dz^2$ where

$$\varphi(z) := |u_x|^2 - |u_y|^2 - 2i\langle u_x, u_y\rangle.$$

In the present section we will establish various natural estimates for this quantity, when the domain is S^2 or a disc. Our main goal is to be able to control the product $e_\partial(u)e_{\bar\partial}(u)$ of the ∂ and $\bar\partial$-energy densities, and the connection here is the easily-verified identity

$$|\varphi(z)|^2 = \psi^2(x,y), \tag{2.11}$$

where the function ψ is defined by

$$\psi(x,y) := |u \times u_x + u_y|.|u \times u_x - u_y|. \tag{2.12}$$

It is worth stressing that it is these estimates which inject global information into our theory, and exploit the fact that the domain is S^2 rather than some higher genus surface in our main theorems. In contrast, it is less important that the target is S^2, and the results below have analogues applying to maps into arbitrary targets, of arbitrary dimension.

Lemma 2.5. *Suppose that* $u : S^2 \cong \mathbb{R}^2 \cup \{\infty\} \to S^2 \hookrightarrow \mathbb{R}^3$ *is smooth, with* $E(u) < M$. *Let* $\psi : \mathbb{R}^2 \to \mathbb{R}$ *be defined as in* (2.12).

(a) *There exists a universal constant* C *such that*

$$\mathrm{Area}_{\mathbb{R}^2}\{\mathbf{x} \in \mathbb{R}^2 \ : \ \psi^2(\mathbf{x}) > \eta\} \leq \frac{C}{\eta} M \|\mathcal{T}(u)\|^2_{L^2(S^2)},$$

for all $\eta > 0$.

(b) *For all* $q \in [1,2)$, *there exists* $C = C(q)$ *such that*

$$\|(e_\partial(u)e_{\bar\partial}(u))^{\frac{1}{2}}\|_{L^q(S^2)} \leq CM^{\frac{1}{2}}\|\mathcal{T}(u)\|_{L^2(S^2)}.$$

We stress that the notation $\mathrm{Area}_{\mathbb{R}^2}$ refers to area with respect to the standard metric on $\mathbb{R}^2$. In contrast, we occasionally write Area_{S^2} to compute with respect to the $\sigma^2(dx^2 + dy^2)$ metric.

These estimates are strong forms of the statement that any harmonic map from S^2 has vanishing Hopf differential, and is therefore (weakly) conformal. We record here the following well-known consequence of this fact, due to Wood and Lemaire (see [2, (11.5)]) to which we have already referred.

Lemma 2.6. *The harmonic maps between* 2*-spheres are precisely the rational maps and their complex conjugates* (*i.e. rational in* z *or* $\bar z$). *In particular such a map* u *has energy given by*

$$E(u) = 4\pi|\deg(u)|.$$

LEMMA 2.7. *Suppose that $u : D_\gamma \to S^2 \hookrightarrow \mathbb{R}^3$ is smooth, with $\gamma \in (0,1]$, and has $E(u, D_\gamma) < M \geq 1$, and $\|\mathcal{T}\sigma\|_{L^2(D_\gamma)} \leq 1$ where $\mathcal{T}$ is defined as in (2.1). Then there exists a universal constant C such that for any measurable $\Omega \subset D_{\gamma/2}$, there holds the estimate*

$$\|\psi\|_{L^1(\Omega)} \leq \frac{C}{\gamma} M |\Omega|^{\frac{1}{2}},$$

where $|\Omega|$ here represents the area in $\mathbb{R}^2$ of the set Ω.

Let us take z to be a stereographic coordinate when our domain is S^2, and the normal complex coordinate $x+iy$ when we work on D_γ. The spherical metric is given by $\sigma^2|dz|^2$, and the volume form is $\sigma^2 dx \wedge dy = \sigma^2 \frac{i}{2} dz \wedge d\bar{z}$. The basic fact underpinning these lemmata is that

$$\varphi_{\bar{z}} := \frac{1}{2}(\varphi_x + i\varphi_y) = \langle u_x - iu_y, \Delta u\rangle = \sigma^2 \langle u_x - iu_y, \mathcal{T}(u)\rangle, \tag{2.13}$$

which is easily verified by direct calculation. In particular, any harmonic map from an orientable surface has holomorphic Hopf differential, which must then vanish if the domain has genus zero. Thus we have the main ingredient of Lemma 2.6.

A stronger consequence of (2.13) is that $|\varphi_{\bar{z}}| \leq \sigma^2 |\nabla u|.|\mathcal{T}|$, and hence that

$$\|\varphi_{\bar{z}}\|_{L^1(\Sigma)} \leq 2\sqrt{2} E(u, \Sigma)^{\frac{1}{2}} \|\sigma\mathcal{T}\|_{L^2(\Sigma)}, \tag{2.14}$$

for any measurable $\Sigma \subset \mathbb{R}^2$, since $\sigma \leq 2$.

Proof of Lemma 2.5. An application of Cauchy's integral formula to φ, over the domain D_r yields

$$\varphi(w) = \frac{1}{2\pi i} \int_{D_r} \frac{\varphi_{\bar{z}}(z)}{z-w} dz \wedge d\bar{z} + \frac{1}{2\pi i} \int_{\partial D_r} \frac{\varphi(z)}{z-w} dz, \tag{2.15}$$

where ∂D_r is given an anticlockwise orientation. Since φdz^2 is a quadratic differential on the sphere, the function $\varphi(z)$ must decay like $\frac{1}{|z|^2}$ as $|z| \to \infty$, and therefore the boundary term of (2.15) vanishes in the limit $r \to \infty$ to give

$$\varphi(w) = \frac{1}{2\pi i} \int_{\mathbb{R}^2} \frac{\varphi_{\bar{z}}(z)}{z-w} dz \wedge d\bar{z},$$

and the corresponding inequality

$$|\varphi(w)| \leq \frac{1}{\pi} \int_{\mathbb{R}^2} \frac{|\varphi_{\bar{z}}(z)|}{|z-w|} dx \wedge dy. \tag{2.16}$$

To prove part (a) of the lemma, we proceed by invoking part (v) of Lemma 2.2 which tells us that

$$|\varphi(w)|^2 \leq C \|\varphi_{\bar{z}}\|_{L^1(\mathbb{R}^2)} \mathcal{M}|\varphi_{\bar{z}}|(w), \tag{2.17}$$

for some universal C. The fundamental theorem for the maximal function — see [12, Th. 1b] — tells us that

$$\mathrm{Area}_{\mathbb{R}^2}\{\mathbf{x}\in\mathbb{R}^2\ :\ \mathcal{M}|\varphi_{\bar z}|(\mathbf{x})>\eta\}\leq\frac{C}{\eta}\|\varphi_{\bar z}\|_{L^1(\mathbb{R}^2)},$$

for all $\eta>0$, where C is universal. Therefore, using (2.17) and (2.14) this gives

$$\begin{aligned}\mathrm{Area}_{\mathbb{R}^2}\{\mathbf{x}\in\mathbb{R}^2\ :\ |\varphi(\mathbf{x})|^2>\eta\}&\leq\frac{C}{\eta}\|\varphi_{\bar z}\|^2_{L^1(\mathbb{R}^2)}\\&\leq\frac{C}{\eta}E(u,\mathbb{R}^2)\|\sigma\mathcal{T}\|^2_{L^2(\mathbb{R}^2)}\leq\frac{C}{\eta}M\|\mathcal{T}(u)\|^2_{L^2(S^2)},\end{aligned}$$

for all $\eta>0$ (and for some new universal constants C) which in view of (2.11) is simply part (a) of the lemma.

For part (b) of the lemma, we apply part part (ii) of Lemma 2.2 to (2.16), and use (2.14), to give

$$\|\varphi\|_{L^q(D)}\leq C\|\varphi_{\bar z}\|_{L^1(\mathbb{R}^2)}\leq C\,E(u,\mathbb{R}^2)^{\frac12}\|\sigma\mathcal{T}\|_{L^2(\mathbb{R}^2)}\leq C\,M^{\frac12}\|\mathcal{T}(u)\|_{L^2(S^2)},\tag{2.18}$$

for $q\in[1,2)$ and $C=C(q)$. Meanwhile we have, by definition, that

$$(e_\partial(u)e_{\bar\partial}(u))^{\frac12}=\frac{1}{4\sigma^2}|u\times u_x+u_y|.|u\times u_x-u_y|=\frac{1}{4\sigma^2}\psi=\frac{1}{4\sigma^2}|\varphi|.$$

Therefore, if we denote the hemisphere in S^2 corresponding to the disc $D\in\mathbb{R}^2$ (via stereographic projection) by S^2_+, we may appeal to (2.18) and calculate that

$$\begin{aligned}\|(e_\partial(u)e_{\bar\partial}(u))^{\frac12}\|_{L^q(S^2_+)}&=\left(\int_D\left(\frac{1}{4\sigma^2}|\varphi|\right)^q\sigma^2\right)^{\frac1q}\\&\leq\left(\int_D|\varphi|^q\right)^{\frac1q}\leq C\,M^{\frac12}\|\mathcal{T}(u)\|_{L^2(S^2)},\end{aligned}$$

for $q\in[1,2)$ and $C=C(q)$, since $\sigma\geq1$ on D. Repeating this estimate on the opposite hemisphere S^2_-, and combining, yields part (b) of the lemma. □

Proof of Lemma 2.7. Let $\zeta:D_\gamma\to[0,1]$ be a smooth cut-off function, with compact support, and with the properties that $\zeta\equiv1$ on $D_{\frac{\gamma}{2}}$ and $|\nabla\zeta|\leq\frac{4}{\gamma}$. Now, an application of Cauchy's theorem gives us

$$\zeta(w)\varphi(w)=\frac{1}{2\pi i}\int_D\frac{(\zeta\varphi)_{\bar z}(z)}{z-w}dz\wedge d\bar z,$$

and by calculating

$$|(\zeta\varphi)_{\bar z}|\leq\zeta|\varphi_{\bar z}|+|\zeta_{\bar z}\varphi|\leq|\varphi_{\bar z}|+\frac{2}{\gamma}\psi,$$

this gives us

$$|\zeta\varphi|(w) \leq \frac{1}{\pi}\int_D \frac{|\varphi_{\bar{z}}|(z)}{|z-w|}dx\wedge dy + \frac{2}{\gamma\pi}\int_D \frac{\psi(z)}{|z-w|}dx\wedge dy.$$

We are then able to apply part part (iv) of Lemma 2.2 to each term, and deduce that

$$\|\zeta\varphi\|_{L^1(\Omega)} \leq C|\Omega|^{\frac{1}{2}}\left(\|\varphi_{\bar{z}}\|_{L^1(D)} + \frac{1}{\gamma}\|\psi\|_{L^1(D)}\right),$$

for some universal C. But now (2.14) and our hypothesis $\|\mathcal{T}\sigma\|_{L^2(D_\gamma)} \leq 1$ implies

$$\|\varphi_{\bar{z}}\|_{L^1(D)} \leq CM^{\frac{1}{2}}\|\sigma\mathcal{T}\|_{L^2(D)} \leq CM^{\frac{1}{2}},$$

and Young's inequality (and the definitions of ψ, E_∂ and $E_{\bar{\partial}}$) immediately give us

$$\|\psi\|_{L^1(D)} \leq 2(E_\partial(u,D) + E_{\bar{\partial}}(u,D)) = 2E(u,D) \leq 2M,$$

and so we may conclude that

$$\|\psi\|_{L^1(\Omega)} = \|\zeta\varphi\|_{L^1(\Omega)} \leq C|\Omega|^{\frac{1}{2}}(M^{\frac{1}{2}} + \frac{1}{\gamma}M) \leq \frac{C}{\gamma}M|\Omega|^{\frac{1}{2}}$$

for universal constants C, since $\Omega \subset D_{\gamma/2}$, $M > 1$ and $\gamma \in (0,1]$. □

2.2. *Neck analysis.* In this section we derive an energy decay estimate to be applied to annular regions of the domain surrounding bubbles. Coupled with estimates of the degree of concentration of antiholomorphic bubbles which develop on holomorphic body maps (in terms of the tension) this analysis will guarantee the existence of dyadic annuli around such bubbles, on which the energy is extremely small (again in terms of the tension) and this will allow us to perform a programme of analytic surgery. Indeed, we will be able to use this smallness of energy to make energy estimates on different portions of a bubble tree and still be able to control 'boundary terms' arising from regions where the components join together.

There is a series of recent papers concerned with controlling the oscillation of maps over annular neck regions (although not exclusively in terms of the tension) which contain techniques on which we can draw for our purposes. Parker [6] made an analysis of neck regions in the context of bubbling in sequences of harmonic maps. Qing and Tian [8] extended these results to the case of almost harmonic maps. An alternative proof was then given by Lin and Wang [5]. In some sense, our route in this section is via stengthened versions of the key estimates in the final paper mentioned, and we adopt their notation where possible.

Remark 2.8. For consistency and simplicity, we phrase our results for maps into S^2; however, analogous results hold for arbitrary compact target manifolds, with essentially the same proof.

The following lemma is the only result from this section which we shall use elsewhere in this paper — and is thus the only result which need be understood on a first reading.

LEMMA 2.9. *For $\gamma \in (0,1]$, $M > 0$ and $r \in (0, e^{-4}]$, let us suppose that $u : D_\gamma \to S^2 \hookrightarrow \mathbb{R}^3$ is smooth and satisfies the constraints*

$$E(u, D_\gamma) = \frac{1}{2}\|\nabla u\|^2_{L^2(D_\gamma)} < M; \qquad \|\mathcal{T}\varsigma\|_{L^2(D_\gamma)} \leq 1, \tag{2.19}$$

where

$$\mathcal{T} := \frac{1}{\varsigma^2}\left(\Delta u + u|\nabla u|^2\right) \tag{2.20}$$

on D_γ, and $\varsigma : D_\gamma \to [1,2]$ is smooth. We assign notation to the 'fat' annulus $\Sigma := D_\gamma \backslash D_{\gamma r^2}$ and the 'dyadic' annulus $\Omega := D_{\gamma e r} \backslash D_{\gamma r}$.

Then there exist $\delta > 0$ universal and $K > 0$ dependent only on M, such that whenever

$$E(u, \Sigma) < \delta$$

and

$$\|\mathcal{T}\varsigma\|^2_{L^2(\Sigma)} < \delta,$$

then we have the estimates

$$E(u, \Omega) \leq K\, r \tag{2.21}$$

and

$$\operatorname{osc}(u, \Omega) := \sup_{\mathbf{x}, \mathbf{a} \in \Omega} |u(\mathbf{x}) - u(\mathbf{a})| \leq K r^{\frac{1}{2}}. \tag{2.22}$$

Remark 2.10. The tension $\mathcal{T}$ here is with respect to the metric $\varsigma^2(dx^2 + dy^2)$. The reader may imagine ς to be the usual conformal factor σ during the proof, but in applications, we will want to analyse on annuli, in a stereographic chart, which are not centred at the origin, and consequently ς will typically be some translation of σ (i.e. $\varsigma(\mathbf{x}) := \sigma(\mathbf{x} - \mathbf{a})$). Assuming that the annulus lies within the unit disc D inside the stereographic chart, then we can guarantee that the constraint $\varsigma \in [1,2]$ holds.

In practice, the lemma will be applied for $0 < r \ll 1$.

The remainder of this section will be devoted to the proof of Lemma (2.9). We will normally not work with the usual (x, y) coordinates on the domain Σ, but with the conformally equivalent cylindrical coordinates (t, θ) defined by

$$t = -\ln\sqrt{x^2 + y^2}; \qquad \tan\theta = \frac{y}{x},$$

with $(t, \theta) \in I \times S^1$, where $I := (-\ln\gamma, -\ln(\gamma r^2)]$ throughout this section, and θ will normally be assumed to take values in $[0, 2\pi)$. Rewriting (2.20) in these coordinates gives

$$\hat{\mathcal{T}} := \mathcal{T}\varsigma^2 e^{-2t} = u_{tt} + u_{\theta\theta} + u(|u_t|^2 + |u_\theta|^2). \tag{2.23}$$

Now we may see $\hat{\mathcal{T}}$ as the tension of u from the cylinder $I \times S^1$ with the standard cylinder metric $(dt^2 + d\theta^2)$, and in this framework we may apply a 'small-energy' estimate in the spirit of the work of Sacks and Uhlenbeck [9]. For example, the following lemma is a special case of a minor adaptation of [1, Lemma 2.1].

Lemma 2.11. *There exist universal constants* $\delta_0 \in (0,1]$ *and* $C > 0$ *such that any map* $u \in W^{2,2}([-1,2]\times S^1, S^2)$ *which satisfies* $\frac{1}{2}\|\nabla u\|^2_{L^2([-1,2]\times S^1)} < \delta_0$ *must obey the inequality*

$$\|u - \bar{u}\|_{W^{2,2}([0,1]\times S^1)} \leq C\left(\|\nabla u\|_{L^2([-1,2]\times S^1)} + \|\hat{\mathcal{T}}\|_{L^2([-1,2]\times S^1)}\right),$$

where $\bar{u}$ *is the average value of* u *over* $[-1,2]\times S^1$, *and*

$$\hat{\mathcal{T}} := u_{tt} + u_{\theta\theta} + u(|u_t|^2 + |u_\theta|^2).$$

We will use this lemma in the second part of the following elementary result.

Corollary 2.12. *There exists a universal constant* C *such that for any map* $u \in C^\infty([-1,2]\times S^1, S^2)$ *and for any* $t \in [0,1]$, *there holds the estimate*

$$\sup_{S^1} |u_\theta|^2(t,\cdot) \leq C \int_{\{t\}\times S^1} |u_{\theta\theta}|^2. \tag{2.24}$$

If, in addition, u *satisfies* $\frac{1}{2}\|\nabla u\|^2_{L^2([-1,2]\times S^1)} < \delta_0$ *(where* δ_0 *originates in Lemma* 2.11*) then*

$$\int_{\{t\}\times S^1} (|u_\theta|^2 + |u_t|^2) \leq C\left(\|\nabla u\|_{L^2([-1,2]\times S^1)} + \|\hat{\mathcal{T}}\|_{L^2([-1,2]\times S^1)}\right)^2. \tag{2.25}$$

Proof of Corollary 2.12. Given any $t \in [0,1]$ and $\theta_0 \in S^1$, let $\hat{n} \in \mathbb{R}^3$ be the fixed unit vector in the direction of $u_\theta(t,\theta_0)$. Well, since

$$\int_{\{t\}\times S^1} \langle u_\theta(t,\cdot), \hat{n}\rangle = \int_{\{t\}\times S^1} \langle u(t,\cdot), \hat{n}\rangle_\theta = 0,$$

the function $\langle u_\theta(t,\cdot), \hat{n}\rangle$ must take the value zero somewhere on S^1, and therefore

$$\begin{aligned}|u_\theta(t,\theta_0)| &= \langle u_\theta(t,\theta_0), \hat{n}\rangle \leq \operatorname{osc}(\langle u_\theta(t,\cdot), \hat{n}\rangle, S^1)\\ &\leq \int_{\{t\}\times S^1} |\langle u_{\theta\theta}(t,\cdot), \hat{n}\rangle| \leq \int_{\{t\}\times S^1} |u_{\theta\theta}|,\end{aligned}$$

by the Fundamental Theorem of Calculus. Then (2.24) follows immediately via the Cauchy-Schwarz inequality.

Meanwhile, with a view to proving (2.25) we define the function $f : [0,1] \to \mathbb{R}$ by

$$f(t) := \int_{\{t\}\times S^1} (|u_\theta|^2 + |u_t|^2).$$

Clearly, we have

$$\min_{[0,1]} f \leq \int_0^1 f(t)dt = \int_{[0,1]\times S^1} (|u_\theta|^2 + |u_t|^2) = \|\nabla u\|^2_{L^2([0,1]\times S^1)}. \tag{2.26}$$

But

$$f'(t) = 2\int_{\{t\}\times S^1} (\langle u_\theta, u_{\theta t}\rangle + \langle u_t, u_{tt}\rangle),$$

and so by Young's inequality

$$|f'(t)| \leq \int_{\{t\}\times S^1} (|u_\theta|^2 + |u_{\theta t}|^2 + |u_t|^2 + |u_{tt}|^2),$$

which allows us to deduce that

$$\begin{aligned} \operatorname{osc}(f,[0,1]) &\leq \int_0^1 |f'(t)| \leq \int_{[0,1]\times S^1} (|u_\theta|^2 + |u_{\theta t}|^2 + |u_t|^2 + |u_{tt}|^2) \\ &\leq C\left(\|\nabla u\|_{L^2([-1,2]\times S^1)} + \|\hat{\mathcal{T}}\|_{L^2([-1,2]\times S^1)}\right)^2, \end{aligned} \tag{2.27}$$

from the Fundamental Theorem of Calculus and Lemma 2.11. Finally, the combination of (2.26) and (2.27) yields

$$|f(t)| \leq C\left(\|\nabla u\|_{L^2([-1,2]\times S^1)} + \|\hat{\mathcal{T}}\|_{L^2([-1,2]\times S^1)}\right)^2,$$

for some new C, which is precisely (2.25). □

In the next lemma, we establish a differential inequality which is satisfied by the quantity

$$\int_{\{t\}\times S^1} |u_\theta|^2,$$

as t varies — i.e. as we move along the cylinder $I \times S^1$. Eventually this will be used to show that this quantity must decay exponentially as we move t towards the centre of the interval I from either end. This lemma resembles [5, Lemma 2.1] but without the $\sup|\nabla u|$ bound as a hypothesis.

Lemma 2.13. *For $\gamma \in (0,1]$ and $r \in (0, e^{-4}]$, let us suppose that $u : \Sigma \to S^2 \hookrightarrow \mathbb{R}^3$ is smooth (where $\Sigma := D_\gamma \backslash D_{\gamma r^2}$ as before) and write*

$$\mathcal{T} = \frac{1}{\varsigma^2}\left(\Delta u + u|\nabla u|^2\right)$$

on Σ, where $\varsigma : \Sigma \to [1,2]$ is smooth. Then there exists a universal constant $\delta > 0$ such that whenever

$$E(u,\Sigma) < \delta,$$

and

$$\|\mathcal{T}\varsigma\|^2_{L^2(\Sigma)} < \delta,$$

there holds the inequality

$$\frac{d^2}{dt^2}\int_{\{t\}\times S^1}|u_\theta|^2 \geq \int_{\{t\}\times S^1}|u_\theta|^2 - 12\int_{\{t\}\times S^1}|\mathcal{T}|^2\varsigma^2 e^{-4t}, \tag{2.28}$$

for all $t \in (-\ln\gamma + 1, -\ln(\gamma r^2) - 1]$.

Proof. Let us adopt the notation $C_t := \{t\} \times S^1$ throughout this proof. We begin by calculating

$$\frac{d}{dt}\int_{C_t}|u_\theta|^2 = 2\int_{C_t}\langle u_\theta, u_{\theta t}\rangle,$$

and hence, by definition of $\hat{\mathcal{T}}$,

$$\begin{aligned}\frac{d^2}{dt^2}\int_{C_t}|u_\theta|^2 &= 2\int_{C_t}(|u_{\theta t}|^2 + \langle u_\theta, u_{\theta tt}\rangle) = 2\int_{C_t}(|u_{\theta t}|^2 - \langle u_{\theta\theta}, u_{tt}\rangle) \\ &= 2\int_{C_t}|u_{\theta t}|^2 + 2\int_{C_t}|u_{\theta\theta}|^2 + 2\int_{C_t}\langle u_{\theta\theta}, u\rangle(|u_t|^2 + |u_\theta|^2) \\ &\quad -2\int_{C_t}\langle u_{\theta\theta}, \hat{\mathcal{T}}\rangle.\end{aligned} \tag{2.29}$$

The third term in the final expression here may be rewritten

$$\begin{aligned}&2\int_{C_t}\langle u_{\theta\theta}, u\rangle(|u_t|^2 + |u_\theta|^2) \\ &\quad = -2\int_{C_t}|u_\theta|^2(|u_t|^2 + |u_\theta|^2) - 2\int_{C_t}\langle u_\theta, u\rangle(2\langle u_t, u_{t\theta}\rangle + 2\langle u_\theta, u_{\theta\theta}\rangle),\end{aligned}$$

and therefore, since $\langle u_\theta, u\rangle = 0$, we have

$$\left|2\int_{C_t}\langle u_{\theta\theta}, u\rangle(|u_t|^2 + |u_\theta|^2)\right| \leq 2\int_{C_t}|u_\theta|^2(|u_t|^2 + |u_\theta|^2). \tag{2.30}$$

Now since $t \in (-\ln\gamma + 1, -\ln(\gamma r^2) - 1]$, we can find some interval $[s-1, s+2] \subset \overline{I}$ with $t \in [s, s+1]$, and by virtue of conformal invariance of energy, we have

$$\frac{1}{2}\|\nabla u\|^2_{L^2([s-1,s+2]\times S^1)} \leq \frac{1}{2}\|\nabla u\|^2_{L^2(I\times S^1)} = \frac{1}{2}\|\nabla u\|^2_{L^2(\Sigma)} < \delta. \tag{2.31}$$

Therefore, if we insist that $\delta < \delta_0$, we may apply Corollary 2.12 over the interval $[s-1, s+2]$ (instead of $[-1, 2]$) by translation of u, and deduce that

$$\begin{aligned}&\int_{C_t}|u_\theta|^2(|u_t|^2 + |u_\theta|^2) \\ &\quad \leq C\left(\int_{C_t}|u_{\theta\theta}|^2\right)\left(\|\nabla u\|_{L^2([s-1,s+2]\times S^1)} + \|\hat{\mathcal{T}}\|_{L^2([s-1,s+2]\times S^1)}\right)^2 \\ &\quad \leq C\left(\int_{C_t}|u_{\theta\theta}|^2\right)\delta \leq \frac{1}{3}\int_{C_t}|u_{\theta\theta}|^2,\end{aligned} \tag{2.32}$$

for some universal constants C — assuming that δ is chosen sufficiently small — where the final line uses both (2.31) and the inequality

$$\begin{aligned}\|\hat{\mathcal{T}}\|^2_{L^2([s-1,s+2]\times S^1)} &= \|\mathcal{T}\varsigma^2 e^{-2t}\|^2_{L^2([s-1,s+2]\times S^1)} \leq 2\|\mathcal{T}\varsigma e^{-t}\|^2_{L^2(I\times S^1)} \\ &= 2\|\mathcal{T}\varsigma\|^2_{L^2(\Sigma)} < 2\delta,\end{aligned}$$

which uses the facts that $\varsigma \leq 2$ and $t \geq 0$. Inequality (2.32) allows us to develop (2.30) to

$$\left|2\int_{C_t}\langle u_{\theta\theta}, u\rangle(|u_t|^2 + |u_\theta|^2)\right| \leq \frac{2}{3}\int_{C_t}|u_{\theta\theta}|^2. \tag{2.33}$$

We now set our sights on the final term of (2.29). The inequalities of Cauchy-Schwarz and Young, and the pointwise inequality $\varsigma \leq 2$ tell us that

$$\begin{aligned}\left|2\int_{C_t}\langle u_{\theta\theta}, \hat{\mathcal{T}}\rangle\right| &\leq \frac{1}{3}\int_{C_t}|u_{\theta\theta}|^2 + 3\int_{C_t}|\mathcal{T}|^2\varsigma^4 e^{-4t} \\ &\leq \frac{1}{3}\int_{C_t}|u_{\theta\theta}|^2 + 12\int_{C_t}|\mathcal{T}|^2\varsigma^2 e^{-4t}.\end{aligned} \tag{2.34}$$

It remains to apply the sum of (2.33) and (2.34) to (2.29) to yield

$$\begin{aligned}\frac{d^2}{dt^2}\int_{C_t}|u_\theta|^2 &\geq 2\int_{C_t}|u_{\theta t}|^2 + 2\int_{C_t}|u_{\theta\theta}|^2 - \left(\int_{C_t}|u_{\theta\theta}|^2 + 12\int_{C_t}|\mathcal{T}|^2\varsigma^2 e^{-4t}\right) \\ &\geq \int_{C_t}|u_{\theta\theta}|^2 - 12\int_{C_t}|\mathcal{T}|^2\varsigma^2 e^{-4t},\end{aligned}$$

which, after an application of the Poincaré inequality

$$\int_{C_t}|u_\theta|^2 \leq \int_{C_t}|u_{\theta\theta}|^2,$$

completes the proof. □

We will deal with the differential inequality of Lemma 2.13 by comparing it to an ordinary differential equation which we analyse in the following lemma.

Lemma 2.14. *Suppose* $\gamma \in (0,1]$ *and* $r \in (0, e^{-4}]$, *and*

$$T_1 \in (-\ln\gamma + 1, -\ln\gamma + 2]; \qquad T_2 \in (-\ln(\gamma r^2) - 2, -\ln(\gamma r^2) - 1].$$

Then if $H : [T_1, T_2] \to [0, \infty)$ *is smooth and satisfies*

$$\int_{T_1}^{T_2} H(t)dt \leq 12, \tag{2.35}$$

and we find a solution $f : [T_1, T_2] \to \mathbb{R}$ *of the* ODE

$$f''(t) - f(t) = -e^{-2t}H(t), \tag{2.36}$$

with given boundary values $f(T_1), f(T_2) \in [0,1]$, *then there exists a universal constant* C *such that*

$$f(t) \leq C\, r$$

for any $t \in (-\ln(\gamma r) - 2, -\ln(\gamma r) + 1]$.

Note that the constraint $r \in (0, e^{-4}]$ guarantees that $T_1 \leq t \leq T_2$. The bizarre hypothesis (2.35) on H arises since we will later set

$$H(t) := 12 \int_{\{t\}\times S^1} |\mathcal{T}|^2 \varsigma^2 e^{-2t},$$

and the constraint on the tension $\|\mathcal{T}\varsigma\|_{L^2(D_\gamma)} \leq 1$ from Lemma 2.9 then will complete the estimate

$$\int_{T_1}^{T_2} H(t)dt \leq 12 \int_{I\times S^1} |\mathcal{T}|^2 \varsigma^2 e^{-2t} = 12 \|\mathcal{T}\varsigma\|^2_{L^2(\Sigma)} \leq 12.$$

Proof. The solution of the ODE (2.36) may be given explicitly as

$$f(t) = Ae^t + Be^{-t} - \frac{1}{2}\int_t^{T_2} H(s)(e^{-s-t} - e^{t-3s})ds,$$

where

$$A := \frac{e^{T_2} f(T_2) - e^{T_1} f(T_1) - Q}{e^{2T_2} - e^{2T_1}},$$

and

$$B := \frac{e^{T_1+2T_2} f(T_1) - e^{2T_1+T_2} f(T_2) + e^{2T_2} Q}{e^{2T_2} - e^{2T_1}},$$

in which

$$Q := \frac{1}{2}\int_{T_1}^{T_2} H(s)e^{-s}(1 - e^{2(T_1-s)})ds \geq 0.$$

We now proceed to estimate the constants A, B and Q. We begin with the nonnegative quantity Q which satisfies

$$Q \leq \frac{1}{2}\int_{T_1}^{T_2} H(s)ds \leq 6,$$

by hypothesis (2.35). We will use this to handle A and B, together with the hypotheses on T_1 and T_2 which we may write as

$$\frac{e}{\gamma} < e^{T_1} \leq \frac{e^2}{\gamma}; \qquad \frac{1}{\gamma r^2 e^2} < e^{T_2} \leq \frac{1}{\gamma r^2 e},$$

and the restrictions

$$\gamma \in (0,1]; \qquad r \in (0, e^{-4}].$$

Thus we have

$$e^{2T_2} - e^{2T_1} \geq \left(\frac{1}{\gamma r^2 e^2}\right)^2 - \left(\frac{e^2}{\gamma}\right)^2 \geq \frac{1}{\gamma^2 r^4 e^4} - \frac{1}{2\gamma^2 r^4 e^4} = \frac{1}{2\gamma^2 r^4 e^4},$$

and hence

$$|A| \leq \frac{e^{T_2} + e^{T_1} + Q}{e^{2T_2} - e^{2T_1}} \leq \frac{\frac{1}{\gamma r^2 e} + \frac{e^2}{\gamma} + 6}{\frac{1}{2\gamma^2 r^4 e^4}} \leq C \left(\frac{\frac{1}{\gamma r^2}}{\frac{1}{\gamma^2 r^4}} \right) = C\gamma r^2,$$

and

$$\begin{aligned} |B| &\leq \frac{e^{T_1+2T_2} + e^{2T_1+T_2} + e^{2T_2}Q}{e^{2T_2} - e^{2T_1}} \\ &\leq \frac{\frac{e^2}{\gamma}\left(\frac{1}{\gamma r^2 e}\right)^2 + \left(\frac{e^2}{\gamma}\right)^2 \frac{1}{\gamma r^2 e} + \left(\frac{1}{\gamma r^2 e}\right)^2 6}{\frac{1}{2\gamma^2 r^4 e^4}} \leq C \left(\frac{\frac{1}{\gamma^3 r^4}}{\frac{1}{\gamma^2 r^4}} \right) = \frac{C}{\gamma}, \end{aligned}$$

for some universal constants C. Finally, we use the hypothesis

$$\frac{1}{\gamma r e^2} < e^t \leq \frac{e}{\gamma r},$$

to estimate

$$\begin{aligned} \left| \frac{1}{2} \int_t^{T_2} H(s)(e^{-s-t} - e^{t-3s}) ds \right| &= \frac{1}{2} e^{-t} \int_t^{T_2} H(s) e^{-s}(1 - e^{-2(s-t)}) ds \\ &\leq \frac{1}{2} e^{-t} \int_t^{T_2} H(s) ds \leq 6e^{-t}, \end{aligned}$$

and conclude that

$$|f(t)| \leq |A|e^t + |B|e^{-t} + 6e^{-t} \leq C \left((\gamma r^2) \frac{1}{\gamma r} + \frac{1}{\gamma}(\gamma r) + \gamma r \right) \leq Cr. \qquad \square$$

We have now compiled enough machinery to prove Lemma 2.9.

Proof. We will choose δ to be no more than the δ of Lemma 2.13 or the δ_0 of Lemma 2.11, and to lie within $(0, \frac{1}{2}]$. Now since

$$\int_{[-\ln\gamma+1, -\ln\gamma+2]\times S^1} |u_\theta|^2 \leq \|\nabla u\|^2_{L^2(\Sigma)} \leq 2\delta \leq 1,$$

we must be able to find $T_1 \in [-\ln\gamma + 1, -\ln\gamma + 2]$ such that

$$\int_{\{T_1\}\times S^1} |u_\theta|^2 \leq 1.$$

Similarly, we may find $T_2 \in [-\ln(\gamma r^2) - 2, -\ln(\gamma r^2) - 1]$ such that

$$\int_{\{T_2\}\times S^1} |u_\theta|^2 \leq 1.$$

Therefore we may fix boundary conditions

$$f(T_i) = \int_{\{T_i\}\times S^1} |u_\theta|^2,$$

for $i = 1, 2$, and invoke Lemma 2.14 with

$$H(t) := 12 \int_{\{t\}\times S^1} |\mathcal{T}|^2 \varsigma^2 e^{-2t},$$

to find that the solution f of the ODE (2.36) will satisfy $f(t) \leq Cr$ for $t \in (-\ln(\gamma r)-2, -\ln(\gamma r)+1]$. We are now in a position to compare the solution f, to the function

$$t \to \int_{\{t\}\times S^1} |u_\theta|^2,$$

since we know this function obeys the differential inequality (2.28) by virtue of Lemma 2.13. Indeed the maximum principle then tells us that

$$\int_{\{t\}\times S^1} |u_\theta|^2 \leq Cr,$$

for $t \in (-\ln(\gamma r) - 2, -\ln(\gamma r) + 1]$, and we may integrate over t to obtain

$$\int_{[-\ln(\gamma r)-2,-\ln(\gamma r)+1]\times S^1} |u_\theta|^2 \leq Cr, \tag{2.37}$$

for some new universal constant C.

Now we have control of the 'angular' energy, but are missing an estimate for the 'radial' energy. The bridge between the two is the Hopf differential (see §2.1.4) since

$$|u_t|^2 \leq |u_\theta|^2 + \left||u_t|^2 - |u_\theta|^2 - 2i\langle u_t, u_\theta\rangle\right| = |u_\theta|^2 + \psi(t,\theta).$$

Exploiting the conformal invariance of both the energy and the L^1 integral of ψ, we may calculate

$$\begin{aligned}\|\nabla u\|^2_{L^2(D_{\gamma re^2}\backslash D_{\gamma re^{-1}})} &= \int_{[-\ln(\gamma r)-2,-\ln(\gamma r)+1]\times S^1} (|u_\theta|^2 + |u_t|^2)\\ &\leq \int_{[-\ln(\gamma r)-2,-\ln(\gamma r)+1]\times S^1} (2|u_\theta|^2 + \psi(t,\theta))\\ &= 2\left(\int_{[-\ln(\gamma r)-2,-\ln(\gamma r)+1]\times S^1} |u_\theta|^2\right) + \|\psi\|_{L^1(D_{\gamma re^2}\backslash D_{\gamma re^{-1}})}.\end{aligned}$$

Then by invoking (2.37) and Lemma 2.7 we may continue to

$$\|\nabla u\|^2_{L^2(D_{\gamma re^2}\backslash D_{\gamma re^{-1}})} \leq C\left(r + \frac{1}{\gamma}|D_{\gamma re^2}\backslash D_{\gamma re^{-1}}|^{\frac{1}{2}}\right) \leq Cr, \tag{2.38}$$

where the constants C are dependent only on the energy bound M. Since $\Omega \subset D_{\gamma re^2}\backslash D_{\gamma re^{-1}}$, we have established estimate (2.21) of Lemma 2.9.

The reason we have sought an energy estimate over a region larger than Ω is that it will help with the oscillation estimate (2.22). Indeed, estimate (2.38) is equivalent to

$$\|\nabla u\|_{L^2([-\ln(\gamma r)-2,-\ln(\gamma r)+1]\times S^1)} \leq Cr^{\frac{1}{2}},$$

and so we may apply Lemma 2.11 to yield

$$\|u-\bar{u}\|_{W^{2,2}([-\ln(\gamma r)-1,-\ln(\gamma r)]\times S^1)} \leq C\left(r^{\frac{1}{2}}+\|\mathcal{T}\varsigma^2 e^{-2t}\|_{L^2([-\ln(\gamma r)-2,-\ln(\gamma r)+1]\times S^1)}\right). \tag{2.39}$$

Concerning the tension term, we have, over the domain of integration, the bound $t\geq -\ln(\gamma r)-2$, or equivalently $e^{-t}\leq \gamma r e^2\leq e^2 r$, and the usual bound $\varsigma\leq 2$, which together tell us that

$$\begin{aligned}\|\mathcal{T}\varsigma^2 e^{-2t}\|_{L^2([-\ln(\gamma r)-2,-\ln(\gamma r)+1]\times S^1)} &\leq Cr\|\mathcal{T}\varsigma e^{-t}\|_{L^2([-\ln(\gamma r)-2,-\ln(\gamma r)+1]\times S^1)}\\ &\leq Cr\|\mathcal{T}\varsigma\|_{L^2(D_\gamma)}\leq Cr\leq Cr^{\frac{1}{2}},\end{aligned}$$

by hypothesis (2.19) for universal constants C. Applying this to estimate (2.39) and exploiting the continuous embedding of $W^{2,2}$ into L^∞, we arrive at the estimate

$$\begin{aligned}\|u-\bar{u}\|_{L^\infty(\Omega)} &= \|u-\bar{u}\|_{L^\infty([-\ln(\gamma r)-1,-\ln(\gamma r)]\times S^1)}\\ &\leq C\|u-\bar{u}\|_{W^{2,2}([-\ln(\gamma r)-1,-\ln(\gamma r)]\times S^1)}\leq Cr^{\frac{1}{2}},\end{aligned}$$

for C dependent only on M, which is equivalent to (2.22). □

2.3. *Consequences of Theorem* 1.1. There are a number of convergence statements for the energies E, $E_{\bar\partial}$ and E_∂ which follow from Theorem 1.1 when we accept the hypotheses of Theorem 1.2. Since we shall need them from the next section onwards, we compile them now into the following lemma.

LEMMA 2.15. *Suppose that $u_n : S^2\to S^2$ is a sequence of maps satisfying the hypotheses of Theorem* 1.2. *Let us denote the set of antiholomorphic bubble points by $\mathcal{A}$, and the holomorphic bubble points by $\mathcal{H}$, and write Ξ for the set of all bubble maps ω (at all points). Then as $n\to\infty$,*

(i)
$$E(u_n)\to E(u_\infty)+\sum_{\omega\in\Xi}E(\omega),$$

(ii) *if $\Omega\subset\subset S^2\backslash\mathcal{A}$ then $E_{\bar\partial}(u_n,\Omega)\to 0$,*

(iii) *if $\Omega\subset\subset S^2\backslash\mathcal{H}$ then $E_\partial(u_n,\Omega)\to E_\partial(u_\infty,\Omega)$.*

Focusing on one bubble point, and adopting the notation of Theorem 1.1 (*with $\mu>0$ permitted to be any positive number for which D_μ contains only the one bubble point*) *we have that*

(iv)
$$\lim_{\nu\downarrow 0}\lim_{n\to\infty}E_\partial(u_n,D_\nu)=\sum_{i=1}^{k}E_\partial(\omega^i),$$

(v)

$$\lim_{\nu\downarrow 0}\lim_{n\to\infty} E_{\bar\partial}(u_n, D_\nu) = \sum_{i=1}^{k} E_{\bar\partial}(\omega^i),$$

(vi) $E_{\bar\partial}\left(u_n, D_\mu \backslash \bigcup_i D_{a_n^i,(\lambda_n^i)^{1/2}}\right) \to 0$ *as* $n\to\infty$,

(vii) $E\left(u_n, D_\mu \backslash \bigcup_i D_{a_n^i,(\lambda_n^i)^{1/2}}\right) \to E(u_\infty, D_\mu)$ *as* $n\to\infty$,

(viii) *if we are analysing an* anti*holomorphic bubble point, then*

$$(u_n)\times(u_n)_x + (u_n)_y \to (u_\infty)\times(u_\infty)_x + (u_\infty)_y$$

in $L^2(D_\mu)$ *as* $n\to\infty$.

Finally, if we fix a bubble point, and fix i *between* 1 *and* k, *then with the notation of part* (iv) *of Theorem* 1.1, *if* $\Omega \subset\subset \mathbb{R}^2\backslash\mathcal{S}$ *and we write* $v_n(x) := u_n(a_n^i + \lambda_n^i x)$ *and abbreviate* ω^i *to* ω, *then*

(ix)

$$|v_n \times (v_n)_x - (v_n)_y|^2 \to |\omega\times\omega_x - \omega_y|^2$$

in $L^1(\Omega)$ *as* $n\to\infty$.

Proof. All of these results follow fairly easily from Theorem 1.1, and we will only sketch the details here. Part (i) of the lemma follows almost immediately from parts (b) and (ii) of Theorem 1.1. In this regard, note that if $\Omega\subset\subset S^2\backslash\{x^1,\ldots,x^m\}$, then we have the convergence $u_n\to u_\infty$ in $W^{2,2}(\Omega)$ as $n\to\infty$, and so $E(u_n,\Omega)\to E(u_\infty,\Omega)$.

In fact, this $W^{2,2}$ convergence (which imples $W^{1,2}$ and L^∞ convergence) also tells us that

$$E_\partial(u_n,\Omega)\to E_\partial(u_\infty,\Omega), \tag{2.40}$$

and similarly $E_{\bar\partial}(u_n,\Omega)\to E_{\bar\partial}(u_\infty,\Omega) = 0$. To prove this, we may work in a small disc D_ν in a stereographic coordinate chart, which corresponds to a subset of Ω. Then we have

$$\begin{aligned}(u_n)\times(u_n)_x + (u_n)_y &- ((u_\infty)\times(u_\infty)_x + (u_\infty)_y)\\ &= (u_n - u_\infty)\times(u_n)_x + u_\infty\times(u_n-u_\infty)_x + (u_n-u_\infty)_y,\end{aligned}$$

and so

$$\begin{aligned}\|(u_n)\times(u_n)_x &+ (u_n)_y - ((u_\infty)\times(u_\infty)_x + (u_\infty)_y)\|_{L^2(D_\nu)}\\ &\le \|u_n - u_\infty\|_{L^\infty(D_\nu)}\|u_n\|_{W^{1,2}(D_\nu)} + 2\|\nabla(u_n-u_\infty)\|_{L^2(D_\nu)}\to 0,\end{aligned}$$

which implies (2.40).

A similar approach using the $W^{2,2}$ convergence of part (iv) of Theorem 1.1 is enough to establish part (ix) of the lemma. Parts (iv) and (v) also follow from this similar approach, coupled with part (ii) of Theorem 1.1. Once equipped with parts (iv) and (v), parts (ii) and (iii) are simple.

For the remaining parts (vi) and (vii), we can use part (iii) of Theorem 1.1. Here the radius $(\lambda_n^i)^{\frac{1}{2}}$ could be any radius r_n for which $r_n(\lambda_n^i)^{-1} \to \infty$ as $n \to \infty$. □

2.4. *Repulsive effects.* Our goal in this section is the bound (1.4) from Theorem 1.2.

2.4.1. *Lower bound for e_∂ off $\mathcal{T}$-small sets.* Given a nonconstant rational map v between 2-spheres, the ∂-energy density $e_\partial(v)$ can be zero only at finitely many points. If u is an almost-harmonic map which is 'close' to v over a region where $e_\partial(v) \neq 0$, then we might hope that $e_\partial(u)$ will also avoid zero. The next estimate will control the size of the set on which $e_\partial(u)$ can decay near to zero, when we take a notion of 'closeness' of u and v which arises naturally in applications.

Looking ahead, the main application of this estimate will be to control the size of any antiholomorphic bubbles which may be attached to a holomorphic body map in an almost harmonic map u. We will be able to argue eventually that an antiholomorphic bubble must lie within a region of the domain where $e_\partial(u)$ is small; otherwise we will have an overlapping region where the $\bar\partial$-energy is large and the ∂-energy density is not small, and thus a region where the Hopf differential is large in some sense which is prohibited by part (a) of Lemma 2.5.

LEMMA 2.16. *Suppose $\beta \in (0,1]$ and that $u : D_{2\beta} \to S^2 \hookrightarrow \mathbb{R}^3$ and $v : D_{2\beta} \to S^2 \hookrightarrow \mathbb{R}^3$ are smooth with*

$$0 = \Delta v + v|\nabla v|^2, \tag{2.41}$$

(*i.e. v harmonic*) *and*

$$\mathcal{T} := \frac{1}{\sigma^2}\left(\Delta u + u|\nabla u|^2\right),$$

as usual (*on $D_{2\beta}$*). *Suppose further that $e_\partial(v)$ is restricted by*

$$|v \times v_x + v_y| \geq 2\alpha,$$

for some $\alpha > 0$, on the whole of D_β.

Then there exist $\varepsilon = \varepsilon(\alpha,\beta) \in (0,1]$, $K = K(\alpha)$ and $\eta > 0$ universal such that if

(H1) $E_\partial(v, D_{2\beta}) := \frac{1}{4}\|v_x - v \times v_y\|^2_{L^2(D_{2\beta})} < \eta$,

(H2) $\|u - v\|_{L^2(D_{2\beta})} < \varepsilon$,

(H3) $\|\mathcal{T}\sigma\|_{L^2(D_{2\beta})} < \varepsilon$,

(H4) $\|(u \times u_x + u_y) - (v \times v_x + v_y)\|^2_{L^2(D_{2\beta})} < \varepsilon$, *all hold, then*

$$\mathrm{Area}_{\mathbb{R}^2}\{\mathbf{x}\in D_\beta\ :\ |u\times u_x+u_y|(\mathbf{x})\leq\alpha\}\leq\exp\left[-\frac{1}{K\|\mathcal{T}\sigma\|^2_{L^2(D_{2\beta})}}\right].$$

To see that this result is sharp, we may take, as a prototype example, the maps $u(z)=|z|^\delta z$ and $v(z)=z$, where we are using the complex coordinate z on $D_{2\beta}$, and a stereographic complex coordinate on the target.

The proof will involve deriving an integral expression for the difference between $e_\partial(u)$ and $e_\partial(v)$, and then invoking the borderline Riesz potential estimate of part (iii), Lemma 2.2.

Proof. For $q\in(1,2)$, let us use the shorthand $s=\frac{2q}{2-q}$ which then takes arbitrary values in $(2,\infty)$. Let us set $\varepsilon=\eta=1$ for the moment; during the course of the proof we will impose a finite number of positive upper bounds for ε (each with no more than a dependence on α and β) and for η (with no dependencies) and end up taking the minima of these. (An additional dependence on q — or a derived exponent — will be suppressed since these exponents will shortly be given explicit values.)

We begin by using $(H1)$ and $(H4)$ to estimate

$$\begin{aligned}E_\partial(u,D_{2\beta})&=\frac{1}{4}\|u\times u_x+u_y\|^2_{L^2(D_{2\beta})}\\&\leq\frac{1}{4}\left(\|v\times v_x+v_y\|_{L^2(D_{2\beta})}+\|(u\times u_x+u_y)-(v\times v_x+v_y)\|_{L^2(D_{2\beta})}\right)^2\\&\leq 2E_\partial(v,D_{2\beta})+\frac{1}{2}\|(u\times u_x+u_y)-(v\times v_x+v_y)\|^2_{L^2(D_{2\beta})}<2\eta+\frac{\varepsilon}{2}.\end{aligned}$$

By using also $(H3)$, we may apply part (a) of Lemma 2.3 to our situation — provided we insist that $\varepsilon\leq\varepsilon_0(s)$ and $\eta\leq\frac{\varepsilon_0}{4}$ so that $2\eta+\frac{\varepsilon}{2}\leq\varepsilon_0(s)$ — and deduce that

$$\|u_x-u\times u_y\|_{L^s(D_{\frac{3\beta}{2}})}<C=C(\beta,s).\tag{2.42}$$

A more direct application of the same part of the lemma — this time to v rather than u — gives us

$$\|v_x-v\times v_y\|_{L^s(D_{\frac{3\beta}{2}})}<C=C(\beta,s).\tag{2.43}$$

We use these two estimates first to get the universal bound

$$\|(u\times u_x+u_y)-(v\times v_x+v_y)\|_{L^s(D_{\frac{3\beta}{2}})}<C=C(\beta,s),$$

(with C changing all the time as usual) which we can now interpolate with the hypothesis $(H4)$. Indeed, given $\lambda\in(0,1)$, Hölder's inequality allows us to interpolate

$$\|f\|_{L^{\frac{2s}{s\lambda+2(1-\lambda)}}}\leq\|f^\lambda\|_{L^{\frac{2}{\lambda}}}\|f^{(1-\lambda)}\|_{L^{\frac{s}{1-\lambda}}}=\|f\|^\lambda_{L^2}\|f\|^{1-\lambda}_{L^s},$$

for any function f, and in our context — with $\lambda = \frac{1}{4}$ — this gives us

$$\|(u \times u_x + u_y) - (v \times v_x + v_y)\|_{L^{\frac{8s}{s+6}}(D_{\frac{3\beta}{2}})} \leq (\varepsilon^{\frac{1}{2}})^{\frac{1}{4}}(C(s,\beta))^{\frac{3}{4}}.$$

Using the shorthand $r = \frac{8s}{s+6}$, we deduce that for every $r \in (2,8)$, there exists $C = C(\beta, r)$ such that

$$\|(u \times u_x + u_y) - (v \times v_x + v_y)\|_{L^r(D_{\frac{3\beta}{2}})} \leq C\varepsilon^{\frac{1}{8}}. \tag{2.44}$$

One of our uses of (2.44) will be to control the quantity

$$\begin{aligned} &u|u_x - u \times u_y|^2 - v|v_x - v \times v_y|^2 \\ &\qquad = (u - v)|v_x - v \times v_y|^2 + u(|u_x - u \times u_y|^2 - |v_x - v \times v_y|^2), \end{aligned}$$

in L^p for $p > 2$. In particular, we calculate on $D_{\frac{3\beta}{2}}$

$$\begin{aligned} &\|u|u_x - u \times u_y|^2 - v|v_x - v \times v_y|^2\|_{L^3} \\ &\quad \leq \|(u - v)|v_x - v \times v_y|^2\|_{L^3} + \|u(|u_x - u \times u_y|^2 - |v_x - v \times v_y|^2)\|_{L^3} \\ &\quad \leq \|(u - v)\|_{L^6}\||v_x - v \times v_y|^2\|_{L^6} \\ &\qquad + \|(u \times u_x + u_y) - (v \times v_x + v_y)\|_{L^6}\|(u \times u_x + u_y) + (v \times v_x + v_y)\|_{L^6} \\ &\quad \leq C(\|(u - v)\|_{L^6} + \|(u \times u_x + u_y) - (v \times v_x + v_y)\|_{L^6}), \end{aligned}$$

(for $C = C(\beta)$) using (2.42) and (2.43), and since $|u - v| \leq 2$ pointwise, and we have the estimate (2.44), and hypothesis $(H2)$, this implies

$$\begin{aligned} \|u|u_x - u \times u_y|^2 - v|v_x - v \times v_y|^2\|_{L^3(D_{\frac{3\beta}{2}})} &\leq C\|(u - v)\|_{L^2}^{\frac{1}{3}} + C\varepsilon^{\frac{1}{8}} \\ &\leq C(\varepsilon^{\frac{1}{3}} + \varepsilon^{\frac{1}{8}}) \leq C\varepsilon^{\frac{1}{8}}, \end{aligned} \tag{2.45}$$

for constants $C = C(\beta)$.

Another use of (2.44) will be to control the L^3 norm of $|(u_x - u \times u_y) - (v_x - v \times v_y)|$. We stress that although we are frequently using facts such as $|u \times u_x + u_y| = |u_x - u \times u_y|$ implicitly, it is not necessarily true that $|(u_x - u \times u_y) - (v_x - v \times v_y)|$ is the same as $|(u \times u_x + u_y) - (v \times v_x + v_y)|$. Consequently we must observe that

$$(u_x - u \times u_y) - (v_x - v \times v_y) = (v - u) \times (u \times u_x + u_y) + v \times (v \times v_x + v_y - (u \times u_x + u_y)),$$

which allows us to estimate

$$\begin{aligned} &\|(u_x - u \times u_y) - (v_x - v \times v_y)\|_{L^3} \\ &\qquad \leq \|v - u\|_{L^6}\|u \times u_x + u_y\|_{L^6} + \|v \times v_x + v_y - (u \times u_x + u_y)\|_{L^3}. \end{aligned}$$

We can now mimick the strategy used to obtain (2.45) and deduce that

$$\|(u_x - u \times u_y) - (v_x - v \times v_y)\|_{L^3(D_{\frac{3\beta}{2}})} \leq C\varepsilon^{\frac{1}{8}}, \tag{2.46}$$

with C dependent on β.

We have now assembled enough estimates for $e_\partial(u)$, $e_\partial(v)$ and related quantities to be able to control the set where $e_\partial(u)-e_\partial(v)$ differs substantially from zero.

Let $\varphi : D_{\frac{3\beta}{2}} \to [0,1]$ be a smooth cut-off function, with compact support, and with the properties that $\varphi \equiv 1$ on D_β and $|\nabla\varphi| \leq \frac{4}{\beta}$. We use this cut-off in the second estimate of Lemma 2.1, applied both to u and to v. The latter is simpler in that there is no tension term since v is harmonic (see (2.41)). Subtracting these two formulae gives the expression

$$2\pi(u\times u_x + u_y - (v\times v_x + v_y))\varphi(a,b) = I + II + III, \tag{2.47}$$

where $(a,b)\in\mathbb{R}^2$ and

$$\begin{aligned}
I &:= -\int_{\mathbb{R}^2} \frac{1}{(x-a)^2+(y-b)^2}\,((y-b)\,\mathcal{T} + (x-a)\,u\times\mathcal{T})\,\sigma^2\varphi\,dx\wedge dy,\\
II &:= \int_{\mathbb{R}^2} \frac{(y-b)}{(x-a)^2+(y-b)^2}\,(u|u_x - u\times u_y|^2 - v|v_x - v\times v_y|^2)\varphi\,dx\wedge dy,\\
III &:= -\int_{\mathbb{R}^2} \frac{1}{(x-a)^2+(y-b)^2}\\
&\quad\cdot\Big(((x-a)\varphi_x + (y-b)\varphi_y)(u\times u_x + u_y - (v\times v_x + v_y))\\
&\quad -((x-a)\varphi_y - (y-b)\varphi_x)(u_x - u\times u_y - (v_x - v\times v_y))\Big)\,dx\wedge dy.
\end{aligned}$$

We now estimate each term I, II and III independently. The first term has the integral bound

$$|I| \leq 2\int_{D_{2\beta}} \frac{1}{|\mathbf{x}-\mathbf{a}|}|\mathcal{T}|\sigma\,dx\wedge dy, \tag{2.48}$$

where we have traded one power of σ for the factor 2 since $\sigma\leq 2$ pointwise. Meanwhile, the final two terms may be seen to be small uniformly in $\mathbf{a}$. We have, using Hölder's inequality,

$$\begin{aligned}
|II| &\leq \int_{D_{\frac{3\beta}{2}}} \frac{1}{|\mathbf{x}-\mathbf{a}|}\left|u|u_x - u\times u_y|^2 - v|v_x - v\times v_y|^2\right| dx\wedge dy,\\
&\leq \left(\int_{D_{\frac{3\beta}{2}}} \frac{dx\wedge dy}{|\mathbf{x}-\mathbf{a}|^{\frac{3}{2}}}\right)^{\frac{2}{3}} \|u|u_x - u\times u_y|^2 - v|v_x - v\times v_y|^2\|_{L^3(D_{\frac{3\beta}{2}})},
\end{aligned}$$

which is maximised when $\mathbf{a}=(0,0)$, and so we may use (2.45) to yield

$$|II| \leq C(\beta)\varepsilon^{\frac{1}{8}}. \tag{2.49}$$

We handle the third term in a similar manner, again using Hölder's inequality in the estimate

$$\begin{aligned}|III| \leq &\int_{D_{\frac{3\beta}{2}}} \frac{1}{|\mathbf{x}-\mathbf{a}|}\,|\nabla\varphi|(|u\times u_x+u_y-(v\times v_x+v_y)| \\ &+|u_x-u\times u_y-(v_x-v\times v_y)|)dx\wedge dy \\ \leq &\, C(\beta)\Big(\|u\times u_x+u_y-(v\times v_x+v_y)\|_{L^3} \\ &+\|u_x-u\times u_y-(v_x-v\times v_y)\|_{L^3}\Big),\end{aligned}$$

over $D_{\frac{3\beta}{2}}$. We may then bring (2.44) and (2.46) into play to give

$$|III| \leq C(\beta)\varepsilon^{\frac{1}{8}}. \tag{2.50}$$

Equipped with these estimates for I, II and III, we are in a position to return to (2.47) and find that

$$2\pi|u\times u_x+u_y-(v\times v_x+v_y)|\varphi(\mathbf{a}) \leq C(\beta)\varepsilon^{\frac{1}{8}}+2\int_{D_{2\beta}}\frac{1}{|\mathbf{x}-\mathbf{a}|}|\mathcal{T}|\sigma\,dx\wedge dy.$$

In particular, provided we insist on a sufficiently small ε (dependent on α and β) we have

$$|u\times u_x+u_y-(v\times v_x+v_y)|(\mathbf{a}) \leq \frac{\alpha}{2}+\frac{1}{\pi}\int_{D_{2\beta}}\frac{1}{|\mathbf{x}-\mathbf{a}|}|\mathcal{T}|\sigma\,dx\wedge dy, \tag{2.51}$$

for all $\mathbf{a}\in D_\beta$. We cannot hope to bound the right-hand side of this estimate uniformly in $\mathbf{a}$. However, after defining a nonnegative function g on D_2 by

$$g(\mathbf{a}) := \frac{1}{\pi}\int_{D_{2\beta}}\frac{1}{|\mathbf{x}-\mathbf{a}|}|\mathcal{T}|\sigma\,dx\wedge dy,$$

we may invoke part (iii) of Lemma 2.2 to give

$$\int_{D_2}\exp\left[\frac{g(\mathbf{x})}{C_1\|\mathcal{T}\sigma\|_{L^2(D_{2\beta})}}\right]^2 d\mathbf{x} \leq C_2,$$

for some universal constants C_1 and C_2. Therefore, we have control on the area

$$A := \text{Area}_{\mathbb{R}^2}\{\mathbf{x}\in D_\beta \;:\; |g(\mathbf{x})| \geq \frac{\alpha}{2}\},$$

given by the estimate

$$A\,\exp\left[\frac{\alpha}{2C_1\|\mathcal{T}\sigma\|_{L^2(D_{2\beta})}}\right]^2 \leq C_2,$$

or equivalently

$$A \leq C_2\exp\left[-\frac{2}{K\|\mathcal{T}\sigma\|^2_{L^2(D_{2\beta})}}\right],$$

for some constant $K = K(\alpha)$. But now, assuming that we choose ε sufficiently small so that

$$C_2 \exp\left[-\frac{1}{K\varepsilon^2}\right] \leq 1,$$

then by hypothesis $(H3)$ we can absorb the constant C_2 to get

$$A \leq \exp\left[-\frac{1}{K\|\mathcal{T}\sigma\|^2_{L^2(D_{2\beta})}}\right].$$

It remains to conclude using (2.51) that

$$\mathrm{Area}_{\mathbb{R}^2}\{\mathbf{x} \in D_\beta \;:\; |u \times u_x + u_y| \leq \alpha\} \leq A \leq \exp\left[-\frac{1}{K\|\mathcal{T}\sigma\|^2_{L^2(D_{2\beta})}}\right]. \qquad \square$$

2.4.2. *Bubble concentration estimates.* Equipped with the estimates of the previous section — which control the area in which the ∂-energy density can decay to near zero — and the Hopf differential estimates of Section 2.1.4, we may now proceed to quantify the repulsion that exists between antiholomorphic bubbles and holomorphic body maps, and estimate the level of concentration of such a bubble in terms of the tension.

As mentioned briefly in the previous section, our strategy is to argue that the antiholomorphic bubbles must lie essentially within a set on which the ∂-energy density is small (the size of which we can control) since otherwise the product $e_\partial(u)e_{\bar\partial}(u)$ will be larger than is allowed by part (a) of Lemma 2.5.

The following lemma resolves (1.4) of Theorem 1.2.

LEMMA 2.17. *Suppose u_n is a sequence satisfying the hypotheses of Theorem* 1.2. *Then at any antiholomorphic bubble point x^j, there exist constants $C > 0$ and $N \in \mathbb{N}$ such that, using the notation of Theorem* 1.1, *we have*

$$\lambda_n^i \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right], \tag{2.52}$$

for all $i \in \{1, \ldots, k\}$ and $n \geq N$.

Proof. According to the hypotheses of Theorem 1.2, we already have a stereographic coordinate chart chosen around x^j, and have a sequence $a_n^i \to 0 \in \mathbb{R}^2$ about which we rescale to extract the bubble.

Let us choose $\beta \in (0,1)$ sufficiently small so that the disc $D_{2\beta}$ contains no holomorphic bubble points and no points where $|\nabla u_\infty| = 0$ (this is possible by the hypotheses of Theorem 1.2) and also so that

$$E_\partial(u_\infty, D_{2\beta}) < \eta,$$

where η is taken to be as in Lemma 2.16. (Recall that u_∞ is assumed to be holomorphic.) By our hypothesis that u_n is as in Theorem 1.2 (and therefore with the hypotheses and convergence of Theorem 1.1) we know that $\mathcal{T}(u_n) \to 0$ in $L^2(S^2)$, and hence that $\|\mathcal{T}(u_n)\sigma\|_{L^2(D_{2\beta})} \to 0$ as $n \to \infty$, and also that $u_n \to u_\infty$ in $L^2(D_{2\beta})$ since weak convergence in $W^{1,2}$ implies strong convergence in L^2. Moreover, we recall the convergence

$$\| (u_n \times (u_n)_x + (u_n)_y) - (u_\infty \times (u_\infty)_x + (u_\infty)_y) \|_{L^2(D_{2\beta})} \to 0,$$

from part (viii) of Lemma 2.15.

By virtue of these convergence statements, we may apply — for sufficiently large n — Lemma 2.16 with $u = u_n$ and $v = u_\infty$, which tells us that

(2.53)
$$\begin{aligned}\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x} \in D_\beta \ : \ 2\sigma e_\partial(u_n)^{\frac{1}{2}}(\mathbf{x}) \le \alpha\right\} &\le \exp\left[-\frac{1}{K\|\mathcal{T}(u_n)\sigma\|^2_{L^2(D_{2\beta})}}\right] \\ &\le \exp\left[-\frac{1}{K\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right],\end{aligned}$$

where K depends on α, and

$$\alpha := \frac{1}{2}\min_{D_\beta}\left(2\sigma e_\partial(u_\infty)^{\frac{1}{2}}\right)$$

is positive since $4\sigma^2 e_\partial(u_\infty) = 4\sigma^2 e(u_\infty) = 2|\nabla u_\infty|^2 \neq 0$ on $D_{2\beta}$ by our hypothesis on β above.

Let us focus on one particular bubble — i.e. fix $i \in \{1, \dots, k\}$ — and adopt the abbreviations $a_n = a_n^i$, $\lambda_n = \lambda_n^i$ and $\omega = \omega^i$. Since $E(\omega, \mathbb{R}^2) \ge 4\pi$ we may choose $r > 0$ sufficiently large so that $E_{\bar\partial}(\omega, D_r) = E(\omega, D_r) \ge 2\pi$. (We recall that ω is antiholomorphic.) Now we choose $\delta > 0$ sufficiently small so that

$$\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x} \in D_r \ : \ 4\sigma^2 e_{\bar\partial}(\omega) > 2\delta\right\} \ge 3\delta.$$

By virtue of part (ix) of Lemma 2.15 we know that

$$\|4\sigma^2 e_{\bar\partial}(v_n) - 4\sigma^2 e_{\bar\partial}(\omega)\|_{L^1(\Omega)} \to 0,$$

where $v_n(x) := u_n(a_n + \lambda_n x)$, for some $\Omega \subset D_r$ with $\mathrm{Area}_{\mathbb{R}^2}(D_r \backslash \Omega) < \delta$. Therefore, since

$$\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x} \in \Omega \ : \ 4\sigma^2 e_{\bar\partial}(\omega) > 2\delta\right\} \ge 2\delta,$$

we have, for sufficiently large n, that

$$\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x} \in \Omega \ : \ 4\sigma^2 e_{\bar\partial}(v_n) > \delta\right\} \ge \delta,$$

and in particular that

$$\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x} \in D_r \ : \ 4\sigma^2 e_{\bar\partial}(v_n) > \delta\right\} \ge \delta.$$

The equivalent statement for the unscaled u_n is

$$\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x}\in D_{a_n,\lambda_n r}\ :\ 4\sigma^2 e_{\bar\partial}(u_n) > \frac{\delta}{\lambda_n^2}\right\} \geq \delta\lambda_n^2,$$

and so provided n is large enough to ensure that $D_{a_n,\lambda_n r}\subset D_\beta$, we have that

$$\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x}\in D_{\beta}\ :\ 4\sigma^2 e_{\bar\partial}(u_n) > \frac{\delta}{\lambda_n^2}\right\} \geq \delta\lambda_n^2. \tag{2.54}$$

Now suppose that we may pass to a subsequence in n so that

$$\delta\lambda_n^2 \geq 2\exp\left[-\frac{1}{K\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right], \tag{2.55}$$

for all n. Then by (2.53) and (2.54) we have, for sufficiently large n, that

$$\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x}\in D_{\beta}\ :\ 4\sigma^2 e_{\bar\partial}(u_n) > \frac{\delta}{\lambda_n^2}\quad \text{and}\quad 4\sigma^2 e_{\partial}(u_n) > \alpha^2\right\} \geq \frac{\delta\lambda_n^2}{2}.$$

In particular, writing $\psi_n := 4\sigma^2(e_\partial(u_n)e_{\bar\partial}(u_n))^{\frac{1}{2}}$ to correspond with the ψ of Section 2.1.4, this implies that

$$\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x}\in D_{\beta}\ :\ \psi_n^2(\mathbf{x}) > \frac{\delta\alpha^2}{\lambda_n^2}\right\} \geq \frac{\delta\lambda_n^2}{2}, \tag{2.56}$$

for sufficiently large n. However, by part (a) of Lemma 2.5 we know that

$$\begin{aligned}\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x}\in D_{\beta}\ :\ \psi_n^2(\mathbf{x}) > \frac{\delta\alpha^2}{\lambda_n^2}\right\} &\leq \frac{C}{(\frac{\delta\alpha^2}{\lambda_n^2})}M\|\mathcal{T}(u_n)\|^2_{L^2(S^2)} \\ &= \frac{CM}{\delta\alpha^2}\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}\,\lambda_n^2,\end{aligned} \tag{2.57}$$

for some universal C, and since $\|\mathcal{T}(u_n)\|_{L^2(S^2)}\to 0$ as $n\to\infty$, we see that for large enough n, the statements (2.56) and (2.57) will contradict each other.

Therefore we cannot have (2.55) and we must, for large enough n, have

$$\lambda_n^2 < \frac{2}{\delta}\exp\left[-\frac{1}{K\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right].$$

Assuming we take n sufficiently large so that

$$\frac{2}{\delta}\exp\left[-\frac{1}{2K\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right] \leq 1,$$

we then have

$$\lambda_n^2 < \exp\left[-\frac{1}{2K\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right],$$

which completes the proof. □

2.5. *Quantization effects.* Our goal in this section is to apply the repulsion estimates of the previous section to prove the quantization estimate (1.3) of Theorem 1.2.

2.5.1. *Control of* $e_{\bar\partial}$. In this section, we prove two lemmata controlling the size of $e_{\bar\partial}(u_n)$ in terms of the tension, for a sequence u_n satisfying the hypotheses of Theorem 1.2. The first of these will control the L^q norm of $e_{\bar\partial}(u_n)^{\frac{1}{2}}$, for $q \in (1,2)$, and will apply over any compact subdomain of S^2 not containing any holomorphic bubble points or points where $|\nabla u_\infty| = 0$. The second lemma will improve this to L^2 control of $e_{\bar\partial}(u_n)^{\frac{1}{2}}$ — or equivalently L^1 control of $e_{\bar\partial}(u_n)$ — but now over any compact subdomain of S^2 not containing any *anti*holomorphic bubble points.

Recall that in Theorem 1.2 we have assumed, without loss of generality, that u_∞ is holomorphic and hence that $e_{\bar\partial}(u_\infty) \equiv 0$. Therefore, since u_n converges to u_∞ in some sense, we expect $e_{\bar\partial}(u_n)$ to decay when measured in an appropriate way. Despite this, we note that we cannot expect to control $e_{\bar\partial}(u_n)$ in L^1 over any antiholomorphic bubble points since the antiholomorphic bubbles hold at least 4π of $\bar\partial$-energy.

As we shall see, having control of the $\bar\partial$-energy on most of the domain S^2 is part of showing that the energy on the body component of the bubble tree is becoming quantized.

LEMMA 2.18. *Suppose u_n is a sequence satisfying the hypotheses of Theorem* 1.2. *Then given any $q \in [1,2)$, and any compact subset Ω of the domain S^2 which contains no holomorphic bubble points and no points where $|\nabla u_\infty| = 0$, there exist constants C and N such that*

$$\|e_{\bar\partial}(u_n)^{\frac{1}{2}}\|_{L^q(\Omega)} \leq C\|\mathcal{T}(u_n)\|_{L^2(S^2)},$$

for all $n \geq N$.

The strategy of the proof of Lemma 2.18 will be to control the ∂-energy density from below — except on a set of controlled size — and then invoke the Hopf differential estimate from part (b) of Lemma 2.5. The difficulty will lie in controlling the size of the exceptional set where the ∂-energy density is small; for this we shall require Lemma 2.16.

LEMMA 2.19. *Suppose u_n is a sequence satisfying the hypotheses of Theorem* 1.2. *Then given any compact subset Ω of the domain S^2 which contains no* anti*holomorphic bubble points, there exist constants C and N such that*

$$E_{\bar\partial}(u_n, \Omega) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}, \tag{2.58}$$

for all $n \geq N$.

We stress that although we are making a $\bar{\partial}$-energy estimate over a *sub*domain of S^2, there is no 'boundary' term in (2.58). If this were permitted, we would proceed in a more direct fashion, in the spirit of Lemma 2.3.

In order to prove Lemma 2.19, we will apply Lemma 2.18 where it is applicable and then bootstrap the estimate with part (c) of Lemma 2.3. The bootstrapping procedure requires a certain smallness of $\bar{\partial}$-energy over the region being considered, and this will fail across antiholomorphic bubble points where $\bar{\partial}$-energy concentrates in the sequence $\{u_n\}$.

We are not concerned about the dependencies of the constants C and N in the lemmata above, other than that they are independent of n. In particular, they may depend on the sequence u_n.

Proof of Lemma 2.18. We begin by picking a compact subset $\hat{\Omega}$ of the domain S^2 whose interior contains Ω, and yet which still contains no holomorphic bubble points and no points where $|\nabla u_\infty| = 0$. These conditions guarantee that

$$\delta := \frac{1}{16}\min_{\hat{\Omega}} e_\partial(u_\infty)$$

will be positive.

Suppose now that we take some geodesic disc in the domain S^2 which lies within $\hat{\Omega}$, and is no larger than a hemisphere. This disc corresponds to a disc $D_{2\beta}$ in a stereographic chart, with $\beta \in (0, \frac{1}{2}]$. Suppose further that β is sufficiently small that $E_\partial(u_\infty, D_{2\beta}) < \eta$, with η taken from Lemma 2.16.

Since u_n satisfies the hypotheses of Theorem 1.2, and therefore satisfies the hypotheses and convergence of Theorem 1.1, we know that $\mathcal{T}(u_n) \to 0$ in $L^2(S^2)$ as $n \to \infty$, and also that $u_n \to u_\infty$ in $L^2(D_{2\beta})$ since weak convergence in $W^{1,2}$ implies strong convergence in L^2. Moreover, part (viii) of Lemma 2.15 equipped us with the convergence

$$\| (u_n \times (u_n)_x + (u_n)_y) - (u_\infty \times (u_\infty)_x + (u_\infty)_y) \|_{L^2(D_{2\beta})} \to 0.$$

Finally, we observe that since $\sigma \geq 1$ (the disc $D_{2\beta}$ lies within the unit disc D) we must — by definition of δ — have $4\sigma^2 e_\partial(u_\infty) \geq 64\,\delta$, or equivalently $2\sigma e_\partial(u_\infty)^{\frac{1}{2}} \geq 8\delta^{\frac{1}{2}}$; we may then invoke Lemma 2.16 (for sufficiently large n) with $u = u_n$ and $v = u_\infty$, to find that

$$\mathrm{Area}_{\mathbb{R}^2}\left\{\mathbf{x} \in D_\beta \ : \ 2\sigma e_\partial(u_n)^{\frac{1}{2}} \leq 4\delta^{\frac{1}{2}}\right\} \leq \exp\left[-\frac{1}{K\|\mathcal{T}(u_n)\sigma\|^2_{L^2(D_{2\beta})}}\right].$$

Using the facts that $\sigma \leq 2$, and $\|\mathcal{T}(u_n)\sigma\|_{L^2(D_{2\beta})} \leq \|\mathcal{T}(u_n)\|_{L^2(S^2)}$, we then have

$$\mathrm{Area}_{\mathbb{R}^2}\{\mathbf{x} \in D_\beta \ : \ e_\partial(u_n) \leq \delta\} \leq \exp\left[-\frac{1}{K\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right].$$

Switching to measuring the area of sets with the S^2 metric $\sigma^2(dx^2+dy^2)$ rather than the $\mathbb{R}^2$ metric (dx^2+dy^2) may then cost us a factor of no more than four (since $\sigma \leq 2$) and thus we find that

$$\text{Area}_{S^2}\{\mathbf{x} \in D_\beta \ : \ e_\partial(u_n) \leq \delta\} \leq 4 \exp\left[-\frac{1}{K\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right]. \tag{2.59}$$

We now wish to patch together local estimates of the form (2.59) into an estimate over Ω. By the compactness of Ω, we may find a finite number m discs in S^2 corresponding to discs D_β above (for varying stereographic coordinate charts) which cover the set Ω, and with the property that the union of the discs $D_{2\beta}$ lie within the slightly larger set $\hat{\Omega}$. Adding together the estimates of the form (2.59) which hold for each disc D_β, we then have

$$\text{Area}_{S^2}\{\mathbf{x} \in \Omega \ : \ e_\partial(u_n) \leq \delta\} \leq 4m \exp\left[-\frac{1}{K\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right].$$

Since $\mathcal{T}(u_n) \to 0$ in $L^2(S^2)$, this then implies, for sufficiently large n, that

$$\text{Area}_{S^2}\{\mathbf{x} \in \Omega \ : \ e_\partial(u_n) \leq \delta\} \leq \left(\|\mathcal{T}(u_n)\|_{L^2(S^2)}\right)^{\frac{2q}{2-q}}, \tag{2.60}$$

with $q \in [1,2)$ as in the lemma.

It is now time to invoke part (b) of Lemma 2.5 which tells us, for each n, that

$$\|(e_\partial(u_n)e_{\bar\partial}(u_n))^{\frac{1}{2}}\|_{L^q(S^2)} \leq C\|\mathcal{T}(u_n)\|_{L^2(S^2)},$$

for some constant $C = C(q, M)$. By splitting Ω into

$$\Omega_n^+ := \{\mathbf{x} \in \Omega \ : \ e_\partial(u_n) > \delta\}$$

and

$$\Omega_n^- := \{\mathbf{x} \in \Omega \ : \ e_\partial(u_n) \leq \delta\}$$

we may then calculate

$$\begin{aligned}
\|e_{\bar\partial}(u_n)^{\frac{1}{2}}\|_{L^q(\Omega)} &\leq \|(\delta^{-1}e_\partial(u_n)e_{\bar\partial}(u_n))^{\frac{1}{2}}\|_{L^q(\Omega_n^+)} + \|e_{\bar\partial}(u_n)^{\frac{1}{2}}\|_{L^q(\Omega_n^-)} \\
&\leq \delta^{-\frac{1}{2}}\|(e_\partial(u_n)e_{\bar\partial}(u_n))^{\frac{1}{2}}\|_{L^q(S^2)} + \|e_{\bar\partial}(u_n)^{\frac{1}{2}}\|_{L^2(\Omega_n^-)}\|1\|_{L^{\frac{2q}{2-q}}(\Omega_n^-)} \\
&\leq \delta^{-\frac{1}{2}}C\|\mathcal{T}(u_n)\|_{L^2(S^2)} + M^{\frac{1}{2}}(\text{Area}_{S^2}(\Omega_n^-))^{\frac{2-q}{2q}},
\end{aligned}$$

where we have used Hölder's inequality on the second term, and $C = C(q, M)$. Applying (2.60) which estimates the size of the set Ω_n^-, and allowing C to carry dependency on δ (i.e. on Ω and u_∞) we may then conclude that

$$\|e_{\bar\partial}(u_n)^{\frac{1}{2}}\|_{L^q(\Omega)} \leq C\|\mathcal{T}(u_n)\|_{L^2(S^2)}. \qquad \square$$

Proof of Lemma 2.19. We must first handle the possibility that u_∞ is a constant map. In this case, we have $|\nabla u_\infty| = 0$ at all points in the domain, and therefore by the hypotheses on antiholomorphic bubbles imposed in Theorem 1.2 we find that all bubbles must necessarily be holomorphic. By part (ii) of Lemma 2.15 this implies that $E_{\bar\partial}(u_n) \to 0$ as $n \to \infty$, and we may apply our estimate from [15, Lemma 1] (which holds for $\bar\partial$-energies as well as ∂-energies by composing the map with an orientation reversing isometry of the target) to deduce that

$$E_{\bar\partial}(u_n) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)},$$

for sufficiently large n and some universal constant C. (Note that in the language of the present paper, this estimate follows in a similar manner to Lemma 2.3 albeit with a more global approach.) In particular, the proof is complete in this case.

We now deal with the case that u_∞ is not a constant map, and in particular we now have $|\nabla u_\infty| \neq 0$ except at isolated points.

Let us pick a compact subset $\hat\Omega$ of the domain S^2 whose interior contains Ω, and yet which still contains no antiholomorphic bubble points. By part (ii) of Lemma 2.15, we know that

$$E_{\bar\partial}(u_n, \hat\Omega) \to 0 \tag{2.61}$$

as $n \to \infty$. Suppose that $a \in \Omega$. Our first task is to establish the estimate of the lemma in a small neighbourhood of a. Let us stereographically project about a and consider a disc $D_{2\mu}$ (representing a neighbourhood of a) with $\mu \in (0, \frac{1}{2}]$ sufficiently small so that $\overline{D_{2\mu}}$ corresponds to a region in $\hat\Omega$ which contains no holomorphic bubble points, and no points where $|\nabla u_\infty| = 0$, with the possible exception of the point a itself. (Recall that both holomorphic bubble points and points where $|\nabla u_\infty| = 0$ are isolated.) Fixing $q \in (1, 2)$, we may invoke Lemma 2.18 over the region $D_{2\mu}\backslash D_\mu$ to control the L^q norm of $e_{\bar\partial}(u_n)^{\frac{1}{2}}$. In particular, we have

$$\|(u_n)_x + (u_n) \times (u_n)_y\|^2_{L^q(D_{2\mu}\backslash D_\mu)} \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)},$$

for sufficiently large n, and some constant C. By virtue of the convergence (2.61) we may then invoke part (c) of Lemma 2.3 for sufficiently large n, to establish that

$$\begin{aligned} E_{\bar\partial}(u_n, D_\mu) &\leq C\left(\|\mathcal{T}(u_n)\sigma\|^2_{L^2(D_{2\mu})} + \|(u_n)_x + (u_n) \times (u_n)_y\|^2_{L^q(D_{2\mu}\backslash D_\mu)}\right) \\ &\leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}. \end{aligned} \tag{2.62}$$

It remains to repeat this process on a finite number of discs like D_μ, covering Ω, where each corresponding $D_{2\mu}$ lies within $\hat\Omega$. Adding the resulting estimates of the form (2.62), we reach our conclusion that

$$E_{\bar{\partial}}(u_n, \Omega) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)},$$

for sufficiently large n. □

2.5.2. *Analysis of neighbourhoods of antiholomorphic bubbles.* Whilst the previous section was concerned with obtaining $\bar{\partial}$-energy estimates on fixed (independent of n) compact subsets of the domain, in this section we locate necks around each antiholomorphic bubble — or appropriate groups of them — and derive $\bar{\partial}$-energy estimates right up to those necks. Meanwhile, we establish estimates of the energy *on* the necks, and ∂-energy estimates within the necks — i.e. over the antiholomorphic bubbles. The neck regions contract around the bubbles in the limit $n \to \infty$, and in particular, they depend on n.

Amongst the results we will apply in this section are the estimates of Section 2.4.2 controlling the size of any antiholomorphic bubbles, the neck analysis of Section 2.2 which will then furnish us with a dyadic neck with extremely small energy (in terms of the tension) and the estimates for e_∂ and $e_{\bar{\partial}}$ in parts (b) and (d) of Lemma 2.3 which will be applied inside and outside of the neck regions respectively. Part (d) of Lemma 2.3 will leave us with an unwanted boundary term which is dealt with using Lemma 2.19.

There is a technical complication involved here which arises from the possibility of having more than one bubble developing at the same point. To extract the maximum amount of information from the neck analysis of Lemma 2.9, we must find 'fat' annuli $D_\gamma \backslash D_{\gamma r^2}$ (with $0 < r \ll 1$) surrounding the bubbles, but not actually containing any. Given two bubbles developing at the same point (in the limit $n \to \infty$) we must decide whether they are both to be grouped within the same fat annulus, or whether they will have individual annuli. The answer will depend on whether these bubbles develop in close proximity to each other (in a renormalised sense) or not, and is the subject of the technical Lemma 2.21 later in this section.

LEMMA 2.20. *Suppose that u_n is a sequence satisfying the hypotheses of Theorem 1.2, and that $a \in S^2$ is an antiholomorphic bubble point about which we take a stereographic coordinate chart (sending $a \in S^2$ to $0 \in \mathbb{R}^2$).*

Then we can find $\mu > 0$ sufficiently small (in particular so that $\overline{D_{2\mu}}$ contains no bubble points other than $0 \in \mathbb{R}^2$) and a constant $C > 0$, and some $l \in \{1, \dots k\}$ (where k is the number of bubbles developing at $0 \in \mathbb{R}^2$ just as in Theorem 1.2) so that after passing to some subsequence in n, there exist a sequence of points $b_n^i \to 0 \in \mathbb{R}^2$ in D_μ and a decreasing sequence of numbers $\xi_n \downarrow 0$ such that

(a) *The discs $D_{b_n^i, e\xi_n}$ are disjoint, for $i \in \{1, \dots, l\}$, and lie within D_μ,*

(b)

$$\xi_n \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right],$$

and writing

$$B_n := \bigcup_i D_{b^i_n,\xi_n} \qquad\qquad A_n := \left(\bigcup_i D_{b^i_n,e\xi_n}\right)\backslash B_n,$$

we have

(i) $E_{\bar\partial}(u_n, D_\mu\backslash(A_n\cup B_n)) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}$,

(ii) $E_\partial(u_n, B_n) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}$,

(iii) $E(u_n, A_n) \leq C(\xi_n)^{\frac{1}{3}}$,

(iv) $\mathrm{osc}(u_n, D_{b^i_n,e\xi_n}\backslash D_{b^i_n,\xi_n}) \leq 1$, *for each* i.

As mentioned above, we shall require the following technical lemma during the proof of Lemma 2.20, which will group bubbles which are developing so close to each other that they cannot each be enclosed by a 'fat' annulus.

LEMMA 2.21. *Suppose, just as in Lemma* 2.20, *that* u_n *is a sequence satisfying the hypotheses of Theorem* 1.2, *and that* $a \in S^2$ *is an antiholomorphic bubble point about which we take a stereographic coordinate chart* (*sending* $a \in S^2$ *to* $0 \in \mathbb{R}^2$). *Using the notation of Theorem* 1.1, *we know that we may associate bubble data* $a^i_n \to 0 \in \mathbb{R}^2$, $\lambda^i_n \downarrow 0$ *and* $\omega^i : S^2 \cong \mathbb{R}^2\cup\{\infty\} \to S^2$ (*for* $i \in \{1,\ldots,k\}$) *to the bubble point* $a \in S^2$.

Then we may pass to a subsequence in n, *and find* l *new sequences of points* $b^i_n \to 0 \in \mathbb{R}^2$ (*for some* l *and all* i *with* $1 \leq i \leq l \leq k$) *a constant* $C > 0$ *and a decreasing sequence* $\eta_n \downarrow 0$ *such that*

(i) *The discs* $D_{b^i_n,\eta_n^{1/2}}$ *are disjoint, for* $i \in \{1,\ldots,l\}$,

(ii)

$$\bigcup_{i=1}^k D_{a^i_n,(\lambda^i_n)^{1/2}} \subset \bigcup_{i=1}^l D_{b^i_n,\eta_n},$$

(iii)

$$\eta_n \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right] \leq 1.$$

Proof of Lemma 2.20. Let us choose $\mu \in (0, \frac{1}{2}]$ sufficiently small so that $\overline{D_{2\mu}}$ contains no other bubble points (other than $0 \in \mathbb{R}^2$) and so that both

$$E(u_\infty, D_\mu) < \frac{\delta}{2},$$

where δ is taken from Lemma 2.9, and

$$E(u_\infty, D_{2\mu}) < \frac{\varepsilon_1}{2},$$

with ε_1 as in Lemma 2.3.

By parts (vi), (vii) and (iii) of Lemma 2.15, we know that

$$E_{\bar\partial}\left(u_n, D_\mu\backslash\left(\bigcup_{i=1}^k D_{a_n^i,(\lambda_n^i)^{1/2}}\right)\right) \to 0, \tag{2.63}$$

$$E\left(u_n, D_\mu\backslash\left(\bigcup_{i=1}^k D_{a_n^i,(\lambda_n^i)^{1/2}}\right)\right) \to E(u_\infty, D_\mu) < \frac{\delta}{2}, \tag{2.64}$$

and

$$E_\partial(u_n, D_{2\mu}) \to E_\partial(u_\infty, D_{2\mu}) = E(u_\infty, D_{2\mu}) < \frac{\varepsilon_1}{2}. \tag{2.65}$$

Therefore, for sufficiently large n we will have

$$E_{\bar\partial}\left(u_n, D_\mu\backslash\left(\bigcup_{i=1}^k D_{a_n^i,(\lambda_n^i)^{1/2}}\right)\right) \leq \varepsilon_3, \tag{2.66}$$

where ε_3 is taken from Lemma 2.3 (and we remove its q dependence by fixing $q = \frac{13}{12}$)

$$E\left(u_n, D_\mu\backslash\left(\bigcup_{i=1}^k D_{a_n^i,(\lambda_n^i)^{1/2}})\right)\right) < \delta, \tag{2.67}$$

and

$$E_\partial(u_n, D_{2\mu}) < \varepsilon_1. \tag{2.68}$$

Moreover, since $\mathcal{T}(u_n) \to 0$ in $L^2(S^2)$, we may assume (for sufficiently large n) that

$$\|\mathcal{T}(u_n)\sigma\|^2_{L^2(D_\mu)} < \delta.$$

Let us now invoke Lemma 2.21, which generates sequences of points $b_n^i \to 0 \in \mathbb{R}^2$ and a decreasing sequence $\eta_n \downarrow 0$, with the properties (i), (ii) and (iii) of that lemma. Properties (i) and (ii) combined with (2.67) guarantee that for each $i \in \{1, \ldots, l\}$ (and sufficiently large n) the annulus

$$\Sigma_n^i := D_{b_n^i,\eta_n^{1/2}}\backslash D_{b_n^i,\eta_n}$$

satisfies

$$E(u_n, \Sigma_n^i) < \delta,$$

and therefore, for sufficiently large n, we may apply Lemma 2.9 to Σ_n^i — where we are setting $\gamma = (\eta_n)^{\frac{1}{2}}$ and $\gamma r^2 = \eta_n$, i.e. $r = (\eta_n)^{\frac{1}{4}}$ — to find that

$$E(u_n, D_{b_n^i, e\xi_n} \backslash D_{b_n^i, \xi_n}) \leq C(\eta_n)^{\frac{1}{4}} = C(\xi_n)^{\frac{1}{3}}, \tag{2.69}$$

and

$$\operatorname{osc}(u_n, D_{b_n^i, e\xi_n} \backslash D_{b_n^i, \xi_n}) \leq C(\eta_n)^{\frac{1}{8}} \leq 1, \tag{2.70}$$

where we make the definition $\xi_n := (\eta_n)^{\frac{3}{4}}$. Notice that now we have defined the sequences b_n^i and ξ_n, we already know that parts (a) and (b) of Lemma 2.20 are true for sufficiently large n. The two statements (2.69) and (2.70) give us parts (iii) and (iv) of Lemma 2.20.

The next stage is to apply part (b) of Lemma 2.3 with $b = b_n^i$, $\nu = \xi_n$ and $u = u_n$ to get the estimate

$$E_\partial(u_n, D_{b_n^i, \xi_n}) \leq C\left(\|\mathcal{T}(u_n)\|^2_{L^2(S^2)} + E(u_n, D_{b_n^i, e\xi_n} \backslash D_{b_n^i, \xi_n})\right),$$

which may be summed to

$$E_\partial(u_n, B_n) \leq C\left(\|\mathcal{T}(u_n)\|^2_{L^2(S^2)} + E(u_n, A_n)\right). \tag{2.71}$$

Since we have just proved part (iii) of the lemma, we know that

$$E(u_n, A_n) \leq C(\xi_n)^{\frac{1}{3}} \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right] \leq \|\mathcal{T}(u_n)\|^2_{L^2(S^2)},$$

(where the constant C is changing throughout as usual) for sufficiently large n. Combining this with (2.71) yields part (ii) of the lemma.

It remains to prove part (i) of the lemma. For this, we have part (d) of Lemma 2.3, which we apply with $b^i = b_n^i$, $u = u_n$ and $\nu = \xi_n$ to find that

$$\begin{aligned} &E_{\bar\partial}(u_n, D_\mu \backslash (A_n \cup B_n)) \\ &\quad \leq C\left(\|\mathcal{T}(u_n)\|^2_{L^2(S^2)} + E_{\bar\partial}(u_n, D_{2\mu} \backslash D_\mu) + \xi_n^{-\frac{4(q-1)}{q}} E(u_n, A_n)\right). \end{aligned} \tag{2.72}$$

The second term on the right-hand side can be handled with the estimates developed in Section 2.5.1. Indeed, an application of Lemma 2.19 tells us that

$$E_{\bar\partial}(u_n, D_{2\mu} \backslash D_\mu) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}, \tag{2.73}$$

for sufficiently large n. Meanwhile, we handle the final term of (2.72) with part (iii) of the present lemma — which we have proved above. Recalling that we have set $q = \frac{13}{12}$, we have that

$$\xi_n^{-\frac{4(q-1)}{q}} E(u_n, A_n) \leq C(\xi_n)^{-\frac{4}{13}} (\xi_n)^{\frac{1}{3}} = C(\xi_n)^{\frac{1}{39}}.$$

By using the upper bound for ξ_n of part (b) of the present lemma — which is also dealt with above — we then know that

$$\xi_n^{-\frac{4(q-1)}{q}} E(u_n, A_n) \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right] \leq \|\mathcal{T}(u_n)\|^2_{L^2(S^2)}, \tag{2.74}$$

for some new C and sufficiently large n. A combination of (2.72), (2.73) and (2.74) then establishes part (i) of the lemma for sufficiently large n. □

We conclude this section with the proof of Lemma 2.21 which generates an appropriate grouping of any antiholomorphic bubbles developing at each point. We use a finite iteration procedure which at each step will group two bubbles which are developing too close to each other.

Proof of Lemma 2.21. Our starting point is the bubble concentration estimate of Lemma 2.17, which tells us that

$$\lambda_n^i \leq \exp\left[-\frac{1}{C_0\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right], \tag{2.75}$$

for all $i \in \{1, \ldots, k\}$ and sufficiently large n. Here C_0 is a constant whose value will now be fixed during this proof. Let us set

$$\eta_n = \exp\left[-\frac{1}{3C_0\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right],$$

and $l = k$, and $b_n^i = a_n^i$ for the moment. For sufficiently large n — and thus for all n after passing to a subsequence — we have $\eta_n \leq 1$, and this then ensures that part (ii) as well as part (iii) hold with our present choices of sequences. The remainder of the proof will be an iterative procedure designed to modify the points b_n^i, reducing l by one each time, and alter η_n, so that parts (ii) and (iii) of the lemma continue to hold, but so that eventually part (i) will also hold.

It is possible that for sufficiently large n, part (i) already holds — i.e. that the discs $D_{b_n^i,\eta_n^{1/2}}$ are disjoint — in which case we have finished the proof. If this is not the case, then after passing to some subsequence in n and relabelling in i, we may assume that $D_{b_n^{l-1},\eta_n^{1/2}}$ intersects $D_{b_n^l,\eta_n^{1/2}}$ for all n. We may now replace these two discs by one, whose centre is at the midpoint of b_n^{l-1} and b_n^l. More precisely, we redefine b_n^{l-1} to be the point $\frac{1}{2}(b_n^{l-1} + b_n^l)$, and reduce l by one (which has the effect of throwing away the old b_n^l). At the same time, we replace the old sequence η_n by $2(\eta_n)^{\frac{1}{2}}$.

Having made this modification, we find that the set

$$\bigcup_{i=1}^{l} D_{b_n^i,\eta_n}$$

with the new definitions of l, b_n^i and η_n, must be a superset of the same union with the old definitions. Therefore, part (ii) of the lemma remains true. Moreover, we clearly have

$$\eta_n \leq 2\exp\left[-\frac{1}{6C_0\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right] \leq \exp\left[-\frac{1}{7C_0\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right],$$

for sufficiently large n, and so part (iii) of the lemma remains true.

Now let us iterate the above procedure. Since l decreases by one each iteration, after at most $l-1$ steps, the discs $D_{b_n^i,\eta_n^{1/2}}$ will be disjoint, and all parts of the lemma will be satisfied. □

2.5.3. *Neck surgery and energy quantization.* In this section, we collect a few results which will enable us to take a bubble tree, decompose it into collections of bubble maps and body map — without altering the energy too much — and then show that the energy of each component is near to a multiple of 4π. These results assume that we can control the energy over some dyadic 'neck' annulus around the bubbles (eventually in terms of the tension) and this difficult issue is addressed elsewhere in the paper.

The neck surgery lemma is a simple result which says that we can chop out one or more bubbles (see the map w^1 below) or isolate them (see w^2) and be sure that the dyadic 'neck region' cannot increase its energy by more than a universal factor.

LEMMA 2.22. *Suppose that $s > 0$ and $u : \mathbb{R}^2 \to S^2$ is smooth, with*

$$\operatorname{osc}(u,\Omega) := \sup_{\mathbf{x},\mathbf{a}\in\Omega} |u(\mathbf{x}) - u(\mathbf{a})| \leq 1,$$

where $\Omega := D_{es}\backslash D_s$. Then there exist smooth maps $w^i : \mathbb{R}^2 \to S^2$, and points $a^i \in S^2$ — for $i = 1,2$ — such that

$$w^1(x) = \begin{cases} a^1 & x \in D_s \\ u(x) & x \in \mathbb{R}^2\backslash D_{es} \end{cases}, \tag{2.76}$$

and

$$w^2(x) = \begin{cases} u(x) & x \in D_s \\ a^2 & x \in \mathbb{R}^2\backslash D_{es} \end{cases},$$

and with energy constrained on the remaining dyadic annulus Ω according to

$$E(w^i,\Omega) \leq CE(u,\Omega), \tag{2.77}$$

for some universal C, and $i = 1,2$.

Proof. By the conformal invariance of energy, we may scale the domain and assume that $s = 1$. Meanwhile, by virtue of the restriction on the oscillation of u within Ω, the image of Ω under u must lie within some hemisphere of the target. Therefore, we may compose u with an appropriate stereographic projection $S^2 \to \mathbb{R}^2 \cup \{\infty\}$ to get a new map $\mathbb{R}^2 \to \mathbb{R}^2 \cup \{\infty\}$ — which we still call u — with the property that $u(\Omega) \subset D$, where D is the unit disc in $\mathbb{R}^2$ as usual.

Let us define

$$\bar{u} := \frac{1}{|\Omega|} \int_\Omega u \in D,$$

and choose some smooth cut-off function $\varphi : \mathbb{R}^2 \to [0,1]$ such that $|\nabla \varphi| \leq 1$ and

$$\varphi(x) = \begin{cases} 0 & x \in D \\ 1 & x \in \mathbb{R}^2 \backslash D_e \end{cases}.$$

Then we can define

$$w^1(x) := \varphi(x)u(x) + (1 - \varphi(x))\bar{u} = \varphi(x)(u(x) - \bar{u}) + \bar{u},$$

which will satisfy (2.76) with a^1 corresponding to $\bar{u}$, together with the range constraint $w^1(\Omega) \subset D$.

Given this definition for w^1, we would now like to show that it satisfies (2.77). In order to do this, we cling to the viewpoint that w^1 maps Ω into $D \subset \mathbb{R}^2$ (rather than into a hemisphere of S^2 which will be the ultimate viewpoint) and in this framework,

$$E(w^1, \Omega) = \frac{1}{2} \int_\Omega \rho^2(w^1)|\nabla w^1|^2,$$

where $\rho(w^1) = \frac{2}{1+|w^1|^2}$ is the conformal factor for the target — analagous to σ for the domain. We may then control the integrand by

$$\begin{aligned} \rho^2(w^1)|\nabla w^1|^2 \leq 4|\nabla w^1|^2 = 4|\nabla [\varphi(u - \bar{u})]|^2 &\leq 8(|\nabla \varphi|^2 |u - \bar{u}|^2 + |\varphi|^2 |\nabla u|^2) \\ &\leq 8(|u - \bar{u}|^2 + |\nabla u|^2), \end{aligned}$$

and thus the integral by

$$E(w^1, \Omega) \leq 4 \int_\Omega (|u - \bar{u}|^2 + |\nabla u|^2) \leq C \int_\Omega |\nabla u|^2,$$

for some universal C, by Poincaré's inequality. Moreover, since u maps Ω into D, on which $\rho \geq 1$, we then have

$$E(w^1, \Omega) \leq C \int_\Omega \rho^2(u)|\nabla u|^2 = C\, E(u, \Omega),$$

which is precisely (2.77).

The proof for w^2 is identical to the proof for w^1 except that we replace φ by $1 - \varphi$. □

In applications, this lemma will be combined with an energy quantization lemma which we now describe. The set $\hat{\Sigma}\backslash\Sigma$ below will correspond to a neck region, whilst Σ will correspond either to a bubble region which we are trying to isolate (in (2.79) below) or to a body region from which we would like to remove bubbles (in (2.78) below). We will not have to worry about controlling the value of $\deg(w)$ below; the important point is that it is an integer.

Lemma 2.23. *Suppose that u and w are both smooth maps $S^2 \to S^2$, and that we have two (measurable) sets $\Sigma \subset \hat{\Sigma} \subset S^2$ such that $u \equiv w$ on Σ and $w \equiv a$ on $S^2\backslash\hat{\Sigma}$, for some point $a \in S^2$. Then*

$$|E(u,\Sigma) - 4\pi \deg(w)| \leq 2E_{\bar{\partial}}(u,\Sigma) + 3E(w,\hat{\Sigma}\backslash\Sigma), \tag{2.78}$$

and

$$|E(u,\hat{\Sigma}) + 4\pi \deg(w)| \leq 2E_{\partial}(u,\Sigma) + 3E(w,\hat{\Sigma}\backslash\Sigma) + E(u,\hat{\Sigma}\backslash\Sigma). \tag{2.79}$$

Proof. In order to prove (2.78) we need only the hypotheses on u, w and a, and the fact that

$$E(w) - 4\pi \deg(w) = 2E_{\bar{\partial}}(w), \tag{2.80}$$

which follows from (1.9) and (1.10). This equips us completely for the calculation

$$\begin{aligned} |E(u,\Sigma) - 4\pi \deg(w)| &= |E(w,\Sigma) - 4\pi \deg(w)| \\ &\leq |E(w) - 4\pi \deg(w)| + E(w,\hat{\Sigma}\backslash\Sigma) \\ &= 2E_{\bar{\partial}}(w) + E(w,\hat{\Sigma}\backslash\Sigma) \\ &\leq 2E_{\bar{\partial}}(w,\Sigma) + 2E_{\bar{\partial}}(w,\hat{\Sigma}\backslash\Sigma) + E(w,\hat{\Sigma}\backslash\Sigma) \\ &\leq 2E_{\bar{\partial}}(u,\Sigma) + 3E(w,\hat{\Sigma}\backslash\Sigma), \end{aligned}$$

which is precisely (2.78).

Meanwhile, the second part (2.79) is similar but uses the fact that

$$E(w) + 4\pi \deg(w) = 2E_{\partial}(w),$$

in place of (2.80), in the calculation

$$\begin{aligned} |E(u,\hat{\Sigma}) + 4\pi \deg(w)| &\leq |E(u,\hat{\Sigma}) - E(w,\hat{\Sigma})| + |E(w) + 4\pi \deg(w)| \\ &\leq E(u,\hat{\Sigma}\backslash\Sigma) + E(w,\hat{\Sigma}\backslash\Sigma) + 2E_{\partial}(w) \\ &\leq E(u,\hat{\Sigma}\backslash\Sigma) + E(w,\hat{\Sigma}\backslash\Sigma) + 2E_{\partial}(w,\Sigma) + 2E_{\partial}(w,\hat{\Sigma}\backslash\Sigma) \\ &\leq 2E_{\partial}(u,\Sigma) + 3E(w,\hat{\Sigma}\backslash\Sigma) + E(u,\hat{\Sigma}\backslash\Sigma). \end{aligned}$$

□

2.5.4. *Assembly of the proof of Theorem* 1.2. We are now in a position to assemble the proof of Theorem 1.2. The bubble concentration estimate (1.4) has already been established in Lemma 2.17. This leaves the energy quantization estimate 1.3, and for this we will appeal mainly to the analysis of the neighbourhoods of antiholomorphic bubbles made in Lemma 2.20, the control of $E_{\bar{\partial}}$ away from antiholomorphic bubbles made in Lemma 2.19 and the surgery and quantization results of Section 2.5.3.

We should first deal with the case that there are no antiholomorphic bubbles. In this case, we may apply Lemma 2.19 with $\Omega = S^2$ giving

$$E_{\bar{\partial}}(u_n) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)},$$

for some C and sufficiently large n. By subtracting the identities (1.9) and (1.10) and dividing by two, we know that

$$E_{\bar{\partial}}(u_n) = \frac{1}{2}(E(u_n) - 4\pi \deg(u_n)),$$

and so we may conclude that

$$|E(u_n) - 4\pi k| \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)},$$

with integral k, a new constant C, and sufficiently large n. Therefore we may now assume that there *are* some antiholomorphic bubbles.

Let us define a set $\Upsilon \subset S^2$ to be the whole of the domain S^2, with open discs removed around each antiholomorphic bubble point corresponding to the discs D_μ arising in Lemma 2.20. We also define — for each n — a set $\Gamma_n \subset S^2$ to be the complement in S^2 of the union (over each antiholomorphic bubble point) of all subsets corresponding to $B_n \cup A_n$ arising in Lemma 2.20, and a set $\hat{\Gamma}_n \subset S^2$ to be the complement in S^2 of the union of all subsets corresponding to B_n. In particular, we have $\Upsilon \subset \Gamma_n \subset \hat{\Gamma}_n$ for each n.

An application of Lemma 2.19 tells us that

$$E_{\bar{\partial}}(u_n, \Upsilon) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}, \tag{2.81}$$

for some C and sufficiently large n. This estimate can then be improved by making an application of part (i) of Lemma 2.20 to each antiholomorphic bubble point and summing the resulting estimates with (2.81). This process yields the estimate

$$E_{\bar{\partial}}(u_n, \Gamma_n) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}, \tag{2.82}$$

for some constant C, after we have passed to a subsequence in n.

We may also sum the estimates from part (ii) of Lemma 2.20 over each antiholomorphic bubble point and deduce that

$$E_{\partial}(u_n, S^2 \backslash \hat{\Gamma}_n) \leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}, \tag{2.83}$$

for some constant C.

The analogous procedure for the estimate of part (iii) of Lemma 2.20 — coupled with part (b) of that same lemma — tells us that

$$E(u_n, \hat{\Gamma}_n \backslash \Gamma_n) \leq \exp\left[-\frac{1}{C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}}\right] \leq \|\mathcal{T}(u_n)\|^2_{L^2(S^2)}, \tag{2.84}$$

for some constant C, and sufficiently large n.

Having compiled estimates (2.82), (2.83) and (2.84), we may now perform surgery on the maps u_n. Our wish is to obtain one map w_n^1 — for each n — which captures the holomorphic part of the bubble tree which forms in the limit of large n, and also a number of maps — for each n — which capture the antiholomorphic parts of the bubble tree. To do this, we apply Lemma 2.22 at *each* antiholomorphic bubble point, and for *each* annulus

$$\Omega := D_{b_n^i, e\xi_n} \backslash D_{b_n^i, \xi_n},$$

for $i = 1, \ldots, l$ where we have adopted notation from Lemma 2.20. Notice that we are using part (iv) of Lemma 2.20 to fulfill the hypotheses of Lemma 2.22.

At each application of Lemma 2.22, we use the existence of the map w^1 to remove one or more antiholomorphic bubbles from u_n. After removing them all — at all antiholomorphic bubble points — we are left with a smooth map $w_n^1 : S^2 \to S^2 \hookrightarrow \mathbb{R}^3$ satisfying

$$w_n^1(x) = u_n(x) \qquad \text{for } x \in \Gamma_n,$$

with w_n^1 constant on each connected component of $S^2 \backslash \hat{\Gamma}_n$. Of course, the information obtained from Lemma 2.22 is the estimate (2.77), and if we sum this estimate over each application of the lemma, we find that

$$E(w_n^1, \hat{\Gamma}_n \backslash \Gamma_n) \leq C E(u_n, \hat{\Gamma}_n \backslash \Gamma_n), \tag{2.85}$$

for some C dependent only on M.

Meanwhile, at *each* application of Lemma 2.22 — i.e. at each antiholomorphic bubble point and for each $i = 1, \ldots, l$ — we are given a map $w_n^2 : S^2 \to S^2 \hookrightarrow \mathbb{R}^3$ which isolates one or more antiholomorphic bubbles. More precisely, at each application, w_n^2 satisfies

$$w_n^2(x) = u_n(x) \qquad \text{for } x \in D_{b_n^i, \xi_n},$$

with w_n^2 constant on $S^2 \backslash D_{b_n^i, e\xi_n}$, and we have the estimate

$$E(w_n^2, D_{b_n^i, e\xi_n} \backslash D_{b_n^i, \xi_n}) \leq C E(u_n, D_{b_n^i, e\xi_n} \backslash D_{b_n^i, \xi_n}), \tag{2.86}$$

for some universal C.

We may now apply the partial energy quantization results of Lemma 2.23. To begin with, we make a single application of (2.78) from that lemma, with $\Sigma = \Gamma_n$, $\hat{\Sigma} = \hat{\Gamma}_n$, $u = u_n$ and $w = w_n^1$. This tells us that for some nonnegative integer k, we have

$$|E(u_n, \Gamma_n) - 4\pi k| \leq 2E_{\bar{\partial}}(u_n, \Gamma_n) + 3E(w_n^1, \hat{\Gamma}_n \backslash \Gamma_n),$$

and by virtue of (2.82), (2.85) and (2.84) this improves to

$$|E(u_n,\Gamma_n)-4\pi k|\leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}, \tag{2.87}$$

for some C.

Now we make several applications of (2.79) from Lemma 2.23, with $\Sigma = D_{b^i_n,\xi_n}$, $\hat{\Sigma} = D_{b^i_n,e\xi_n}$, $u = u_n$ and $w = w^2_n$ (for each antiholomorphic bubble point, and each valid i). We learn that

$$\begin{aligned}|E(u_n,D_{b^i_n,e\xi_n})-4\pi k|\leq{}& 2E_\partial(u_n,D_{b^i_n,\xi_n})\\ &+3E(w^2_n,D_{b^i_n,e\xi_n}\backslash D_{b^i_n,\xi_n})+E(u_n,D_{b^i_n,e\xi_n}\backslash D_{b^i_n,\xi_n}),\end{aligned}$$

for some integer k. Note that the map w^2_n is different for each different bubble point and each different i. Combining with (2.86) then yields

$$|E(u_n,D_{b^i_n,e\xi_n})-4\pi k|\leq 2E_\partial(u_n,D_{b^i_n,\xi_n})+C\,E(u_n,D_{b^i_n,e\xi_n}\backslash D_{b^i_n,\xi_n}),$$

for some universal C, and we may sum this estimate over i and then over all bubble points, to find that

$$|E(u_n,S^2\backslash\Gamma_n)-4\pi k|\leq 2E_\partial(u_n,S^2\backslash\hat{\Gamma}_n)+C\,E(u_n,\hat{\Gamma}_n\backslash\Gamma_n),$$

for some new integer k.

Now we may invoke (2.83) and (2.84) to yield

$$|E(u_n,S^2\backslash\Gamma_n)-4\pi k|\leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}.$$

We conclude by adding this estimate to (2.87) giving

$$|E(u_n)-4\pi k|\leq C\|\mathcal{T}(u_n)\|^2_{L^2(S^2)},$$

for some integer k, some constant C, and sufficiently large n. □

3. Heat flow - the proof of Theorem 1.7

This section is devoted to proving the exponential uniform asymptotic convergence of Theorem 1.7. The main ingredient will be the quantization estimate of Theorem 1.2, coupled with the ideas of our previous work [15]. The main remaining difficulty will be to work around the restrictions on the bubble tree imposed in Theorem 1.2.

THEOREM 3.1. *Suppose that* $u_\infty : S^2 \to S^2$ *is some nonconstant holomorphic map, that* $M > 0$, *and that* $B_1, \ldots, B_m \subset S^2$ *are closed disjoint geodesic balls in* S^2.

Given this data, if we define Q^ε *to be the set of smooth maps* $u : S^2 \to S^2 \hookrightarrow \mathbb{R}^3$ *satisfying*

(i) $E(u) < M$,

(ii) $\|u - u_\infty\|_{W^{1,2}(S^2 \setminus \cup B_i)} < \varepsilon$,

(iii) *either* $E_{\bar{\partial}}(u, B_i) < 2\pi$ *or* $E_\partial(u, B_i) < 2\pi$ *for each* B_i,

(iv) $E_{\bar{\partial}}(u, B_i) < 2\pi$ *for any* B_i *containing a point where* $|\nabla u_\infty| = 0$,

then there exist constants $\varepsilon_0 > 0$ *and* C *such that any map* $u \in Q^{\varepsilon_0}$ *satisfies*

$$|E(u) - 4\pi k| \leq C\|\mathcal{T}(u)\|^2_{L^2(S^2)},$$

for some integer k.

It will shortly become clear why we have used so much familiar notation in the above theorem.

Proof. Suppose that this result is not true for some u_∞, M and sets B_i. Then there must exist a sequence of smooth maps $u_n : S^2 \to S^2$ such that $u_n \in Q^{\frac{1}{n}}$ but so that

$$|E(u_n) - 4\pi k| \geq n\|\mathcal{T}(u_n)\|^2_{L^2(S^2)}, \tag{3.1}$$

for each n and integral k. In particular, since there is always some $k \in \mathbb{Z}$ which makes the left-hand side of (3.1) no more than 2π, we have that $\|\mathcal{T}(u_n)\|^2_{L^2(S^2)} \leq \frac{2\pi}{n}$. Now since $E(u_n) < M$ for each n, and $\mathcal{T}(u_n) \to 0$ in $L^2(S^2)$, we may apply Theorem 1.1 and pass to a subsequence which undergoes bubbling as described in that theorem. Note that despite passing to a subsequence, we are still guaranteed that $u_n \in Q^{\frac{1}{n}}$, and (3.1) will still hold.

Observe that part (ii) of the definition of Q^ε, and the fact that $u_n \in Q^{\frac{1}{n}}$, tell us that

$$u_n \to u_\infty \qquad \text{in } W^{1,2}(S^2 \setminus \cup B_i), \tag{3.2}$$

and we may deduce that the u_∞ of Theorem 3.1 agrees with the u_∞ of Theorem 1.1. Let us consider what type of bubbling is possible in the sequence u_n. By (3.2), we see that all bubble points must lie within one of the balls B_i. Meanwhile, part (iii) of the definition of Q^ε tells us that we cannot have both holomorphic and antiholomorphic bubbles developing within the same ball B_i; for example if a holomorphic bubble develops in B_1 then we know that $\limsup_{n\to\infty} E_\partial(u_n, B_1) \geq 4\pi$, by part (iv) of Lemma 2.15 and the fact that any holomorphic bubble must have ∂-energy of at least 4π.

In particular, we cannot have both holomorphic and antiholomorphic bubbles developing at the same point. Finally, part (iv) of the definition of Q^ε tells us that we cannot have an antiholomorphic bubble developing at a point where $|\nabla u_\infty| = 0$.

What we have established above is that the sequence u_n satisfies the hypotheses of Theorem 1.2. Consequently, a subsequence must satisfy the quantization estimate (1.3) which contradicts (3.1) and completes the proof. □

We now switch our attention to solutions $u : S^2 \times [0,\infty) \to S^2 \hookrightarrow \mathbb{R}^3$ of the heat equation (1.5) with initial map u_0, as found by Theorem 1.5.

Let us adopt the notation $B(x,R)$ to denote the closed geodesic ball in S^2 centred at $x \in S^2$ of radius R. For consistency with the rest of the paper, we shall write $M := E(u_0) + 1$ which then provides an upper bound for $E(u(t))$ for all $t \geq 0$.

We shall require some regularity estimates below, which state that if energy does not concentrate too much in a flow u then we can control its C^k norms. Estimates of this form originate in the work of Struwe [13, Lemma 3.10′].

LEMMA 3.2. *There exists $\varepsilon_1 > 0$ such that whenever we have a solution $u : S^2 \times [0,T) \to S^2$ of the heat equation* (1.5) (*with* $1 \leq T \leq \infty$) *satisfying*

$$\sup_{(x,t)\in\Omega\times[0,T)} E(u(t), B(x,R)) < \varepsilon_1,$$

for some $R \in (0, \frac{\pi}{2})$ and $\Omega \subset\subset S^2$, then for all $k \in \mathbb{N}$, there exists a constant C dependent on k, R, and M so that

$$\|u\|_{C^k(\Omega\times[1,T))} \leq C.$$

We will not repeat the proof from [13]. Note that the dependency on $E(u_0)$ has been changed to the upper bound M. Moreover, the dependency on T has been removed by iterating a fixed time length estimate over different time intervals.

We shall also require some control on how fast energy can flow into and out of local regions.

LEMMA 3.3. *Suppose that u is a solution of the heat equation* (1.5) *and $R \in (0, \frac{\pi}{4})$. Then there exists a constant $C = C(R,M)$ such that if $0 \leq t_0 \leq t$ and $x \in S^2$, then*

(i) $E(u(t), B(x,R)) \leq E(u(t_0), B(x,2R)) + C\int_{t_0}^t \|\mathcal{T}(u(s))\|_{L^2(S^2)} ds,$

(ii) $E_\partial(u(t), B(x,R)) \leq E_\partial(u(t_0), B(x,2R)) + C\int_{t_0}^t \|\mathcal{T}(u(s))\|_{L^2(S^2)} ds,$

(iii) $E_{\bar\partial}(u(t), B(x,R)) \leq E_{\bar\partial}(u(t_0), B(x,2R)) + C\int_{t_0}^t \|\mathcal{T}(u(s))\|_{L^2(S^2)} ds,$

whilst if $0 \leq t \leq t_0$ and the flow u is regular during the time interval $[t, t_0]$, then

$$E(u(t), B(x,R)) \leq E(u(t_0), B(x,2R)) + [E(u(t)) - E(u(t_0))] + C\int_t^{t_0} \|\mathcal{T}(u(s))\|_{L^2(S^2)} ds. \tag{3.3}$$

Proof. Once $x \in S^2$ has been chosen, we take a stereographic coordinate chart about this point, and choose $\mu = \tan\frac{R}{2}$ so that the disc $\overline{D_\mu}$ in that chart corresponds to the geodesic ball $B(x,R)$ of the lemma. Note that the ball $B(x,2R)$ corresponds to the disc $\overline{D_{\tan R}}$ which contains the disc $D_{2\mu}$, and is contained in the unit disc D.

Let us choose a smooth cut-off function $\varphi : D_{2\mu} \to [0,1]$ with compact support in $D_{2\mu}$, and $\varphi \equiv 1$ in D_μ. Then we may calculate

$$\begin{aligned}
\frac{d}{dt}\int_{D_{2\mu}} \frac{\varphi}{4}|u\times u_x - u_y|^2 \\
&= \frac{d}{dt}\int_{D_{2\mu}} \frac{\varphi}{4}\left(|u_x|^2+|u_y|^2-2\langle u, u_x\times u_y\rangle\right)\\
&= \int_{D_{2\mu}} \frac{\varphi}{2}\left(\langle u_x,u_{xt}\rangle + \langle u_y,u_{yt}\rangle - \langle u,u_{xt}\times u_y\rangle - \langle u, u_x\times u_{yt}\rangle\right),
\end{aligned}$$

where we have used the fact that u_t is orthogonal to $u_x\times u_y$. Let us now integrate each term on the right-hand side by parts. Pairing the first two terms, we have

$$\begin{aligned}
\int_{D_{2\mu}} \frac{\varphi}{2}\left(\langle u_x,u_{xt}\rangle + \langle u_y,u_{yt}\rangle\right) &= -\frac{1}{2}\int_{D_{2\mu}}\left(\varphi\langle\Delta u,u_t\rangle + \varphi_x\langle u_x,u_t\rangle+\varphi_y\langle u_y,u_t\rangle\right)\\
&= -\frac{1}{2}\int_{D_{2\mu}}\left(\varphi\sigma^2|\mathcal{T}|^2 + \varphi_x\langle u_x,\mathcal{T}\rangle+\varphi_y\langle u_y,\mathcal{T}\rangle\right),
\end{aligned}$$

where we have used the equation $u_t = \mathcal{T} = \frac{1}{\sigma^2}(\Delta u)^T$. Meanwhile, the final two terms satisfy

$$\begin{aligned}
-\frac{1}{2}\int_{D_{2\mu}}\varphi\left(\langle u,u_{xt}\times u_y\rangle + \langle u,u_x\times u_{yt}\rangle\right)\\
&= \frac{1}{2}\int_{D_{2\mu}}\left(\varphi_x\langle u,u_t\times u_y\rangle + \varphi_y\langle u,u_x\times u_t\rangle\right).
\end{aligned}$$

Combining all these expressions, we then have

$$\begin{aligned}
\frac{d}{dt}\int_{D_{2\mu}}\varphi e_{\bar\partial}(u)\sigma^2 + \frac{1}{2}\int_{D_{2\mu}}\varphi\sigma^2|\mathcal{T}|^2\\
&= \frac{1}{2}\int_{D_{2\mu}}\left(-\varphi_x\langle u_x,\mathcal{T}\rangle - \varphi_y\langle u_y,\mathcal{T}\rangle + \varphi_x\langle u,\mathcal{T}\times u_y\rangle + \varphi_y\langle u,u_x\times\mathcal{T}\rangle\right).
\end{aligned}$$

Whilst it is now possible to rearrange the right-hand side of this expression into a more natural form, we are content here with the simple estimate

$$\begin{aligned}
\left|\frac{d}{dt}\int_{D_{2\mu}}\varphi e_{\bar\partial}(u)\sigma^2 + \frac{1}{2}\int_{D_{2\mu}}\varphi\sigma^2|\mathcal{T}|^2\right| &\leq C(\mu)M^{\frac{1}{2}}\|\mathcal{T}\|_{L^2(D_{2\mu})}\\
&\leq C(\mu,M)\|\mathcal{T}\sigma\|_{L^2(D_{2\mu})},
\end{aligned}$$

where we are free to assume that $\sigma \geq 1$, since $2\mu \leq 1$. Rewriting this inequality as

$$-\frac{1}{2}\int_{D_{2\mu}} \varphi\sigma^2|\mathcal{T}|^2 - C\|\mathcal{T}\sigma\|_{L^2(D_{2\mu})} \leq \frac{d}{dt}\int_{D_{2\mu}} \varphi e_{\bar\partial}(u)\sigma^2 \leq C\|\mathcal{T}\sigma\|_{L^2(D_{2\mu})},$$

and weakening to

$$-\frac{1}{2}\|\mathcal{T}(u(t))\|^2_{L^2(S^2)} - C\|\mathcal{T}(u(t))\|_{L^2(S^2)} \leq \frac{d}{dt}\int_{D_{2\mu}} \varphi e_{\bar\partial}(u)\sigma^2 \leq C\|\mathcal{T}(u(t))\|_{L^2(S^2)}, \tag{3.4}$$

we may integrate the second inequality over the time interval $[t_0, t]$ to establish part (iii) of the lemma. An analogous calculation for E_∂ instead of $E_{\bar\partial}$ settles part (ii) of the lemma. Part (i) follows by summing parts (ii) and (iii).

If we sum (3.4) with the analogous version for e_∂, we find that

$$-\|\mathcal{T}(u(t))\|^2_{L^2(S^2)} - C\|\mathcal{T}(u(t))\|_{L^2(S^2)} \leq \frac{d}{dt}\int_{D_{2\mu}} \varphi e(u)\sigma^2 \leq C\|\mathcal{T}(u(t))\|_{L^2(S^2)}. \tag{3.5}$$

We integrate the first of these over time, but now we assume that $t \leq t_0$, and integrate over $[t, t_0]$. Assuming that the flow u is regular during this period, we know from (1.6) that

$$\int_t^{t_0} \|\mathcal{T}(u(s))\|^2_{L^2(S^2)} ds = E(u(t)) - E(u(t_0)). \tag{3.6}$$

The final part of the lemma, inequality (3.3), then follows. □

We are now in a position to prove Theorem 1.7. Our strategy will be to let the heat flow run until such time that it is well within an appropriate space Q^ε (as in Theorem 3.1); whilst within this space, we may exploit the quantization estimate of Theorem 3.1 to prove that the flow cannot move around too much — and in particular, we will be able to show that it cannot escape that space Q^ε. Therefore we have the quantization estimate for all time, which will force the energy to decay to its limiting value exponentially fast. We know from previous work — see [11] and [15] — that the only scenario in which the flow can keep moving indefinitely is when the energy decays extremely slowly, as we describe below. Therefore we establish that the flow is 'rigid' and converges uniformly in time in each possible sense.

Proof of Theorem 1.7. We have assumed in the statement of the theorem that at the sequence of times t_n, the bubbling obeys the hypotheses of Theorem 1.2. In particular, we are assuming that the body map u_∞ is holomorphic rather than antiholomorphic.

Let us observe that the theorem is a statement about the asymptotics of the flow, and therefore we may always assume that the flow has no finite time blow-up, by considering only the time beyond any such singularities.

With the notation of Theorem 1.1, the sequence $u(t_n)$ has bubbling at points $x^1, \ldots, x^m$. Let us choose $r > 0$ sufficiently small so that

- each closed ball $B(x^i, 2r)$ contains only one bubble point,
- $E(u_\infty, B(x^i, 2r)) \leq \frac{\pi}{2}$ for each i,
- if x^i is an antiholomorphic bubble point, then $|\nabla u_\infty| \neq 0$ throughout the closed ball $B(x^i, r)$.

Note that we could then reduce r to any smaller positive value whilst preserving the above conditions. If we now set $B_i := B(x^i, r)$, and consider the space Q^ε from Theorem 3.1 with u_∞ and M as in the flow under consideration (considered at times t_n) then by virtue of the bubbling convergence as described in Theorem 1.1 and parts (ii) and (iii) of Lemma 2.15, we have, for arbitrary $\varepsilon > 0$, $u(t_n) \in Q^\varepsilon$ for sufficiently large n (depending on ε).

In fact, with ε_0 as in Theorem 3.1, and for $\delta \in (0, \frac{\varepsilon_0}{3})$ to be chosen shortly (by imposing a finite number of positive upper bounds) we can find $N \in \mathbb{N}$ so that $E(u(t_N - 1)) - \overline{E} < \delta$, which implies that

$$E(u(t)) - \overline{E} < \delta, \tag{3.7}$$

for any $t \geq t_N - 1$ — and also so that

(a) $\|u(t_n) - u_\infty\|_{W^{1,2}(S^2 \setminus \cup_i B(x^i, \frac{1}{2}r))} < \delta$,

(b) either $E_{\bar\partial}(u(t_n), B(x^i, 2r)) < \pi$ or $E_\partial(u(t_n), B(x^i, 2r)) < \pi$ for each i,

(c) $E_{\bar\partial}(u(t_n), B(x^i, 2r)) < \pi$ for any i for which $B(x^i, r)$ contains a point where $|\nabla u_\infty| = 0$,

for $n \geq N$. Note that since $\delta < \frac{\varepsilon_0}{3}$, conditions (a)–(c) tell us more than that $u(t_n) \in Q^{\frac{\varepsilon_0}{3}}$.

Let us remark that the bounds we impose on δ in the sequel should only depend on information such as ε_0, the ε_1 from Lemma 3.2, M, u_∞ and the balls B_i. They clearly must not depend on N which itself depends on δ.

Suppose now that we set $T \in (0, \infty]$ to be the largest value for which $u(t) \in Q^{\varepsilon_0}$ for every $t \in [t_N, t_N + T)$. We claim that if δ is chosen sufficiently small, then we will have $T = \infty$.

Let us analyse the flow over the time interval $[t_N, t_N + T)$. Since $u(t) \in Q^{\varepsilon_0}$, we may appeal to Theorem 3.1 which tells us that

$$|E(u(t)) - 4\pi k| \leq C\|\mathcal{T}(u(t))\|^2_{L^2(S^2)},$$

for some integer k which may at first glance depend on t. However, note that by (3.7), if we assume that $\delta < 2\pi$, then $E(u(t)) - \overline{E} < 2\pi$ and $\overline{E} \in 4\pi\mathbb{Z}$ will be the optimal value of k; i.e. we will have

$$0 \leq E(u(t)) - \overline{E} \leq C\|\mathcal{T}(u(t))\|^2_{L^2(S^2)}. \tag{3.8}$$

An immediate consequence of this, using (1.6), is that

$$\left[E(u(t)) - \overline{E}\right] \leq -C\frac{d}{dt}\left[E(u(t)) - \overline{E}\right],$$

which forces the exponential decay

$$E(u(t)) - \overline{E} \leq \left[E(u(t_N)) - \overline{E}\right] \exp\left[-\frac{t-t_N}{C}\right], \tag{3.9}$$

for $t \in [t_N, t_N + T)$. An alternative application of (3.8) completes the calculation

$$\begin{aligned}
-\frac{d}{dt}\left[E(u(t)) - \overline{E}\right]^{\frac{1}{2}} &= -\frac{1}{2}\left[E(u(t)) - \overline{E}\right]^{-\frac{1}{2}}\frac{d}{dt}\left[E(u(t)) - \overline{E}\right] \\
&= \frac{1}{2}\left[E(u(t)) - \overline{E}\right]^{-\frac{1}{2}}\|\mathcal{T}(u(t))\|^2_{L^2(S^2)} \\
&\geq \frac{1}{2\sqrt{C}}\|\mathcal{T}(u(t))\|_{L^2(S^2)},
\end{aligned}$$

which may be integrated to give

$$\begin{aligned}
\int_s^t \|\mathcal{T}(u(\xi))\|_{L^2(S^2)} d\xi &\leq C\left(\left[E(u(s)) - \overline{E}\right]^{\frac{1}{2}} - \left[E(u(t)) - \overline{E}\right]^{\frac{1}{2}}\right) \\
&\leq C\left[E(u(s)) - \overline{E}\right]^{\frac{1}{2}} \leq C\delta^{\frac{1}{2}},
\end{aligned} \tag{3.10}$$

for $t_N \leq s \leq t \leq t_N + T$, thanks to (3.7).

Our mission now is to show that the restriction given in (3.10) is enough to establish that if we chose δ to be sufficiently small, then $u(t)$ will lie 'well within' Q^{ε_0} during the time interval $[t_N, t_N + T)$, in the sense that $u(t)$ will persist within Q^{ε_0} beyond any finite time $t_N + T$. This contradiction will establish that $T = \infty$ and allow us to appeal to (3.10) for all $t \geq t_N$.

First let us note that since $E(u(t)) < M$ is ensured for all t, part (i) of the definition of Q^ε (from Theorem 3.1) will never prevent $u(t)$ from lying within Q^{ε_0}.

Next let us turn to parts (iii) and (iv) of the definition of Q^ε. By appealing to Lemma 3.3 and (3.10), we know that for $t_N \leq s \leq t \leq t_N + T$, we must have

$$\begin{aligned}
E_\partial(u(t), B_i) &\leq E_\partial(u(s), B(x^i, 2r)) + C\int_s^t \|\mathcal{T}(u(\xi))\|_{L^2(S^2)} d\xi \\
&\leq E_\partial(u(s), B(x^i, 2r)) + C\delta^{\frac{1}{2}},
\end{aligned}$$

for various constants C. An analogous inequality with E_∂ replaced by $E_{\bar\partial}$ will also hold. Coupling these estimates (setting $s = t_N$) with (b) and (c) above (within this proof) we can be sure, provided we choose δ small enough, that

(b′) either $E_{\bar\partial}(u(t), B_i) < \frac{3\pi}{2}$ or $E_\partial(u(t), B_i) < \frac{3\pi}{2}$ for each i, (c′) $E_{\bar{\partial}}(u(t), B_i) < \frac{3\pi}{2}$ for any i for which B_i contains a point where $|\nabla u_\infty| = 0$,

for $t \in [t_N, t_N + T)$. Therefore, both parts (iii) and (iv) of the definition of Q^{ε_0} (from Theorem 3.1) will continue to hold for $u(t)$ as t passes beyond any finite time $t_N + T$.

It remains (in our attempt to force $T = \infty$) to show that part (ii) of the definition of Q^{ε_0} will continue to hold for $u(t)$ as t passes beyond any finite time $t_N + T$. In fact, we will prove that

$$\|u(t) - u_\infty\|_{W^{1,2}(S^2 \setminus \cup B_i)} \leq \frac{2\varepsilon_0}{3}, \tag{3.11}$$

for $t \in [t_N, t_N + T)$. Recall that by part (a) above, we have

$$\|u(t_N) - u_\infty\|_{W^{1,2}(S^2 \setminus \cup_i B(x^i, \frac{1}{2}r))} < \delta. \tag{3.12}$$

Our strategy is to control the quantity $u(t) - u_\infty$ in L^2, and then interpolate with bounds on the higher derivatives of u which will follow from Lemma 3.2.

The L^2 control follows easily from (3.10). Indeed, for $t_N \leq s \leq t < t_N + T$, we have

$$\begin{aligned} \|u(t) - u(s)\|_{L^2(S^2)} &\leq \int_s^t \left\|\frac{\partial u}{\partial t}\right\|_{L^2(S^2)} = \int_s^t \|\mathcal{T}(u(\xi))\|_{L^2(S^2)} d\xi \\ &\leq C\left[E(u(s)) - \overline{E}\right]^{\frac{1}{2}} \leq C\delta^{\frac{1}{2}}. \end{aligned} \tag{3.13}$$

Therefore by (3.12), we must have

$$\|u(t) - u_\infty\|_{L^2(S^2 \setminus \cup B_i)} \leq \delta + C\delta^{\frac{1}{2}}, \tag{3.14}$$

for $t \in [t_N, t_N + T)$. This L^2 estimate will be combined with some higher regularity estimates to which we now turn. By virtue of part (a) above, we know that there exists $R > 0$ dependent on u_∞, ε_1 and r, such that provided we chose δ sufficiently small (compared with ε_1) we must have

$$\sup_{x \in S^2 \setminus \cup B_i} E(u(t_N), B(x, 2R)) < \frac{\varepsilon_1}{2}.$$

We stress that part (a) is used to relieve R of a dependence on N. Application of Lemma 3.3 then provides similar control for different times. Part (i) of that lemma, and (3.10), tell us that

$$\begin{aligned} \sup_{(x,t) \in (S^2 \setminus \cup B_i) \times [t_N, t_N + T)} E(u(t), B(x, R)) &< \frac{\varepsilon_1}{2} + C \int_{t_N}^{t_N + T} \|\mathcal{T}(u(\xi))\|_{L^2(S^2)} d\xi \\ &\leq \frac{\varepsilon_1}{2} + C\delta^{\frac{1}{2}} < \varepsilon_1, \end{aligned} \tag{3.15}$$

for various constants $C = C(R, M, u_\infty, B_i)$, and sufficiently small δ. Meanwhile, (3.3) from Lemma 3.3 tells us that

$$\sup_{(x,t)\in(S^2\setminus\cup B_i)\times[t_N-1,t_N)} E(u(t), B(x,R)) < \frac{\varepsilon_1}{2} + [E(u(t_N-1)) - E(u(t_N))] + C\int_{t_N-1}^{t_N} \|\mathcal{T}(u(\xi))\|_{L^2(S^2)} d\xi,$$

and since

$$\int_{t_N-1}^{t_N} \|\mathcal{T}(u(\xi))\|_{L^2(S^2)} d\xi \leq \left(\int_{t_N-1}^{t_N} \|\mathcal{T}(u(\xi))\|^2_{L^2(S^2)} d\xi\right)^{\frac{1}{2}} = [E(u(t_N-1)) - E(u(t_N))]^{\frac{1}{2}},$$

we may apply (3.7) and deduce that

$$\sup_{(x,t)\in(S^2\setminus\cup B_i)\times[t_N-1,t_N)} E(u(t), B(x,R)) < \frac{\varepsilon_1}{2} + \delta + C\delta^{\frac{1}{2}} < \varepsilon_1, \tag{3.16}$$

for sufficiently small δ. The estimates (3.15) and (3.16) now serve to satisfy the hypotheses of Lemma 3.2 over the time interval $[t_N - 1, t_N + T)$, and therefore

$$\|u\|_{C^k((S^2\setminus\cup B_i)\times[t_N,t_N+T))} \leq C, \tag{3.17}$$

for some constant C dependent on k, M, and $R = R(u_\infty, r)$.

Let us now interpolate between these C^k bounds, and the L^2 estimates of (3.13) and (3.14). We know that if Σ is a $C^{1,1}$ domain in $\mathbb{R}^n$, and $u \in W^{k,p}(\Sigma)$, then there exists a constant $C = C(k, \Sigma)$ such that for any multi-index β with $0 < |\beta| < k$, we have

$$\|D^\beta u\|_{L^p(\Sigma)} \leq C\|u\|^{1-\frac{|\beta|}{k}}_{L^p(\Sigma)} \|u\|^{\frac{|\beta|}{k}}_{W^{k,p}(\Sigma)}.$$

This follows from [4, Theorem 7.28] after judicious choice of ε.

A first application of this estimate tells us that

$$\begin{aligned}\|\nabla(u(t) - u_\infty)\|_{L^2(S^2\setminus\cup B_i)} &\leq C\|u(t) - u_\infty\|^{\frac{1}{2}}_{L^2(S^2\setminus\cup B_i)} \|u(t) - u_\infty\|^{\frac{1}{2}}_{W^{2,2}(S^2\setminus\cup B_i)} \\ &\leq C(\delta + C\delta^{\frac{1}{2}})^{\frac{1}{2}},\end{aligned}$$

for $t \in [t_N, t_N + T)$, where we have used (3.14) and the C^2 estimate from (3.17). Therefore, provided δ is sufficiently small, we have proved (3.11), and may finally deduce that $T = \infty$. We have finished constraining δ, and may therefore consider N to be fixed for good.

Since (3.9) now holds for $t \in [t_N, \infty)$, we can permit C to depend on t_N and u (not t) and deduce that

$$E(u(t)) - \overline{E} \leq C \exp\left[-\frac{t}{C}\right],$$

for $t \geq t_N$, which implies (1.7).

Part (i) of the theorem follows from (3.13) — now we have $T = \infty$ — by setting $t = t_n$ and sending n to infinity, which yields

$$\|u_\infty - u(s)\|_{L^2(S^2)} \leq C \left[E(u(s)) - \overline{E}\right]^{\frac{1}{2}},$$

for $s \geq t_N$.

Part (ii) then involves interpolation. We are free to prove it for $\Omega = S^2 \backslash \cup B_i$ since the balls B_i could have been made smaller in order to ensure that the Ω from the theorem was a subset of $S^2 \backslash \cup B_i$. Our interpolation estimate in this case tells us that

$$\|u(t) - u_\infty\|_{W^{l,2}(S^2 \backslash \cup B_i)} \leq C \|u(t) - u_\infty\|^{\frac{1}{2}}_{L^2(S^2 \backslash \cup B_i)} \|u(t) - u_\infty\|^{\frac{1}{2}}_{W^{2l,2}(S^2 \backslash \cup B_i)},$$

which we may couple with our C^k estimate (3.17) and part (i) of the theorem to yield

$$\|u(t) - u_\infty\|_{W^{l,2}(S^2 \backslash \cup B_i)} \leq C \left[E(u(t)) - \overline{E}\right]^{\frac{1}{4}},$$

for $t \geq t_N$, which implies part (ii) of the theorem. Note that this particular part of the theorem does not hold for $t \geq 0$, since the C^k norm of $u(t)$ may become infinite sometime during the time interval $[0, t_N]$ if there is finite time blow-up.

Part (iii) combines local energy control (as in Lemma 3.3) with part (ii) and (3.10). First, note that by part (ii) (with $k = 1$) we have for any $\Omega \subset\subset S^2 \backslash \{x^1, \ldots, x^m\}$ that

$$|E(u(t), \Omega) - \limsup_{s \to \infty} E(u(s), \Omega)| = |E(u(t), \Omega) - E(u_\infty, \Omega)| \leq C |E(u(t)) - \overline{E}|^{\frac{1}{4}},$$

even for $t \geq 0$. In particular, it suffices to prove part (iii) when B is a small ball $B(x, R)$ for some bubble point x, and $R > 0$ is sufficiently small so that $\overline{B(x, 2R) \backslash B(x, R)}$ lies within $S^2 \backslash \{x^1, \ldots, x^m\}$. Let us delve into the proof of Lemma 3.3. Adopting the notation there, it suffices to prove that

$$\left| \int_{D_{2\mu}} \varphi e(u(t)) \sigma^2 - \limsup_{s \to \infty} \int_{D_{2\mu}} \varphi e(u(s)) \sigma^2 \right| \leq C |E(u(t)) - \overline{E}|^{\frac{1}{4}},$$

by virtue, again, of part (ii) of the present theorem. This follows without difficulty by integrating (3.5) and invoking (3.10) and (3.6).

There remain parts (a) and (b) of the theorem. Part (a) follows rapidly from the L^2 convergence implied by part (i), together with the fact that $E(u(t)) < M$. Further details are available in [15] if required. Part (b) is a weaker form of part (ii). □

Remark 3.4. Let us record that even in the general case of harmonic map flow from a compact domain surface to a compact target of arbitrary dimension, whenever we can prove that

$$E(u(t)) - \overline{E} \leq C \|\mathcal{T}(u(t))\|^2_{L^2},$$

for sufficiently large t and some constant C, then we can deduce uniform convergence of the flow at infinite time. Such an inequality cannot hold in general; see [15].

Acknowledgements. This work was carried out mainly at the Mathematical Sciences Research Institute in Berkeley. The author would like to thank MSRI, and especially David Hoffman and Hugo Rossi for their hospitality and support. The problem addressed in Theorem 1.7 was originally brought to the author's attention by Leon Simon. Many useful conversations on this topic with Mario Micallef, and detailed comments of the referee are also gratefully acknowledged.

Mathematics Institute, University of Warwick, Coventry, UK
E-mail address: topping@maths.warwick.ac.uk
Web address: http://www.maths.warwick.ac.uk/∼topping

References

[1] W.-Y. Ding and G. Tian, Energy identity for a class of approximate harmonic maps from surfaces, *Comm. Anal. Geom.* **3** (1995), 543–554.

[2] J. Eells and L. Lemaire, A report on harmonic maps, *Bull. London Math. Soc.* **10** (1978), 1–68.

[3] J. Eells and J. H. Sampson, Harmonic mappings of Riemannian manifold, *Amer. J. Math.* **86** (1964), 109–169.

[4] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order* (Second edition), Springer-Verlag, New York (1983).

[5] F.-H. Lin and C.-Y. Wang, Energy identity of harmonic map flows from surfaces at finite singular time, *Calc. Var. Partial Differential Equations* **6** (1998), 369–380.

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

[7] J. Qing, On singularities of the heat flow for harmonic maps from surfaces into spheres, *Comm. Anal. Geom.* **3** (1995), 297–315.

[8] J. Qing and G. Tian, Bubbling of the heat flows for harmonic maps from surface, *Comm. Pure Appl. Math.* **50** (1997), 295–310.

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

[10] G. Segal, The topology of spaces of rational functions, *Acta Math.* **143** (1979), 39–72.

[11] L. Simon, Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems, *Ann. of Math.* **118** (1983), 525–571.

[12] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Prinecton, NJ (1970).

[13] M. Struwe, On the evolution of harmonic mappings of Riemannian surfaces, *Comment. Math. Helv.* **60** (1985) 558–581.

[14] P. M. Topping, The harmonic map heat flow from surfaces, Ph.D. Thesis, University of Warwick (1996).

[15] ———, Rigidity in the harmonic map heat flow, *J. Differential Geom.* **45** (1997), 593–610.

[16] P. M. Topping, Bubbling of almost-harmonic maps between 2-spheres at points of zero energy density, preprint (2002).

[17] C.-Y. Wang, Bubble phenomena of certain Palais-Smale sequences from surfaces to general targets, *Houston J. Math.* **22** (1996), 559–590.

[18] W. P. Ziemer, *Weakly Differentiable Functions*, *Graduate Texts in Math.* **120**, Springer-Verlag, New York (1989).

(Received October 1, 1999)