ON A NONLOCAL ISOPERIMETRIC PROBLEM ON THE TWO-SPHERE

IHSAN TOPALOGLU

Abstract. In this article we analyze the minimization of a nonlocal isoperimetric problem (NLIP) posed on the 2-sphere. After establishing the regularity of the free boundary of minimizers, we characterize two critical points of the functional describing (NLIP): the single cap and the double cap. We show that when the parameter controlling the influence of the nonlocality is small, the single cap is not only stable but also is the global minimizer of (NLIP) for all values of the mass constraint. In other words, in this parameter regime, the global minimizer of the (NLIP) coincides with the global minimizer of the local isoperimetric problem on the 2-sphere. Furthermore, we show that in certain parameter regimes the double cap is an unstable critical point.

1. Introduction

There is currently much interest in the area of pattern formation for ordered structures on curved surfaces. Examples of this interest range from biology to material science: covering virus and radiolaria architecture, colloid encapsulation for possible drug delivery; and most relevant to this article, self-assembly in thin block copolymer melt films confined to the surface of a sphere (cf. [2] and references therein). There is an extensive literature on the mathematical analysis of phase separation of block copolymers and their sharp interface limit leading to a nonlocal isoperimetric problem, but to our knowledge, all of these investigations have been carried out on either the flat-tori or bounded domains in the Euclidean space (cf. [4, 5, 19, 21, 28] and references therein). Here, we have chosen to focus on the nonlocal isoperimetric problem on the two-sphere in order to present what is perhaps the first rigorous attempt to analyze a block copolymer related problem on a manifold with nonzero curvature.

We start by defining the nonlocal isoperimetric problem (NLIP) on the sphere: For fixed $m \in (-1,1)$

$$\text{minimize} \quad E_\gamma(u) := \frac{1}{2}\int_{\mathbb{S}^2} |\nabla u| + \gamma \int_{\mathbb{S}^2} |\nabla v|^2 \, d\mathcal{H}^2_x, \tag{1.1}$$

over all $u \in BV(\mathbb{S}^2, \{\pm 1\})$ satisfying

$$\frac{1}{4\pi}\int_{\mathbb{S}^2} u \, d\mathcal{H}^2_x = m$$

and v satisfying

$$-\Delta v = u - m \;\text{ on } \mathbb{S}^2 \quad \text{with} \quad \int_{\mathbb{S}^2} v \, d\mathcal{H}^2_x = 0. \tag{1.2}$$

Here $\mathbb{S}^2$ is the 2-sphere, Δ denotes the Laplace–Beltrami operator on $\mathbb{S}^2$ and $\mathcal{H}^2$ denotes the 2-dimensional Hausdorff measure. Also, the first term in E_γ is half of the total variation of u, which is defined by

$$\int_{\mathbb{S}^2} |\nabla u| := \sup\left\{ \int_{\mathbb{S}^2} u(x)\operatorname{div}X(x)\, d\mathcal{H}^2_x \;:\; X \in \mathcal{X}(\mathbb{S}^2),\, \|X\|_{L^\infty(\mathbb{S}^2)} \leqslant 1 \right\},$$

where $\mathcal{X}(\mathbb{S}^2)$ denotes the set of all vector fields of class C^∞ on $\mathbb{S}^2$ and $\mathrm{div}X$ refers to the divergence operator relative to $\mathbb{S}^2$, see e.g. [8, 10]. Then, the first term in E_γ indeed computes the perimeter of the set $\{x \in \mathbb{S}^2 : u(x) = 1\}$ since u takes on only values ± 1 and the perimeter of a set Ω in $\mathbb{S}^2$ is given by

$$\mathrm{Per}_{\mathbb{S}^2}(\Omega) := \int_{\mathbb{S}^2} |\nabla \chi_\Omega|,$$

where χ_Ω denotes the characteristic function of the set Ω. Finally, let us note that throughout this paper ∇ denotes the gradient relative to the sphere $\mathbb{S}^2$ whereas $\nabla_{\partial\Omega}$ denotes the gradient relative to the submanifold $\partial\Omega$ of a subset $\Omega \subset \mathbb{S}^2$.

The problem (NLIP) arises, up to a constant factor, as the Γ-limit as $\varepsilon \to 0$ of the well-studied Ohta–Kawasaki sequence of functionals $E_{\varepsilon,\gamma}$ which model microphase separation of diblock copolymers, [3, 20]:

$$E_{\varepsilon,\gamma}(u) := \begin{cases} \int_{\mathbb{S}^2} \frac{\varepsilon}{2}|\nabla u|^2 + \frac{(1-u^2)^2}{4\varepsilon} + \gamma\,|\nabla v|^2\,dx & \text{if } u \in H^1(\mathbb{S}^2) \\ & \text{ and } \int_{\mathbb{S}^2} u\,dx = m, \\ & \\ +\infty & \text{otherwise}, \end{cases} \tag{1.3}$$

where again v satisfies (1.2). There is an extensive literature exploring the energy landscape for $E_{\varepsilon,\gamma}$ in two and three dimensions, whether posed on the flat torus (i.e. with periodic boundary conditions) or on a general domain with homogeneous Neumann data, cf. e.g. [5, 6, 22, 24] and references therein. The picture is quite rich and complicated, with the diffuse interface sometimes bounding one or more strips, wriggled strips, discs or ovals. Again though, we are not aware of work on Ohta–Kawasaki posed on $\mathbb{S}^2$.

Independent of its connection to Ohta–Kawasaki, (NLIP) also attracts interest as a rather canonical nonlocal perturbation of the classical isoperimetric problem. Indeed, as a model for pattern formation, (NLIP) sets up a basic competition between low surface area (the perimeter term) and high oscillation (the nonlocal term) and much the same richness exists for the energy landscape of (NLIP) defined on the flat torus or on a general domain with homogeneous Neumann data. In three dimensions, computations reveal a wide array of stable critical points, with the free boundary $\partial\{x : u(x) = 1\}$ consisting of one or more pairs of parallel planes, one or more spheres, cylinders or even hypersurfaces resembling more exotic triply periodic constant mean curvature surfaces such as gyroids, depending on where in the (m, γ) parameter space one looks. With few exceptions, however, rigorous proofs of stability for particular patterns are rare (cf. e.g. [4, 25, 26, 27, 28]).

Patterns emerging by phase separation of diblock copolymers have also been investigated numerically on spherical surfaces. In [33], the authors explore the phase separation on spherical surfaces by solving the time-dependent Cahn–Hilliard equation modified for diblock copolymers using a finite volume method, whereas in [2], the authors develop a numerical method for solving the self-consistent field theory equations in spherical geometries and address detailed numerical simulations of both lamellar and hexagonal ordering of a spherical thin film of diblock copolymer. Here we have chosen to focus on the problem (1.1) defined on the two-sphere $\mathbb{S}^2$ with the hope of initiating a rigorous study of block copolymer models on curved surfaces.

This article represents a sequel to the recent work [30] on the characterization of global minimizers of (NLIP) on the flat two-torus. There we show that the minimizers are lamellar for an interval of m-values containing $m = 0$ when γ is sufficiently small. The main idea in [30] uses the Γ-convergence of E_γ and relies on the fact that the lamellar critical point is the global minimizer of the (local) periodic isoperimetric problem for the same interval of m-values when $\gamma = 0$.

In the case of the two-sphere, when $\gamma = 0$, the problem reduces the (local) isoperimetric problem and it is known that the single spherical cap, i.e., the set with boundary consisting of a single circle, is the global minimizer for any $m \in (-1, 1)$ (cf. [17, 29]). In light of this, we are able to carefully adapt the arguments in [30] and show that, for γ sufficiently small, the single cap is not only a stable critical point of E_γ but also remains as the global minimizer of the perturbed nonlocal problem (NLIP). Perhaps more interesting than its global minimality, analyzing the stability of the single cap we obtain a critical gamma value that we will denote by γ_c where the single cap is stable for all $\gamma < \gamma_c$ and it is unstable for all $\gamma > \gamma_c$. Hence, this value γ_c serves as the borderline of stability of the single cap. We also establish instability of the double cap, i.e., the set whose boundary consists of two identical parallel circles, for sufficiently small and sufficiently large γ-values. Indeed, we will show that in the $m = 0$ regime the double cap is unstable for all $\gamma > 0$. Looking at (NLIP) on a curved surface clearly requires the inclusion of a curvature term in the second variation of E_γ. However, the explicit knowledge of the Green's function for the Laplace–Beltrami operator on $\mathbb{S}^2$ enables us to carry out the calculation of the second variation of E_γ about the single cap and the double cap rather explicitly.

As in [30], for proving the global minimality of the single cap we use a "corralling" argument. After establishing the regularity of the free boundary of local minimizers we first show that if the set where a minimizer equals one is not uniformly close to that of the lamellar critical point, then necessarily, it has too much perimeter. Such a corralling step often occurs when working with sets of finite perimeter in the L^1-topology, cf. e.g. [15, 31], but here, the argument is in some ways more subtle due to the nonlocal term in the energy which prefers multi-component competitors. Next, we show that competitors that are uniformly close are in fact C^2-close. Finally, exploiting the known stability of the lamellar critical point in the sense of second variation, we eliminate competitors that are C^2-close. This scenario of converting stability to either local or global minimality in the context of (local) volume-constrained least area problems also arises in the works [11, 18].

Our arguments for both global minimality and stability/instability of critical points require the regularity of minimizers of (NLIP). Given the well-developed regularity theory for isoperimetric domains, cf. e.g. [10, 16], the issue here is to obtain a good estimate on the "excess-like" quantity that measures how far a set is from minimizing perimeter in a ball in terms of the radius of that ball, cf. (2.17). We show that even with the inclusion of the nonlocal term, it is still possible to obtain an estimate of order $\mathcal{O}(R^{1+\epsilon})$ for some $\epsilon > 0$, hence allowing us to invoke the standard theory.

The paper is organized as follows: The next section is devoted to the regularity of minimizers of (NLIP). We state the formulas for the first and second variation of E_γ in Section 3, and establish criteria for testing criticality and stability. In the fourth section we will investigate the single cap and demonstrate its stability and global minimality for any $m \in (-1, 1)$ and for γ sufficiently small. Moreover we will show that the single cap critical point is unstable for large γ-values. In Section 5 we introduce the double cap in two configurations. After a comparison of total energies of these configurations depending on the (m, γ)-regime, we prove that in both configurations the double cap is unstable for small and large values of γ. Finally, in the sixth section we discuss some open problems.

2. Regularity

In this section, we establish the regularity of the set $\partial\{x \in \mathbb{S}^2 \,:\, u(x) = 1\}$ for any local minimizer u of (NLIP). For simplicity of presentation only let us give an equivalent reformulation of (NLIP) as a functional defined on sets of finite perimeter:

$$\text{locally minimize} \quad \mathcal{E}_\gamma(A) := \int_{\mathbb{S}^2} |\nabla \chi_A| + \gamma \int_{\mathbb{S}^2} |\nabla v_A|^2 \, d\mathcal{H}^2_x \tag{2.1}$$

over $A \subset \mathbb{S}^2$ such that $\mathcal{H}^2(A) = 2\pi(1+m)$, where $-\Delta v_A = u_A - m$ and $u_A := \chi_A - \chi_{A^c}$.

We will phrase the regularity result in terms of L^1-local minimizers, by which we mean any set $\Omega \subset \mathbb{S}^2$ of finite perimeter such that

$$\mathcal{E}_\gamma(\Omega) \leqslant \mathcal{E}_\gamma(A) \quad \text{provided} \quad \int_{\mathbb{S}^2} |\chi_\Omega - \chi_A| \, d\mathcal{H}^2_x < \delta \tag{2.2}$$

for some $\delta > 0$.

The main tool in the proof of regularity is the following lemma, cf. [9, Lemma 2.1], whose proof easily adapts to the case of a two-sphere.

Lemma 2.1. *Let $L \subset \mathbb{S}^2$ be a Borel set, and let $D \subset \mathbb{S}^2$ be an open domain such that*

$$\int_D |\nabla \chi_L| > 0.$$

There exists two positive constants k_0 and l_0, depending only on D and $D \cap L$, such that for every k, $|k| < k_0$, there exists a set F, with $F = L$ outside D and

$$\begin{gathered} \mathcal{H}^2(F) = \mathcal{H}^2(L) + k, \\ \int_D |\nabla \chi_F| \leqslant \int_D |\nabla \chi_L| + l_0 |k|, \\ \int_D |\chi_F - \chi_L| \, dx \leqslant l_0 |k| \int_D |\nabla \chi_L|. \end{gathered} \tag{2.3}$$

Proof. From the definition of perimeter on a bounded subset

$$\int_D |\nabla \chi_L| = \sup \left\{ \int_D \chi_L(x) \mathrm{div} X(x) \, d\mathcal{H}^2_x \,:\, X \in \mathcal{X}_c(D), \, \|X\|_{L^\infty(D)} \leqslant 1 \right\}$$

we can conclude that there exists a vector field $X \in \mathcal{X}_c(D)$ with $\|X\|_{L^\infty(D)} \leqslant 1$ such that

$$\int_D \chi_L \mathrm{div} X \, d\mathcal{H}^2_x \geqslant \frac{1}{2} \int_D |\nabla \chi_L| > 0.$$

Here $\mathcal{X}_c(D)$ denotes the set of all compactly supported vector fields of class C^∞ on D.

Then we construct the corresponding flow deformation of L, whose boundary instantaneously moves according to the perturbation vector field X, by letting $\Psi : \mathbb{S}^2 \times (-\tau, \tau) \to \mathbb{S}^2$ solve

$$\frac{\partial \Psi}{\partial t} = X(\Psi), \qquad \Psi(x, 0) = x, \tag{2.4}$$

for some $\tau > 0$ and by defining $L_t := \Psi(L, t)$.

For t small, Ψ is a diffeomorphism. Hence, letting $J\Psi$ denote the Jacobian of Ψ in local coordinates, we have the following change of variables formulas

$$\begin{gathered} \mathcal{H}^2(L_t) = \int_L J\Psi \, d\mathcal{H}^2_x, \\ \int_D |\nabla \chi_{L_t}| = \int_D |J\Psi (D\Psi)^{-1} \eta| |\nabla \chi_L|, \end{gathered} \tag{2.5}$$

where $\eta \in L^1(\mathbb{S}^2;\mathbb{S}^2)$ is obtained by differentiating the vector-valued measure $\nabla\chi_L$ with respect to the measure $|\nabla\chi_L|$ (cf. [10, Lemma 10.1]). On the other hand, expanding in t we find that

$$\begin{aligned} J\Psi = \det(D\Psi(x,t)) &= \det\left(I + t\nabla X + \frac{1}{2}t^2\nabla Z + o(t^2)\right) \\ &= 1 + t\mathrm{div}X + \frac{1}{2}t^2 K(x,t) + o(t^2), \end{aligned} \tag{2.6}$$

where $Z := \left.\dfrac{\partial^2\Psi}{\partial t^2}\right|_{t=0}$ has i-th component given by $Z^{(i)} = \sum_j X^{(i)}_{x_j} X^{(j)}$ and K is defined by

$$K(x,t) := \mathrm{trace}\nabla Z + (\mathrm{trace}\nabla X)^2 - \mathrm{trace}(\nabla X)^2.$$

Moreoever,

$$(D\Psi)^{-1} = I - tH(x,t)$$

for some H, where both $|K|$ and $|H|$ are bounded by a constant depending only on $\sup|\nabla X|$, and therefore only on D and $D\cap L$. With these, the proof of the lemma follows as shown in [9, Lemma 2.1]. □

We now state the regularity result for local minimizers on $\mathbb{S}^2$:

Proposition 2.2. *Let Ω be an L^1-local minimizer of* (2.1). *Then $\partial\Omega$ is of class $C^{3,\alpha}$ for some $\alpha \in (0,1)$.*

Proof. Let Ω be an L^1-local minimizer of $\mathcal{E}_\gamma$ and let x_0 be any point of $\partial\Omega$. Then let $D \subset\subset U$ be such that $x_0 \notin \overline{D}$ and

$$\int_D |\nabla\chi_\Omega| > 0.$$

By Lemma 2.1, there exist two positive constants k_0 and l_0, depending only on D and $D\cap\Omega$, such that for every k, $|k| < k_0$, there exists a set F, with $F = \Omega$ outside D and satisfying (2.3).

Fix $R > 0$ such that

$$\omega_2 R^2 < k_0, \quad \left(1 + l_0\int_D |\nabla\chi_\Omega|\right)\omega_2 R^2 < \delta \quad \text{and} \quad \overline{B}_R(x_0)\cap\overline{D} = \emptyset, \tag{2.7}$$

where ω_2 is the measure of the unit geodesic ball on $\mathbb{S}^2$ and δ comes from (2.2). Moreover, let $\tilde{F}$ minimize perimeter in $B_R(x_0)$ subject to the boundary values of Ω, i.e.,

$$\int_{B_R(x_0)} |\nabla\chi_{\tilde{F}}| \leqslant \int_{B_R(x_0)} |\nabla\chi_F|$$

for all F such that $F\setminus B_R(x_0) = \Omega\setminus B_R(x_0)$. Without loss of generality, we can assume that $|\tilde{F}\cap B_R(x_0)| \leqslant |\Omega\cap B_R(x_0)|$. Since $\tilde{F}\cap\overline{D} = \Omega\cap\overline{D}$, we can use the same k_0 and l_0 as above with $\tilde{F}$ replacing Ω in (2.3). Hence, for $k := |\Omega| - |\tilde{F}| \leqslant \omega_2 R^2 < k_0$, there exists a set G, with $G = \tilde{F}$ outside D, and

$$\mathcal{H}^2(G) = \mathcal{H}^2(\Omega) = m, \tag{2.8}$$

$$\int_D |\nabla\chi_G| \leqslant \int_D |\nabla\chi_{\tilde{F}}| + CR^2, \tag{2.9}$$

$$\int_{\mathbb{S}^2} |\chi_G - \chi_\Omega|\, d\mathcal{H}^2_x \leqslant C_0 R^2 < \delta, \tag{2.10}$$

where the last condition follows from (2.7) with $C_0 := \left(1 + l_0\int_D |\nabla\chi_\Omega|\right)\omega_2$.

By (2.8), G is a competitor in (2.1) and we have that

$$\int_{\mathbb{S}^2} |\nabla\chi_\Omega| + \gamma \int_{\mathbb{S}^2} |\nabla v_\Omega|^2 \, d\mathcal{H}^2_x \leqslant \int_{\mathbb{S}^2} |\nabla\chi_G| + \gamma \int_{\mathbb{S}^2} |\nabla v_G|^2 \, d\mathcal{H}^2_x. \tag{2.11}$$

Thus, using the fact $\tilde{F} \setminus B_R(x_0) = \Omega \setminus B_R(x_0)$ and $G \setminus D = \tilde{F} \setminus D$, along with (2.9), the inequality (2.11) becomes

$$\begin{aligned}
\int_{\mathbb{S}^2 \setminus (D \cup B_R(x_0))} |\nabla\chi_\Omega| + \int_D |\nabla\chi_{\tilde{F}}| + \int_{B_R(x_0)} |\nabla\chi_\Omega| + \gamma \int_{\mathbb{S}^2} |\nabla v_\Omega|^2 \, d\mathcal{H}^2_x \qquad & \\
\leqslant \int_{\mathbb{S}^2 \setminus (D \cup B_R(x_0))} |\nabla\chi_G| + \int_D |\nabla\chi_G| \qquad & \\
+ \int_{B_R(x_0)} |\nabla\chi_G| + \gamma \int_{\mathbb{S}^2} |\nabla v_G|^2 \, d\mathcal{H}^2_x & \\
= \int_{\mathbb{S}^2 \setminus (D \cup B_R(x_0))} |\nabla\chi_\Omega| + \int_D |\nabla\chi_G| \qquad & \\
+ \int_{B_R(x_0)} |\nabla\chi_{\tilde{F}}| + \gamma \int_{\mathbb{S}^2} |\nabla v_G|^2 \, d\mathcal{H}^2_x & \\
\leqslant \int_{\mathbb{S}^2 \setminus (D \cup B_R(x_0))} |\nabla\chi_\Omega| + \int_D |\nabla\chi_{\tilde{F}}| \qquad & \\
+ \int_{B_R(x_0)} |\nabla\chi_{\tilde{F}}| + \gamma \int_{\mathbb{S}^2} |\nabla v_G|^2 \, d\mathcal{H}^2_x + CR^2. &
\end{aligned}$$

Hence, we get

$$\begin{aligned}
\int_{B_R(x_0)} |\nabla\chi_\Omega| - \int_{B_R(x_0)} |\nabla\chi_{\tilde{F}}| \leqslant \gamma \int_{\mathbb{S}^2} |\nabla v_G|^2 \, d\mathcal{H}^2_x & \\
- \gamma \int_{\mathbb{S}^2} |\nabla v_\Omega|^2 \, d\mathcal{H}^2_x + CR^2. &
\end{aligned} \tag{2.12}$$

Now we estimate the nonlocal parts on the right-hand side of (2.12). To this end, let $w := v_\Omega - v_G$. Then $-\Delta w = u_\Omega - u_G$ with $\int_{\mathbb{S}^2} w \, d\mathcal{H}^2_x = 0$, where $|u_\Omega - u_G|$ is equal to zero in $\mathbb{S}^2 \setminus (B_R(x_0) \cup D)$ and is bounded by 2 in $B_R(x_0) \cup D$. Hence for any $p \geqslant 1$ we have

$$\|u_\Omega - u_G\|_{L^p(\mathbb{S}^2)} \leqslant CR^{2/p}, \tag{2.13}$$

through an appeal to (2.10).

We take

$$p = 2\kappa, \tag{2.14}$$

where κ is less than but as close as needed to 1 so that $1 < p < 2$. Since H^1 imbeds continuously into L^q for any $q < \infty$, using the Poincaré and Hölder inequalities we get that

$$\|w\|_{L^1(\mathbb{S}^2)} \leqslant C\|w\|_{H^1(\mathbb{S}^2)} \leqslant C\|u_\Omega - u_G\|_{L^p(\mathbb{S}^2)}. \tag{2.15}$$

Thus, by combining (2.13)–(2.15) we obtain

$$\|w\|_{L^1(\mathbb{S}^2)} \leqslant C\, R^{1/\kappa},$$

or in other words, since $\kappa < 1$, we have that

$$\int_{\mathbb{S}^2} |v_\Omega - v_G| \, d\mathcal{H}^2_x = \int_{\mathbb{S}^2} |w| \, d\mathcal{H}^2_x \leqslant C\, R^{1+\epsilon} \tag{2.16}$$

for some $\epsilon > 0$.

Now, using (2.10), (2.16) and integration by parts, we obtain the following bound on the difference of the nonlocal parts:

$$\begin{aligned}\int_{\mathbb{S}^2} |\nabla v_G|^2 \, d\mathcal{H}^2_x - \int_{\mathbb{S}^2} |\nabla v_\Omega|^2 \, d\mathcal{H}^2_x \leqslant \int_{\mathbb{S}^2} |u_\Omega - u_G| \, |v_G| \, d\mathcal{H}^2_x \\ + \int_{\mathbb{S}^2} |v_\Omega - v_G| \, |u_\Omega| \, d\mathcal{H}^2_x \\ \leqslant C\, R^{1+\epsilon}.\end{aligned}$$

Returning to (2.12), this implies that

$$\int_{B_R(x_0)} |\nabla \chi_\Omega| - \int_{B_R(x_0)} |\nabla \chi_{\tilde{F}}| \leqslant C\, R^{1+\epsilon}. \tag{2.17}$$

Property (2.17) states that the boundary of the set Ω is almost area-minimizing in any ball. With this property in hand, the results of [14, 32] apply, and we can conclude that $\partial\Omega$ is of class $C^{1,\alpha}$. The $C^{3,\alpha}$ regularity of $\partial\Omega$ then follows from standard elliptic theory. □

3. The first and second variations of (NLIP)

In this section, we will characterize the first and second variations of (NLIP). To fix notation, let us express the functional E_γ as

$$E_\gamma(u) := P(u) + \gamma N(u),$$

where $P : L^1(\mathbb{S}^2) \to \mathbb{R}$ is defined by

$$P(u) := \begin{cases} \frac{1}{2}\int_{\mathbb{S}^2} |\nabla u| & \text{if } u \in BV(\mathbb{S}^2; \pm 1) \text{ and } \frac{1}{4\pi}\int_{\mathbb{S}^2} u \, d\mathcal{H}^2_x = m, \\ \\ +\infty & \text{otherwise,} \end{cases} \tag{3.1}$$

and $N : L^1(\mathbb{S}^2) \to \mathbb{R}$ denotes the functional

$$N(u) := \begin{cases} \int_{\mathbb{S}^2} |\nabla v|^2 \, d\mathcal{H}^2_x & \text{if } u \in BV(\mathbb{S}^2; \pm 1) \text{ and } \frac{1}{4\pi}\int_{\mathbb{S}^2} u \, d\mathcal{H}^2_x = m, \\ \\ +\infty & \text{otherwise,} \end{cases} \tag{3.2}$$

where $v : \mathbb{S}^2 \to \mathbb{R}$ depends on u as the solution to the problem (1.2).

First, let us note that we can write v in terms of the Green's function $G = G(x, y)$ associated with the Poisson problem (1.2). Then, for each $x \in \mathbb{S}^2$, $G(x, y)$ satisfies

$$-\Delta_y G(x,y) = \delta_x - \frac{1}{4\pi} \quad \text{on } \mathbb{S}^2, \qquad \int_{\mathbb{S}^2} G(x,y) \, d\mathcal{H}^2_x = 0,$$

where δ_x is a delta-mass measure supported at x. In particular, one can show, by writing out the Laplace–Beltrami operator in spherical coordinates explicitly, that for $x, y \in \mathbb{S}^2$

$$G(x,y) = -\frac{1}{2\pi} \log|x - y|, \tag{3.3}$$

where $|\cdot|$ denotes the Euclidean norm, that is, $|x-y|$ is the chordal distance between x and y in $\mathbb{R}^3$. The functions G and v are then related by

$$v(x) = -\frac{1}{2\pi} \int_{\mathbb{S}^2} \log(|x-y|) u(y) \, d\mathcal{H}^2_y. \tag{3.4}$$

The first and second variations of P alone computed about a critical point of P have been investigated in [1] by calculating the first and second variations of area subject to fixed volume. However, for the first and second variations of E_γ, one needs to proceed as the authors did in [7], that is, one needs to compute the second variation not about a critical point of P but rather about a critical point of E_γ. To this end, in light of the reformulation given by (2.1), we view E_γ as a functional depending on a set, say $A \subset \mathbb{S}^2$, through the formula

$$u(x) = \begin{cases} 1 & \text{if } x \in A, \\ -1 & \text{if } x \in A^c. \end{cases} \tag{3.5}$$

Here A^c denotes the complement of A, i.e., $A^c = \mathbb{S}^2 \setminus A$, and the 2-dimensional measure of A, $\mathcal{H}^2(A)$, is compatible with the mass constraint on u, so that $|A| = 2\pi(1+m)$.

Given a set $A \subset \mathbb{S}^2$ with C^2 boundary, we define an admissible perturbation of A as a family of sets $\{A_t\}_{t\in(-\tau,\tau)}$ for some $\tau > 0$ so that the sets A_t preserve measure to second order, i.e., $\mathcal{H}^2(A_t) = \mathcal{H}^2(A) + o(t^2)$. In [7], to construct the admissible family $\{A_t\}$, the authors use the ODE given by (2.4) to produce a flow deformation of A whose boundary instantaneously moves according to a perturbation vector field X, and which instantaneously preserves measure. They then apply a correction to insure it instantaneously preserves measure to second order as well. Indeed, in the proof of the following proposition, the perturbation vector field $X = X_f$ is chosen as $X_f = f\,\nu$ on ∂A, for a given smooth function f satisfying the condition $\int_{\partial A} f(x) d\mathcal{H}^1_x = 0$, where ν denotes the outer normal to ∂A that is tangent to $\mathbb{S}^2$.

Now, following the calculations in [7] and replacing the calculations for the local part with the calculations in [1], one easily obtains the following result.

Proposition 3.1. *Let u be a critical point of E_γ given by* (3.5) *such that ∂A is C^2. Let f be any smooth function on ∂A satisfying the condition*

$$\int_{\partial A} f(x) d\mathcal{H}^1_x = 0.$$

Then for v solving (1.2) *we have*

$$H(x) + 4\gamma\, v(x) = \lambda \quad \text{for all } x \in \partial A, \tag{3.6}$$

where λ is a constant and H denotes the geodesic curvature of ∂A.

Moreoever, the second variaton of E_γ about the critical point u is given by

$$\begin{aligned} J(f) := &\int_{\partial A} |\nabla_{\partial A} f|^2 - (1 + \|B_{\partial A}\|^2) f^2 \, d\mathcal{H}^1_x \\ &+ 8\gamma \int_{\partial A}\int_{\partial A} \left(-\frac{1}{2\pi}\log(|x-y|)\right) f(x) f(y)\, d\mathcal{H}^1_x d\mathcal{H}^1_y \\ &+ 4\gamma \int_{\partial A} (\nabla v \cdot \nu) f^2 d\mathcal{H}^1_x. \end{aligned} \tag{3.7}$$

Here $\nabla_{\partial A} f$ denotes the gradient of f relative to the manifold ∂A, $B_{\partial A}$ denotes the second fundamental form of ∂A and ν denotes the unit tangent of $\mathbb{S}^2$ which is normal to ∂A pointing out of A.

We finish this section with three remarks.

Remark. Here we want to note that if the critical point u of E_γ is a local minimizer then Proposition 2.2 applies and we obtain the regularity of ∂A as stated in the hypothesis of the previous proposition.

Remark. In reference to the formula for the second variation above, a critical point u of E_γ is *stable* if $J(f) \geqslant 0$ for all smooth f on ∂A satisfying the condition

$$\int_{\partial A} f(x) d\mathcal{H}^1_x = 0.$$

Remark. We want to note that since ∂A is a curve on $\mathbb{S}^2$, the term $\|B_{\partial A}\|^2$ in (3.7) involving the second fundamental form of ∂A coincides with the square of the geodesic curvature of ∂A, namely, $\|B_{\partial A}\|^2 = H^2$.

4. Single Cap: Stability and Global Minimality

In this section we will analyze the single cap, that is, the set whose boundary consists of a single circle. By rotational symmetries of the sphere we can assume that the boundary of the single cap is parallel to the equator, that is, using spherical coordinates we can assume that the boundary of the single spherical cap is the circle parametrized as $(1, \theta, \phi_0)$ for $\theta \in [0, 2\pi]$ and for some fixed $\phi_0 \in (0, \pi)$. Let us fix some $m \in (-1, 1)$ and define the function u_S describing the single cap as follows:

$$u_S(\phi) = \begin{cases} 1 & \text{if } \phi \in [0, \phi_0], \\ -1 & \text{if } \phi \in (\phi_0, \pi]. \end{cases} \tag{4.1}$$

With this definition, the single cap then is the set

$$S := \{(\cos\theta \sin\phi, \sin\theta\sin\phi, \cos\phi) \in \mathbb{S}^2 : \, u_S(\phi) = 1\}.$$

Also note that, since $\frac{1}{4\pi}\int_{\mathbb{S}^2} u \, d\mathcal{H}^2_x = m$, we get that $\phi_0 = \arccos(-m)$.

Before we start investigating the second variation of E_γ about u_S, we remark that since the function v_S corresponding to u_S through (1.2) clearly depends only on ϕ, one easily sees from (3.6) that u_S is a critical point of E_γ for all $\gamma > 0$.

Now, writing out the Laplace operator in spherical coordinates, plugging into (1.2) and using the fact that v_S' is continuous at ϕ_0 and stays bounded as ϕ approaches 0 and π, we obtain an explicit formula for v_S':

$$v_S'(\phi) = \begin{cases} (1-m)\cot\phi - \frac{1-m}{\sin\phi} & \text{if } \phi \in [0, \phi_0], \\ -(1+m)\cot\phi - \frac{1+m}{\sin\phi} & \text{if } \phi \in (\phi_0, \pi]. \end{cases} \tag{4.2}$$

Since $\phi_0 = \arccos(-m)$, we have that $\|B_{\partial S}\|^2 = \cot^2\phi_0 = \frac{m^2}{1-m^2}$; hence

$$1 + \|B_{\partial S}\|^2 = \frac{1}{1-m^2}.$$

Also, from (4.2) we get that $(\nabla v_S \cdot \nu)|_{\partial S} = v_S'(\phi_0) = \frac{m^2-1}{\sqrt{1-m^2}}$. Thus referring to (3.7), for any smooth function f satisfying $\int_{\partial S} f \, d\mathcal{H}^1_x = 0$, the second variation of E_γ about u_S takes the form

$$\begin{aligned} J_\gamma(f) = & \int_{\partial S} |\nabla_{\partial S} f|^2 - \left(\frac{1}{1-m^2}\right) f^2 \, d\mathcal{H}^1_x \\ & + 8\gamma \int_{\partial S}\int_{\partial S} \left(-\frac{1}{2\pi}\log(|x-y|)\right) f(x) f(y) \, d\mathcal{H}^1_x d\mathcal{H}^1_y \\ & - 4\gamma \frac{1-m^2}{\sqrt{1-m^2}} \int_{\partial S} f^2 \, d\mathcal{H}^1_x. \end{aligned} \tag{4.3}$$

Note that ∂S is a circle of radius $\sqrt{1-m^2}$; hence expressing it in polar coordinates $(\sqrt{1-m^2}\cos\theta, \sqrt{1-m^2}\sin\theta)$ for $\theta \in [0, 2\pi]$ and invoking the identity

$$\log\left([(\cos\theta-\cos\alpha)^2+(\sin\theta-\sin\alpha)^2]^{1/2}\right)=\frac{1}{2}\log(2-2\cos(\theta-\alpha))$$

J_γ then becomes

$$\begin{aligned}
J_\gamma(f) = \sqrt{1-m^2}\int_0^{2\pi} \frac{1}{1-m^2}(f')^2 - \frac{1}{1-m^2}f^2\,d\theta \\
- 8\gamma(1-m^2)\left(\frac{1}{2\pi}\int_0^{2\pi}\int_0^{2\pi}\frac{1}{2}\log(2-2\cos(\theta-\alpha))f(\theta)f(\alpha)\,d\theta d\alpha\right) \\
- 4\gamma(1-m^2)\int_0^{2\pi} f^2\,d\theta.
\end{aligned} \tag{4.4}$$

Now, referring to (4.4), we can prove the following stability/instability result for the single cap S. Indeed this proposition provides a γ-value which defines the borderline for the stability of the single cap solution.

Proposition 4.1. *For any $m \in (-1,1)$, there exists a value γ_c depending only on m such that the function u_S defined as in (4.1) is a stable critical point of E_γ for all $\gamma < \gamma_c$. Moreover u_S is unstable for all values $\gamma > \gamma_c$.*

Proof. First, let us note that for any $\theta \neq \alpha$ we have the following identity (cf. page 190 in [34]):

$$\begin{aligned}
\frac{1}{2}\log(2-2\cos(\theta-\alpha)) \\
= -\left(\cos(\theta-\alpha)+\frac{1}{2}\cos(2\theta-2\alpha)+\frac{1}{3}\cos(3\theta-3\alpha)+\cdots\right).
\end{aligned} \tag{4.5}$$

Also note that the singularity of the logarithm at $\theta = \alpha$ is weak enough not to disturb integrability. Hence, in light of the orthonormality of $\{\sin(n\theta), \cos(n\theta)\}_{n\in\mathbb{N}}$, we get that

$$\begin{aligned}
-\frac{1}{2\pi}\int_0^{2\pi}\int_0^{2\pi}\frac{1}{2}\log(2-2\cos(\theta-\alpha))\sin(n\theta)\sin(m\alpha)\,d\theta d\alpha = \frac{\pi}{2n}\delta_{nm}, \\
-\frac{1}{2\pi}\int_0^{2\pi}\int_0^{2\pi}\frac{1}{2}\log(2-2\cos(\theta-\alpha))\cos(n\theta)\cos(m\alpha)\,d\theta d\alpha = \frac{\pi}{2n}\delta_{nm}, \\
-\frac{1}{2\pi}\int_0^{2\pi}\int_0^{2\pi}\frac{1}{2}\log(2-2\cos(\theta-\alpha))\sin(n\theta)\cos(m\alpha)\,d\theta d\alpha = 0,
\end{aligned} \tag{4.6}$$

where δ_{nm} denotes the Kronecker delta.

As mentioned above, in [7] the second variation of E_γ in Euclidean domains is calculated along a perturbation vector field X. Here, we take $X = X_f$ such that

$$X_f(\theta) = f(\theta)\nu(\theta) \qquad \text{on } \partial S,$$

where ν denotes the outer normal to ∂S that is tangent to $\mathbb{S}^2$ and $\int_0^{2\pi} f(\theta)\,d\theta = 0$. Let us define $g(\theta) := f(\theta) - (a_1\cos\theta + b_1\sin\theta)$, where a_1 and b_1 are the first Fourier coefficients in the expansion of f. Clearly, $\int_0^{2\pi} g(\theta)\,d\theta = 0$. Let us also note that the perturbation of the boundary of the single cap by a vector field of the form $(a_1\cos\theta + b_1\sin\theta)\nu(\theta)$ corresponds to a rotation (cf. [12, Section 4.2]). We will now argue that $J_\gamma(f) > 0$ for $\gamma < \gamma_c$ for some $\gamma_c > 0$, provided $g \not\equiv 0$, that is, provided f does not correspond to a rotation.

Given any smooth function f satisfying $\int_0^{2\pi} f\, d\theta = 0$, let us look at its Fourier expansion given by

$$f(\theta) = \sum_{n=1}^{\infty} a_n \cos(n\theta) + b_n \sin(n\theta) \qquad \text{for } \theta \in [0, 2\pi].$$

Let

$$\gamma_c := \frac{3}{2(1-m^2)^{3/2}}.$$

Plugging f into (4.4) and using (4.6) then yields

$$\begin{aligned} J_\gamma(f) &= \frac{\pi}{\sqrt{1-m^2}} \sum_{n=1}^{\infty} n^2(a_n^2 + b_n^2) - \frac{\pi}{\sqrt{1-m^2}} \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \\ &\qquad + 4\gamma(1-m^2)\pi \sum_{n=1}^{\infty} \frac{1}{n}(a_n^2 + b_n^2) - 4\gamma(1-m^2)\pi \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \\ &= \pi \sum_{n=1}^{\infty} \left[\frac{n^2-1}{\sqrt{1-m^2}} + 4\gamma \left(\frac{1-m^2}{n} - (1-m^2) \right) \right] (a_n^2 + b_n^2) \\ &= \pi \sum_{n=1}^{\infty} \left[\frac{4(n-1)(1-m^2)}{n} \left(\frac{(n+1)n}{4(1-m^2)^{3/2}} - \gamma \right) \right] (a_n^2 + b_n^2) \\ &\geqslant \delta \sum_{n=2}^{\infty} (a_n^2 + b_n^2), \end{aligned}$$

where $\delta := 2\pi(1-m^2)(\gamma_c - \gamma)$. For any $\gamma < \gamma_c$, this implies that

$$J_\gamma(f) > 0;$$

hence, u_S is stable for $\gamma < \gamma_c$.

To establish the instability of u_S let us consider the function $f_2(\theta) := \sin 2\theta$ on $[0, 2\pi]$. Clearly $\int_{\partial S} f_2\, d\mathcal{H}^1_x = 0$. Plugging f_2 into (4.4) and invoking (4.6) yields

$$\begin{aligned} J_\gamma(f_2) &= \frac{3\pi}{\sqrt{1-m^2}} + 2\gamma(1-m^2)\pi - 4\gamma(1-m^2)\pi \\ &= \pi \left[\frac{3}{\sqrt{1-m^2}} - 2\gamma(1-m^2) \right]. \end{aligned}$$

Then, we get that $J_\gamma(f_2) < 0$ for all $\gamma > \gamma_c$, implying the instability of u_S for γ sufficiently large. $\square$

We will now proceed to prove that for γ sufficiently small, the global minimizer u_γ of E_γ coincides with u_S, the single cap.

We first note that we can immediately conclude that any sequence of minimizers $\{u_\gamma\}$ of (1.1) converges, after perhaps a rotation, to the single cap u_S given by (4.1).

Proposition 4.2. *For any $m \in (-1, 1)$, let $\{u_\gamma\}_{\gamma \geqslant 0}$ be a sequence of minimizers of E_γ. Then after perhaps a rotation,*

$$u_\gamma \to u_L \quad \text{in } L^1(\mathbb{S}^2) \text{ as } \gamma \to 0. \tag{4.7}$$

Proof. Since a uniform bound $E_\gamma(u_\gamma) < C$ is immediate in light of the minimality of u_γ, one obtains a uniform BV-bound leading to convergence in L^1 of a subsequence as $\gamma \to 0$. By the standard Γ-convergence argument, this limit must minimize P, the local isoperimetric problem, defined in (3.1). On the sphere, $\mathbb{S}^2$, however, it is known that the global minimizer of P is a single cap for any $m \in (-1,1)$, hence is equal to u_S given by (4.1), after perhaps a rotation, cf. [29, Theorem 3.3]. □

With this in hand, one can also easily establish convergence of the functions v_γ to v_S as in [30, Proposition 3.2].

Proposition 4.3. *For any minimizer u_γ of E_γ, there is a value $\alpha \in (0,1)$ such that the corresponding solution v_γ of (1.2) is bounded in $C^{1,\alpha}$ independent of γ.*

Moreover, for a sequence of minimizers $\{u_\gamma\}_{\gamma\geqslant 0}$ of E_γ satisfying (4.7) we have

$$v_\gamma \to v_S \quad \text{in} \quad H^2(\mathbb{S}^2).$$

In particular, $\displaystyle\int_{\mathbb{S}^2} |\nabla v_\gamma|^2\, d\mathcal{H}^2_x \to \int_{\mathbb{S}^2} |\nabla v_S|^2\, d\mathcal{H}^2_x$ *as* $\gamma \to 0$.

We now can state our result on the global minimality of the single cap.

Theorem 4.4. *Fix any $m \in (-1,1)$. Then for sufficiently small $\gamma > 0$, the minimizers $\{u_\gamma\}$ of E_γ are single caps, that is, $u_\gamma \equiv u_S$ up to rotation.*

Proof. We will prove the theorem in several steps. Let $m \in (-1,1)$ be fixed. Throughout the proof, we denote by S the single cap $\{(\theta,\phi) : 0 \leqslant \theta \leqslant 2\pi,\ 0 < \phi \leqslant \phi_0\} = \{x \in \mathbb{S}^2 : u_S = 1\}$ and by Ω_γ the set $\{x : u_\gamma(x) = 1\}$.

Step 1. We first claim that there cannot exist a sequence of components $S^1_\gamma \subset \Omega_\gamma$ whose area converges to zero as $\gamma \to 0$.

To this end, we write Ω_γ as a union of its connected components, i.e. $\Omega_\gamma = \bigcup_{j=1}^{N_\gamma} S^j_\gamma$. We first note that necessarily, $N_\gamma < \infty$ since otherwise for fixed γ there would have to exist a sequence of components of Ω_γ whose area (and perimeter) approach zero. This would be impossible in light of (3.6) and Proposition 4.3 which imply a (γ-dependent) bound on the L^∞-norm of the curvature H_γ of $\partial\Omega_\gamma$.

Now we assume, by way of contradiction, that for a sequence of $\gamma-$values approaching zero, Ω_γ has a component, say S^1_γ, with $\mathcal{H}^2(S^1_\gamma) \to 0$.

Define $S_\gamma := \Omega_\gamma \setminus S^1_\gamma$. Then $\chi_{S_\gamma} \to \chi_S$ in L^1. Also, note that $\mathrm{Per}_{\mathbb{S}^2}(S^1_\gamma) \to 0$, for if not, that is, if say $c := \liminf_{\gamma\to 0} \mathrm{Per}_{\mathbb{S}^2}(S^1_\gamma)$ with $c > 0$, then since $\liminf \frac{1}{2}\mathrm{Per}_{\mathbb{S}^2}(S_\gamma) \geqslant \frac{1}{2}\mathrm{Per}_{\mathbb{S}^2}(S) = \pi\sqrt{1-m^2}$, we get that

$$\liminf_{\gamma\to 0} E_\gamma(u_\gamma) \geqslant \pi\sqrt{1-m^2} + \frac{c}{2} > \pi\sqrt{1-m^2}.$$

This yields a contradiction to the fact that $E_\gamma(u_\gamma) \to P(u_S) = \frac{1}{2}\mathrm{Per}_{\mathbb{S}^2}(S) = \pi\sqrt{1-m^2}$ by Γ-convergence.

Now, the regularity of $\partial\Omega_\gamma$ asserted in Proposition 2.2 and the fact that $\mathrm{Per}_{\mathbb{S}^2}(S^1_\gamma) \to 0$ imply that we can enclose S^1_γ in a geodesic disk whose radius approaches zero with γ. Shrink the disk until it touches ∂S^1_γ for the first time, and denote the radius of the shrunken disk by r_γ and the point where the disk touches ∂S^1_γ by p_γ. Then we have $H_\gamma(p_\gamma) \geqslant \frac{1}{r_\gamma}$ and so by evaluating the criticality condition (3.6) at $x = p_\gamma$, we see that $\lambda_\gamma \to \infty$ since $\|v_\gamma\|_{L^\infty}$ is bounded independent of γ by Proposition 4.3. Returning to (3.6) for $x \neq p_\gamma$, we conclude that in fact $H_\gamma(x) \to \infty$ for all $x \in \partial\Omega_\gamma$. Moreover, since $H_\gamma(p_\gamma) \geqslant \frac{1}{r_\gamma}$, for γ small enough, we have that, say, $H_\gamma(x) \geqslant \frac{1}{4r_\gamma}$ for all $x \in \partial\Omega_\gamma$ so S^j_γ is contained in a geodesic disk of radius $2r_\gamma$ for each $j \in \{1,\dots,N_\gamma\}$.

For a finer analysis, let $\rho_\gamma^j := \operatorname{diam}(S_\gamma^j)$. Then S_γ^j is contained in a geodesic disk with radius ρ_γ^j. Now, define $\rho_\gamma := \min\{\rho_\gamma^j : j \in \{1, \dots, N_\gamma\}\}$ so that

$$\operatorname{Per}_{\mathbb{S}^2}(S_\gamma^j) \geqslant \rho_\gamma^j \geqslant \rho_\gamma. \tag{4.8}$$

Then let the minimum ρ_γ be attained at, say, $j = j_0$, i.e., ρ_γ is the distance between two points $p_\gamma^{j_0}$, $q_\gamma^{j_0}$ on $\partial S_\gamma^{j_0}$. As $S_\gamma^{j_0}$ is contained in a disk of radius ρ_γ which must be tangent to $\partial S_\gamma^{j_0}$, say, at $p_\gamma^{j_0}$, we see that $H_\gamma(p_\gamma^{j_0}) \geqslant \frac{1}{\rho_\gamma}$. Hence, using the L^∞-bound on v_γ and the criticality condition, we get that

$$\frac{1}{\rho_\gamma} - C_\gamma \leqslant H_\gamma(p_\gamma^{j_0}) + 4\gamma v_\gamma(p_\gamma^{j_0}) = \lambda_\gamma,$$

where C_γ depends only on γ and $\|v_\gamma\|_{L^\infty}$, and C_γ is $\mathcal{O}(\gamma)$. Thus at any point $x \in \partial\Omega_\gamma$ we have

$$H_\gamma(x) + C_\gamma \geqslant H_\gamma(x) + 4\gamma v_\gamma(x) = \lambda_\gamma \geqslant \frac{1}{\rho_\gamma} - C_\gamma,$$

and this gives that

$$H_\gamma(x) \geqslant \frac{1}{\rho_\gamma} - 2C_\gamma \geqslant \frac{1}{2\rho_\gamma}.$$

Thus for any $j \in \{1, \dots, N_\gamma\}$, S_γ^j is contained in a geodesic disk of radius $2\rho_\gamma$. Using this fact we can find a lower bound on N_γ depending on ρ_γ as follows: Since $2\pi(1+m) = \mathcal{H}^2(\Omega_\gamma) = \sum_{j=1}^{N_\gamma} |S_\gamma^j| \leqslant 2\pi(1 - \cos 2\rho_\gamma)N_\gamma$, we get that

$$N_\gamma \geqslant \frac{1+m}{1 - \cos 2\rho_\gamma}.$$

This lower bound on N_γ with (4.8) will then imply that

$$\operatorname{Per}_{\mathbb{S}^2}(\Omega_\gamma) = \sum_{j=1}^{N_\gamma} \operatorname{Per}_{\mathbb{S}^2}(S_\gamma^j) \geqslant \rho_\gamma N_\gamma \geqslant \frac{(1+m)\rho_\gamma}{(1 - \cos 2\rho_\gamma)}.$$

Hence $\operatorname{Per}_{\mathbb{S}^2}(S_\gamma) \to \infty$ as $\gamma \to 0$, which contradicts the fact that $E_\gamma(u_\gamma) \to \pi\sqrt{1-m}$.

Here we want to remark that the above argument also shows that there cannot be a sequence of components of the complement of Ω_γ approaching zero in measure.

Step 2. We claim that Ω_γ consists of precisely one component.

Suppose, for a contradiction, that Ω_γ has at least two components. Without loss of generality, assume that Ω_γ has two components, say $\Omega_\gamma = S^1 \cup S^2$. The proof generalizes easily to the case where Ω_γ has three or more components. Note that, by the mass constraint $\frac{1}{4\pi}\int_{\mathbb{S}^2} u \, d\mathcal{H}_x^2 = m$, we have that

$$\mathcal{H}^2(S^1) + \mathcal{H}^2(S^2) = 2\pi(m+1).$$

Also, by the Bol–Fiala inequality, which is a generalization of the isoperimetric inequality (cf. [29]), for $i = 1, 2$, we have that

$$\operatorname{Per}_{\mathbb{S}^2}^2(S^i) \geqslant 4\pi\mathcal{H}^2(S^i) - (\mathcal{H}^2(S^i))^2.$$

Hence, we get that

$$\begin{aligned}(\mathrm{Per}_{\mathbb{S}^2}(S^1)+\mathrm{Per}_{\mathbb{S}^2}(S^2))^2 &\geqslant \mathrm{Per}^2_{\mathbb{S}^2}(S^1)+\mathrm{Per}^2_{\mathbb{S}^2}(S^2)\\ &\geqslant 4\pi\left(\mathcal{H}^2(S^1)+\mathcal{H}^2(S^2)\right)-\left((\mathcal{H}^2(S^1))^2+(\mathcal{H}^2(S^2))^2\right)\\ &> 8\pi^2(m+1)-4\pi^2(m+1)^2\\ &= 4\pi^2(1-m^2).\end{aligned}$$

Taking the square root of both sides we then obtain

$$\mathrm{Per}_{\mathbb{S}^2}(S^1)+\mathrm{Per}_{\mathbb{S}^2}(S^2) > 2\pi\sqrt{1-m^2}.$$

But again, as in Step 1, this would result in $\liminf_{\gamma\to 0}\mathrm{Per}_{\mathbb{S}^2}(\Omega_\gamma) > 2\pi\sqrt{1-m^2}$, a contradiction to the fact that $E_\gamma(u_\gamma)\to\pi\sqrt{1-m^2}$. Therefore Ω_γ cannot have two or more components; hence, it consists of precisely one component.

Step 3. Let us recall that the the boundary of the spherical cap, ∂S, is parametrized for $0\leqslant\theta\leqslant 2\pi$ as $(1,\theta,\phi_0)$, where $\phi_0=\arccos(-m)$. We now claim that the geodesic curvature H_γ of $\partial\Omega_\gamma$ converges to $\cot\phi_0$ uniformly as $\gamma\to 0$.

Applying the Gauss–Bonnet theorem to Ω_γ (cf. [8]), we get that

$$2\pi(m+1)+\int_{\partial\Omega_\gamma} H_\gamma\,d\mathcal{H}^1_x = 2\pi.$$

Then $\int_{\partial\Omega_\gamma} H_\gamma\,d\mathcal{H}^1_x=-2\pi m$. Now integrating both sides of (3.6) over $\partial\Omega_\gamma$ yields

$$-2\pi m+4\gamma\int_{\partial\Omega_\gamma} v_\gamma\,d\mathcal{H}^1_x=\mathrm{Per}_{\mathbb{S}^2}(\Omega_\gamma)\lambda_\gamma.$$

Since $\mathrm{Per}_{\mathbb{S}^2}(\Omega_\gamma)$ is bounded above and below independent of γ and $\mathrm{Per}_{\mathbb{S}^2}(\Omega_\gamma)\to 2\pi\sqrt{1-m^2}=2\pi\sin\phi_0$ as $\gamma\to 0$, invoking the uniform L^∞-bound on v_γ from Proposition 4.3 we get that $\lambda_\gamma\to\cot\phi_0$ as $\gamma\to 0$. Returning to (3.6), this yields

$$\left|H_\gamma-\frac{2\pi\cos\phi_0}{\mathrm{Per}_{\mathbb{S}^2}(\Omega_\gamma)}\right|=\mathcal{O}(\gamma);$$

hence the uniform convergence of H_γ to $\cot\phi_0$.

Step 4. We next claim that the boundary of Ω_γ is globally the graph of a function, that is, $\partial\Omega_\gamma$ can be expressed as $(1,\theta,\cot\phi_0+g_\gamma(\theta))$ in spherical coordinates for some function $g_\gamma:[0,2\pi]\to\mathbb{R}$.

Let $\Gamma_\gamma(s)=(\theta_\gamma(s),\phi_\gamma(s))$ be a parametrization of $\partial\Omega_\gamma$ by arc-length s. Let $\mathbb{S}^2$ be oriented with the unit normal pointing outward and let σ_γ denote the oriented angle $\angle(\partial_\phi,\frac{d\Gamma_\gamma}{ds})$, where ∂_ϕ is the unit tangent to $\mathbb{S}^2$ given by $(\cos\theta\cos\phi,\sin\theta\cos\phi,-\sin\phi)$ in spherical coordinates and $\frac{d\Gamma_\gamma}{ds}$ denotes the tangent vector to the curve Γ_γ. Then, by [29, Proposition 1.1], we have

$$\begin{aligned}\frac{d\phi_\gamma}{ds}&=\cos\sigma_\gamma,\\ \frac{d\sigma_\gamma}{ds}&=H_\gamma-\cot(\phi_\gamma)\sin(\sigma_\gamma).\end{aligned}\tag{4.9}$$

We will prove the claim of this step by considering two separate cases.

First, let us consider the case where $m\in(-1,1)\setminus\{0\}$. In this case, from Step 3, we have that $H_\gamma\to\cot\phi_0\neq 0$. Suppose, for a contradiction, that $\sigma_\gamma(s_\gamma)=0$ for some s_γ. Also assume, without loss of generality, that $\cot\phi_0 > 0$. The case $\cot\phi_0 < 0$ can be proved in the same way.

Note that $\sin\sigma_\gamma(s_\gamma) = 0$ as we have $\sigma_\gamma(s_\gamma) = 0$. Hence, the second equation of (4.9) at s_γ reads

$$\frac{d\sigma_\gamma}{ds}(s_\gamma) = H_\gamma(s_\gamma).$$

Since $H_\gamma(\gamma)$ is uniformly close to $\cot\phi_0 > 0$ by Step 3, we necessarily have that $\frac{d\sigma_\gamma}{ds}(s_\gamma) > 0$. Now, because of the mass constraint $\frac{1}{4\pi}\int_{\partial S} u_\gamma \, d\mathcal{H}^1_x = m$, we can conclude that $\sigma_\gamma(s_{1,\gamma}) = \sigma_\gamma(s_{2,\gamma}) = \frac{\pi}{2}$ where $s_{1,\gamma} = \arg\min\phi_\gamma(s)$ and $s_{2,\gamma} = \arg\max\phi_\gamma(s)$, that is, the parameters corresponding to the most "northern" and most "southern" points on Γ_γ, respectively. But, since $\frac{d\sigma_\gamma}{ds}(s_\gamma) > 0$, we then get that there has to be another value t_γ such that $\sigma_\gamma(t_\gamma) = 0$ and $\frac{d\sigma_\gamma}{ds}(t_\gamma) \leqslant 0$. Referring to (4.9), this yields a contradiction to the fact that H_γ is uniformly close to a positive number for γ small enough.

Therefore $\sigma_\gamma \neq 0$. Similarly, one reaches a contradiction by assuming $\sigma_\gamma = \pi$ at a point. Thus, the tangent vector of Γ_γ is never parallel to the vector ∂_ϕ and Γ_γ can be expressed globally as the graph of a function depending on θ.

Next, we will prove the claim for the case $m = 0$. By Proposition 4.2 we get that $\chi_{\Omega_\gamma} \to \chi_S$ in L^1; hence $\phi_\gamma \to \frac{\pi}{2}$ in L^1 as $\gamma \to 0$. As above, by looking at the most "northern" point on Γ_γ, we can conclude that $\sigma_\gamma(s_\gamma) = \frac{\pi}{2}$ for some s_γ.

Since $\left|\frac{d\phi_\gamma}{ds}\right| \leqslant 1$ by (4.9), and $0 \leqslant \phi_\gamma \leqslant \pi$, defining $\varphi_\gamma(s) = \phi_\gamma(l_\gamma s)$, where l_γ denotes the arc-length of Ω_γ, and noting that l_γ is uniformly away from zero, we see that $\{\varphi\}_{\gamma>0}$ constitutes a family of uniformly bounded functions defined on $[0,1]$ whose derivatives are also uniformly bounded by a constant independent of γ.

Thus, by the Arzela–Ascoli theorem, ϕ_γ has a subsequence, converging uniformly. Since $\phi_\gamma \to \frac{\pi}{2}$ in L^1, that subsequence, still denoted by ϕ_γ, converges to $\frac{\pi}{2}$ uniformly.

By (4.9), we have that

$$\left|\frac{d\sigma_\gamma}{ds}\right| \leqslant |H_\gamma| + |\cot\phi_\gamma|.$$

Therefore $\frac{d\sigma_\gamma}{ds} \to 0$ uniformly; and as $\sigma_\gamma(s_\gamma) = \frac{\pi}{2}$, we get that $\sigma_\gamma \neq 0$ for γ small enough. Hence, as above, Γ_γ can be expressed globally as the graph of a function of θ.

We can thus define the boundary of Ω_γ as the curve parametrized by $(1, \theta, \cot\phi_0 + g_\gamma(\theta))$ for some function g_γ defined on $[0, 2\pi]$.

Step 5. Now we will show that for small $\gamma > 0$, the minimizers $\{u_\gamma\}$ of E_γ are equal to u_S, i.e., they represent the single spherical cap. For simplicity of presentation only, we will take $m = 0$ in the proof. Since we do not have any restrictions on the parameter regime of m, the proof for an arbitrary $m \in (-1,1)\setminus\{0\}$ follows after minor modifications.

In light of Step 4, note that the set $\Omega_\gamma = \{x \in \mathbb{S}^2 : u_\gamma(x) = 1\}$ takes the form

$$S_1 := \left\{(\cos\theta\sin\phi, \sin\theta\sin\phi, \cos\phi) \in \mathbb{S}^2 : 0 \leqslant \phi \leqslant \frac{\pi}{2} + g_\gamma(\theta), \quad 0 \leqslant \theta \leqslant 2\pi\right\}.$$

We will proceed to show that the global minimality of u_γ is violated for small γ unless $u_\gamma \equiv u_S$. To this end, for each $t \in [0,1]$, define

$$S_t := \left\{(\cos\theta\sin\phi, \sin\theta\sin\phi, \cos\phi) \in \mathbb{S}^2 : 0 \leqslant \phi \leqslant \frac{\pi}{2} + t g_\gamma(\theta), \quad 0 \leqslant \theta \leqslant 2\pi\right\}.$$

Note that, the mass constraint $\int_{\mathbb{S}^2} u_\gamma \, d\mathcal{H}^2_x = 0$ entails $\int_0^{2\pi} g_\gamma(\theta)\, d\theta = 0$; hence, $\mathcal{H}^2(S_t) = 2\pi$ for all $t \in [0,1]$. This, along with the fact that $\chi_{S_t} \to \chi_S$ in L^1 as $t \to 0$ then implies that the family of functions

$$U(x,t) := \begin{cases} 1 & \text{if } x \in S_t, \\ -1 & \text{if } x \in S_t^c, \end{cases}$$

consists of admissible competitors in the minimization of E_γ. Furthermore, let $V(x,t)$ be the solution of $-\Delta V(\cdot,t) = U(\cdot,t)$ subject to $\int_{\mathbb{S}^2} V(x,t)\,d\mathcal{H}^2_x = 0$. Note that $U(x,0) = u_S(x)$, $U(x,1) = u_\gamma(x)$, $V(x,0) = v_S(x)$ and $V(x,1) = v_\gamma(x)$.

Let us define

$$e_\gamma(t) := E_\gamma(U)(t) = \frac{1}{2}\int_{\mathbb{S}^2} |\nabla U(x,t)| + \gamma \int_{\mathbb{S}^2} |\nabla V(x,t)|^2\,d\mathcal{H}^2_x.$$

With this definition, we have that $e_\gamma(0) = E_\gamma(u_S)$ and $e_\gamma(1) = E_\gamma(u_\gamma)$. Taylor's Theorem then implies that

$$e_\gamma(1) = e_\gamma(0) + e'_\gamma(0) + \frac{1}{2}e''_\gamma(\tau) = e_\gamma(0) + \frac{1}{2}e''_\gamma(\tau) \tag{4.10}$$

for some $\tau \in (0,1)$ as u_S is a critical point of E_γ, making $e'_\gamma(0) = 0$.

Now we are going to calculate $e''_\gamma(\tau)$ explicitly.

Since

$$\frac{1}{2}\int_{\mathbb{S}^2} |\nabla U(x,t)| = \int_0^{2\pi} \left(\sin^2\left(\frac{\pi}{2} + t g_\gamma(\theta)\right) + (t g'_\gamma(\theta))^2\right)^{1/2}\,d\theta,$$

a straight-forward calculation of the second derivative evaluated at $t = \tau$ yields

$$\begin{aligned}\frac{d^2}{dt^2}\frac{1}{2}\int_{\mathbb{S}^2} &|\nabla U(x,\tau)| \\ &= \int_0^{2\pi} \frac{\cos^2\left(\frac{\pi}{2} + \tau g_\gamma(\theta)\right) g_\gamma^2(\theta) - \sin^2\left(\frac{\pi}{2} + \tau g_\gamma(\theta)\right) g_\gamma^2(\theta) + (g'_\gamma(\theta))^2}{\left(\sin^2\left(\frac{\pi}{2} + \tau g_\gamma(\theta)\right) + (\tau g'_\gamma(\theta))^2\right)^{1/2}} \\ &\quad - \frac{\left[\sin\left(\frac{\pi}{2} + \tau g_\gamma(\theta)\right)\cos\left(\frac{\pi}{2} + \tau g_\gamma(\theta)\right) g_\gamma(\theta) + \tau (g'_\gamma(\theta))^2\right]^2}{\left(\sin^2\left(\frac{\pi}{2} + \tau g_\gamma(\theta)\right) + (\tau g'_\gamma(\theta))^2\right)^{3/2}}\,d\theta.\end{aligned}$$

Clearly, at $t = 0$ we have

$$\frac{d^2}{dt^2}\frac{1}{2}\int_{\mathbb{S}^2} |\nabla U(x,0)| = \int_0^{2\pi} (g'_\gamma(\theta))^2 - g_\gamma^2(\theta)\,d\theta.$$

Taking the derivative one more time we get that for any $t \in [0,1]$

$$\frac{d^3}{dt^3}\frac{1}{2}\int_{\mathbb{S}^2} |\nabla U(x,t)| \geqslant 0,$$

since $\cos\left(\frac{\pi}{2} + \tau g_\gamma(\theta)\right) g_\gamma(\theta) \leqslant 0$ for any $\theta \in [0,2\pi]$. Then, in particular at $t = \tau$, we obtain the inequality

$$\frac{d^2}{dt^2}\frac{1}{2}\int_{\mathbb{S}^2} |\nabla U(x,\tau)| \geqslant \int_0^{2\pi} (g'_\gamma(\theta))^2 - g_\gamma^2(\theta)\,d\theta. \tag{4.11}$$

Now using the definition of $U(x,t)$, integrating by parts once and switching to spherical coordinates, we can rewrite the nonlocal part of the energy as

$$\begin{aligned}\int_{\mathbb{S}^2} |\nabla V(x,t)|^2\, d\mathcal{H}^2_x &= \int_{\mathbb{S}^2} U(x,t)V(x,t)\, d\mathcal{H}^2_x \\ &= \int_{S_t} V(x,t)\, d\mathcal{H}^2_x - \int_{S_t^c} V(x,t)\, d\mathcal{H}^2_x \\ &= \int_0^{2\pi}\int_0^{\frac{\pi}{2}+tg_\gamma(\theta)} V(\theta,\phi,t)\sin\phi\, d\phi d\theta \\ &\qquad - \int_0^{2\pi}\int_{\frac{\pi}{2}+tg_\gamma(\theta)}^{\pi} V(\theta,\phi,t)\sin\phi\, d\phi d\theta.\end{aligned}$$

Taking two derivatives with respect to t and evaluating at $t=\tau$ we get

$$\begin{aligned}&\frac{d^2}{dt^2}\int_{\mathbb{S}^2} |\nabla V(x,\tau)|^2\, d\mathcal{H}^2_x \\ &\quad = 2\int_0^{2\pi} V_\phi\left(\theta, \frac{\pi}{2}+\tau g_\gamma(\theta), \tau\right)\sin\left(\frac{\pi}{2}+\tau g_\gamma(\theta)\right) g_\gamma^2(\theta) \\ &\qquad + V_t\left(\theta, \frac{\pi}{2}+\tau g_\gamma(\theta), \tau\right)\sin\left(\frac{\pi}{2}+\tau g_\gamma(\theta)\right) g_\gamma(\theta) \\ &\qquad + V\left(\theta, \frac{\pi}{2}+\tau g_\gamma(\theta), \tau\right)\cos\left(\frac{\pi}{2}+\tau g_\gamma(\theta)\right) g_\gamma^2(\theta)\, d\theta.\end{aligned} \tag{4.12}$$

Note that, by [13, Chapter 1], we have

$$\int_{\partial S_\tau}\int_{\partial S_\tau} G(x,y)\zeta(x)\zeta(y) d\mathcal{H}^1_x d\mathcal{H}^1_y = \int_{\mathbb{S}^2} |\nabla\omega|^2 \geqslant 0,$$

where ω is an H^1 weak solution to the equation

$$-\Delta\omega = \mu \quad \text{on } \mathbb{S}^2,$$

and μ is the measure given by $\zeta\mathcal{H}^1\lfloor\partial S_\tau$. Thus adapting the calculations in Step 6 of [30, Theorem 3.3], we get that

$$\begin{aligned}&2\int_0^{2\pi} V_t\left(\theta, \frac{\pi}{2}+\tau g_\gamma(\theta), \tau\right)\sin\left(\frac{\pi}{2}+\tau g_\gamma(\theta)\right) g_\gamma(\theta)\, d\theta \\ &\qquad = 4\int_0^{2\pi}\int_0^{2\pi}\left(-\frac{1}{4\pi}\log(2-2\cos(\theta-\alpha))\right) g_\gamma(\theta)g_\gamma(\alpha)\, d\theta d\alpha \geqslant 0.\end{aligned}$$

Returning to (4.12) then yields

$$\begin{aligned}&\frac{d^2}{dt^2}\int_{\mathbb{S}^2} |\nabla V(x,\tau)|^2\, d\mathcal{H}^2_x \\ &\quad \geqslant 2\int_0^{2\pi}\left[V_\phi\left(\theta, \frac{\pi}{2}+\tau g_\gamma(\theta), \tau\right)\sin\left(\frac{\pi}{2}+\tau g_\gamma(\theta)\right)\right. \\ &\qquad \left. + V\left(\theta, \frac{\pi}{2}+\tau g_\gamma(\theta), \tau\right)\cos\left(\frac{\pi}{2}+\tau g_\gamma(\theta)\right)\right] g_\gamma^2(\theta)\, d\theta \\ &\quad \geqslant -C_0,\end{aligned} \tag{4.13}$$

where C_0 is a positive constant depending on $\|V_\phi\|_{L^\infty}$, $\|V\|_{L^\infty}$ and $\|g_\gamma\|_{L^2}$ and hence is independent of γ or τ.

As in the proof of Proposition 4.1 we will write $g_\gamma(\theta) = c_1 \sin\theta + c_2\cos\theta + \tilde{g}_\gamma(\theta)$ for some function $\tilde{g}_\gamma$. We will assume, for a contradiction, that $\tilde{g}_\gamma \not\equiv 0$, that is, the perturbations S_t of the single cap S do not correspond to rotations. Then (4.11) gives that

$$\frac{d^2}{dt^2}\frac{1}{2}\int_{\mathbb{S}^2} |\nabla U(x,\tau)| \geqslant \int_0^{2\pi} (\tilde{g}_\gamma'(\theta))^2 - \tilde{g}_\gamma^2(\theta)\, d\theta \geqslant C_1,$$

where C_1 is a constant independent of γ and is positive by the Poincaré inequality. Now combining this with (4.13) and (4.10) yields

$$\begin{aligned} e_\gamma(1) &= e_\gamma(0) + \frac{1}{2}e_\gamma''(\tau) \\ &\geqslant e_\gamma(0) + (C_1 - \gamma C_0). \end{aligned}$$

For $\gamma < \frac{C_1}{C_0}$, we then get a contradiction to the minimality of $E_\gamma(u_\gamma) = e_\gamma(1)$ since $\|\tilde{g}_\gamma\|_{L^2} > 0$; hence we conclude that necessarily, $\tilde{g}_\gamma \equiv 0$, that is, $u_\gamma \equiv u_S$ up to rotation. □

5. Double Cap and Its Instability

Another interesting critical point of E_γ is the double cap, that is, the set whose boundary consists of two parallel circles of same radius. The double cap appears in two configurations. In this section we will investigate the relation between those configurations and the instability of the double cap. As in the previous section, because of the rotational symmetries of the sphere, we can assume that the boundary components of the double cap are parallel to the equator, i.e., using spherical coordinates we can assume that the boundary components of the double spherical cap are the circles parametrized by $(1,\theta,\phi_0)$ and $(1,\theta,\pi-\phi_0)$ for $\theta\in[0,2\pi]$ and for some fixed $\phi_0 \in (0,\pi/2)$.

Let us fix some $m \in (-1,1)$ and define the function $u_{D,I}$ describing the double cap of first configuration as follows:

$$u_{D,I}(\phi) = \begin{cases} 1 & \text{if } \phi \in [0,\phi_0], \\ -1 & \text{if } \phi \in (\phi_0, \pi-\phi_0], \\ 1 & \text{if } \phi \in (\pi-\phi_0,\pi]. \end{cases} \tag{5.1}$$

Similarly, the function $u_{D,II}$ describing the double cap of second configuration is defined as

$$u_{D,II}(\phi) = \begin{cases} -1 & \text{if } \phi \in [0,\phi_0], \\ 1 & \text{if } \phi \in (\phi_0, \pi-\phi_0], \\ -1 & \text{if } \phi \in (\pi-\phi_0,\pi]. \end{cases} \tag{5.2}$$

With these definitions, we introduce sets $D_I := \{x \in \mathbb{S}^2 : u_{D,I} = 1\}$ and $D_{II} := \{x\in\mathbb{S}^2 : u_{D,II} = 1\}$. Even though it is the complement of the set D_{II} which is of the form of a double cap on $\mathbb{S}^2$, with an abuse of language, we will call the set D_{II} the double cap of type II for the sake of simplicity. Here we also want to note that, since $\frac{1}{4\pi}\int_{\mathbb{S}^2} u\, d\mathcal{H}^2_x = m$, we have $\phi_0 = \arccos\left(\frac{1-m}{2}\right)$ for type I, and $\phi_0 = \arccos\left(\frac{1+m}{2}\right)$ for type II. Since the functions $v_{D,I}$ and $v_{D,II}$ corresponding to $u_{D,I}$ and $u_{D,II}$, respectively, depend only on ϕ, using the facts that both $v_{D,I}'$ and $v_{D,II}'$ are continuous at ϕ_0 and $\pi-\phi_0$ and stay bounded as ϕ approaches 0, we can solve (1.2) explicitly to obtain:

$$v_{D,I}'(\phi) = \begin{cases} (1-m)\cot\phi - \frac{1-m}{\sin\phi} & \text{if } \phi\in[0,\phi_0], \\ -(1+m)\cot\phi & \text{if } \phi \in (\phi_0,\pi-\phi_0], \\ (1-m)\cot\phi + \frac{1-m}{\sin\phi} & \text{if } \phi\in(\pi-\phi_0,\pi], \end{cases} \tag{5.3}$$

and

$$v'_{D,II}(\phi) = \begin{cases} -(1+m)\cot\phi + \frac{1+m}{\sin\phi} & \text{if } \phi \in [0,\phi_0], \\ (1-m)\cot\phi & \text{if } \phi \in (\phi_0, \pi-\phi_0], \\ -(1+m)\cot\phi - \frac{1+m}{\sin\phi} & \text{if } \phi \in (\pi-\phi_0, \pi]. \end{cases} \tag{5.4}$$

While we will show instability of $u_{D,I}$ and $u_{D,II}$ for certain γ-regimes, it is still interesting to compare the total energies of two configurations of the double cap, namely, $E_\gamma(u_{D,I})$ and $E_\gamma(u_{D,II})$. To this end, adapting the notation in (3.1) and (3.2), let us write

$$E_\gamma(u_{D,I}) = P_I(m) + \gamma\, N_I(m),$$
$$E_\gamma(u_{D,II}) = P_{II}(m) + \gamma\, N_{II}(m).$$

Using (5.3) and (5.4), an easy calculation then yields

$$\begin{aligned} P_I(m) &= 4\pi\sqrt{1-\left(\frac{1-m}{2}\right)^2}, \\ N_I(m) &= 2\pi\left[8m\arccos\left(\frac{1-m}{2}\right) - 4\pi m \right. \\ &\qquad \left. -\pi(1-m)^2 + (3-m)\left(\frac{1-m^2}{\sqrt{1-\left(\frac{1-m}{2}\right)^2}}\right)\right], \end{aligned} \tag{5.5}$$

and

$$\begin{aligned} P_{II}(m) &= 4\pi\sqrt{1-\left(\frac{1+m}{2}\right)^2}, \\ N_{II}(m) &= 2\pi\left[-8m\arccos\left(\frac{1+m}{2}\right) + 4\pi m \right. \\ &\qquad \left. -\pi(1+m)^2 + (3+m)\left(\frac{1-m^2}{\sqrt{1-\left(\frac{1+m}{2}\right)^2}}\right)\right]. \end{aligned} \tag{5.6}$$

Clearly, for $m = 0$, both configurations have the same total energy. Also, after a simple calculation, we see that $E_\gamma(u_{D,I}) = E_\gamma(u_{D,II})$ along the curve

$$\gamma(m) = \frac{P_{II}(m) - P_I(m)}{N_I(m) - N_{II}(m)}$$

defined in the $(m,\gamma)-$plane.

Note that the curve $\gamma(m)$ and the γ-axis divide the $(m,\gamma)-$plane into four regions: A, B, C and D (cf. Figure 1). By (5.5) and (5.6), we get that, in regions A and C, the second configuration has less energy than the first configuration, that is, $E_\gamma(u_{D,II}) < E_\gamma(u_{D,I})$, whereas in regions B and D we have $E_\gamma(u_{D,I}) < E_\gamma(u_{D,II})$.

Now, we will establish the instability of $u_{D,I}$ and $u_{D,II}$ on certain regimes. By the fact that the boundary components of D_I and D_{II} are circles with radius $\sin\phi_0$, referring to

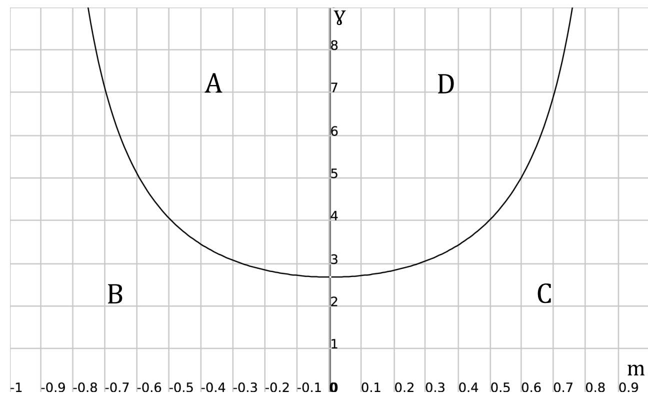

FIGURE 1. $\gamma(m)$ in the $(m,\gamma)-$plane

(5.3) and (5.4) we obtain that

$$\|B_{\partial D_I}\|^2 = \frac{\left(\frac{1-m}{2}\right)^2}{1-\left(\frac{1-m}{2}\right)^2} \qquad \|B_{\partial D_{II}}\|^2 = \frac{\left(\frac{1+m}{2}\right)^2}{1-\left(\frac{1+m}{2}\right)^2},$$

$$(\nabla v_{D,I}\cdot\nu)|_{\partial D_I} = -\frac{1-m^2}{2\sqrt{1-\left(\frac{1-m}{2}\right)^2}} \qquad (\nabla v_{D,II}\cdot\nu)|_{\partial D_{II}} = -\frac{1-m^2}{2\sqrt{1-\left(\frac{1+m}{2}\right)^2}}.$$

Let f be a smooth function defined on ∂D_I such that $\int_{\partial D_I} f\,d\mathcal{H}^1_x = 0$. Then, by the equations for $\|B_{\partial D_I}\|^2$ and $(\nabla v_{D,I}\cdot\nu)|_{\partial D_I}$ given above, (3.7) becomes

$$\begin{aligned} J_{I,\gamma}(f) = &\int_{\partial D_I} |\nabla_{\partial D_I} f|^2 - \left(\frac{1}{1-\left(\frac{1-m}{2}\right)^2}\right) f^2\,d\mathcal{H}^1_x \\ &- 8\gamma \int_{\partial D_I}\int_{\partial D_I} \left(-\frac{1}{2\pi}\log(|x-y|)\right) f(x)f(y)\,d\mathcal{H}^1_x d\mathcal{H}^1_y \\ &- 4\gamma\,\frac{1-m^2}{2\sqrt{1-\left(\frac{1-m}{2}\right)^2}}\int_{\partial D_I} f^2\,d\mathcal{H}^1_x. \end{aligned} \tag{5.7}$$

Using the equations for $\|B_{\partial D_{II}}\|^2$ and $(\nabla v_{D,II}\cdot\nu)|_{\partial D_{II}}$ this time, it is easy to note that $J_{II,\gamma}$ looks exactly the same as (5.7) except m is replaced by $-m$. Hence, the instability result of the next proposition for the double cap holds true for both the double cap of type I and II with a minor modification of switching m by $-m$.

Now we can state the instability result concerning the double cap.

Proposition 5.1. *For any $m\in(-1,1)$, there exists two values γ_0 and γ_1, depending only on m, such that the functions $u_{D,I}$ and $u_{D,II}$, defined in* (5.1) *and* (5.2)*, are unstable critical points of E_γ for all positive $\gamma<\gamma_0$ and for all $\gamma>\gamma_1$.*

Furthermore, when $m=0$, $u_{D,I}$ and $u_{D,II}$ are unstable for all $\gamma>0$.

Proof. As noted above, we will prove this proposition only for $u_{D,I}$. The proof for $u_{D,II}$ follows by switching m by $-m$. For the simplicity of presentation, let us define

$$a := \sqrt{1-\left(\frac{1-m}{2}\right)^2}.$$

Let us also denote the boundary components of D_I by Γ_1 and Γ_2, that is, let $\partial D_I = \Gamma_1 \cup \Gamma_2$.

We will define $f : \partial D_I \to \mathbb{R}$ to be

$$f(x) = \begin{cases} 1 & \text{if } x \in \Gamma_1, \\ -1 & \text{if } x \in \Gamma_2. \end{cases}$$

Then plugging f into (5.7) yields

$$\begin{aligned} J_{I,\gamma}(f) = -\frac{4\pi}{a} + 8\gamma \Bigg[& -\frac{1}{\pi} \int_{\Gamma_1}\int_{\Gamma_1} \log(|x-y|)\, d\mathcal{H}^1_x d\mathcal{H}^1_y \\ & \frac{1}{\pi} \int_{\Gamma_1}\int_{\Gamma_2} \log(|x-y|)\, d\mathcal{H}^1_x d\mathcal{H}^1_y \Bigg] \\ & - 16\pi(1-m^2)\gamma. \end{aligned} \tag{5.8}$$

Now, using the identity (4.5) we get that

$$-\frac{1}{\pi} \int_{\Gamma_1}\int_{\Gamma_1} \log(|x-y|)\, d\mathcal{H}^1_x d\mathcal{H}^1_y = -4\pi a^2 \log a.$$

Also, noting that $|x-y| < 2$ for $x \in \Gamma_1$ and $y \in \Gamma_2$, we have

$$\frac{1}{\pi} \int_{\Gamma_1}\int_{\Gamma_2} \log(|x-y|)\, d\mathcal{H}^1_x d\mathcal{H}^1_y \leqslant 4\pi a^2 \log 2.$$

Hence, going back to (5.8) we obtain that

$$J_{I,\gamma}(f) \leqslant -\frac{4\pi}{a} + 32\pi a^2(\log 2 - \log a)\gamma - 16\pi(1-m^2)\gamma.$$

Taking $\gamma_0 := (4a[2a^2(\log 2 - \log a) - (1-m^2)])^{-1}$ then implies that $J_{I,\gamma} < 0$ for all $\gamma < \gamma_0$, that is, $u_{D,I}$ is unstable for all $\gamma < \gamma_0$.

Next, let $k > 1$ be a positive integer and define $f_k : \partial D_I \to \mathbb{R}$ by

$$f_k(x) = \begin{cases} \sin(k\theta) & \text{if } x = (\cos\theta\sin\phi, \sin\theta\sin\phi, \cos\phi) \in \Gamma_1, \\ 0 & \text{if } x = (\cos\theta\sin\phi, \sin\theta\sin\phi, \cos\phi) \in \Gamma_2. \end{cases}$$

Then plugging f_k into (5.7) and using (4.6) gives

$$J_{I,\gamma}(f_k) = ak^2\pi - \frac{\pi}{a} + \left(\frac{4\pi a^2}{k} - 4(1-m^2)\pi \right)\gamma.$$

Taking $\gamma_1 := \inf_k \frac{ak^3 - a^{-1}k}{4k(1-m^2) - 4a^2}$ then yields $J_{I,\gamma} < 0$ for all $\gamma > \gamma_1$; hence, $u_{D,I}$ is unstable for all $\gamma > \gamma_1$.

Now, for $m = 0$, we have that $a = \sqrt{3}/2$ and the above calculations imply that

$$\gamma_0 = \left(2\sqrt{3} \left[\frac{3}{2} \left(\log 2 - \log \frac{\sqrt{3}}{2} \right) - 1 \right] \right)^{-1} \quad \text{and} \quad \gamma_1 = \frac{4\sqrt{3} - \frac{4}{\sqrt{3}}}{5}.$$

Hence, on one hand, $J_{I,\gamma} < 0$ for

$$\gamma < \gamma_0 \approx 1.13;$$

on the other hand, $J_{I,\gamma} < 0$ for

$$\gamma > \gamma_1 \approx 0.92.$$

Thus, when $m = 0$, $u_{D,I}$ is an unstable critical point of E_γ for all values of $\gamma > 0$. □

6. Closing Remarks

Being the first attempt to analyze (NLIP) on a curved surface, of course, our study leaves many unanswered questions that, we hope, will attract interest in the near future.

(i) From Proposition 4.1 we know that the single spherical cap is a stable critical point of (1.1) for any $m \in (-1,1)$ and for any $\gamma < \gamma_c$. We also know that the single cap is the global minimizer of (NLIP) for sufficiently small γ-values. One can ask then whether the single cap remains as the global minimizer as γ approaches γ_c. In other words, as we increase γ from 0 to γ_c, when does the single cap lose its global minimality? Unfortunately, due to our use of an isoperimetric-type inequality in the proof of Theorem 4.4, at this moment, we are not able to quantify the γ-value which corresponds to the threshold of global minimality.

(ii) One might also wonder about the nature of the bifurcation that occurs from the single cap at $\gamma = \gamma_c$. Referring to the work [24] on wriggled lamellar patterns on the flat-torus, it seems reasonable to guess that the single cap might also bifurcate into a wriggled single cap as γ exceeds γ_c.

(iii) Proposition 5.1 gives a quite "negative" result regarding the double cap. Indeed, we show that for $m = 0$, the double cap is unstable for any $\gamma > 0$. At this point, it is unknown to us whether there is a regime of γ-values between γ_0 and γ_1 for which the double cap critical points are stable whenever $m \neq 0$. In [23] the authors show that striped patterns on a flat-torus possess stability as long as they have a certain number of interfaces. Also, in [2] the hedgehog defect structures (patterns where the diffuse interfaces are parallel annuli) were observed in self-consistent field theory simulations of lamellar block copolymers on a sphere. This suggests that (NLIP) might deliver stable critical points with parallel circles as the interfaces once the number of interfaces exceeds a fixed value.

Finally, we would like to remark that one needs to be careful when drawing an analogy between the lamellar patterns on the flat-torus and the hedgehog patterns on the sphere. One obvious difference is in the method of proving instability of the double cap when γ is small enough. Here, we have used the second variation evaluated at a function which takes on the values 1 and -1 on disconnected boundary components to obtain instability of the double cap for small γ-values. This idea does not work for double stripes on the flat-torus, as such functions correspond to translations and yield zero second variation. To our knowledge, instability of multiple striped patterns for small γ-values on the flat-torus is still an open problem. Another major difference is the dependence on the mass constraint. Our results here, especially the global minimality of the single cap, hold true for all values of m in the interval $(-1,1)$, whereas in the case of the flat-torus, the global minimality of the single stripe is obtained for a smaller interval of m-values (cf. [30, Theorem 3.3]). Indeed, the global minimizer of the local isoperimetric problem on the two-torus is either a disk or a stripe depending on the regime of m-values.

Acknowledgments

The author is grateful to Peter Sternberg for his guidance and support. He also would like to thank Nets Katz for pointing out the identity (4.5) which simplifies the arguments involving the Green's function, and the reviewer for his/her helpful suggestions improving the stability result for the single cap solution.

References

[1] J.L. Barbosa, M. do Carmo and J. Eschenburg, *Stability of hypersurfaces of constant mean curvature in Riemannian manifolds*, Math. Z., **197** (1988), 123–138.

[2] T.L. Chantawansri, A.W. Bosse, A. Hexemer, H.D. Ceniceros, C.J. García-Cervera, E.J. Kramer and G.H. Fredrickson, *Self-consistent field theory simulations of block copolymers assembly on a sphere*, Phys. Rev. E, **75** (2007), 031802.

[3] R. Choksi, *Nonlocal Cahn–Hilliard and isoperimetric problems: Periodic phase separation induced by competing long- and short-term interactions*, in "CRM Proceedings and Lecture Notes: Singularities in PDE and the Calculus of Variations" (eds. S. Alama, L. Bronsard and P.J. Sternberg), American Mathematical Society, (2008), 33–45.

[4] R. Choksi and M.A. Peletier, *Small volume fraction limit of the diblock copolymer problem I: Sharp interface functional*, SIAM J. Math. Anal., **42** (2010), 1334–1370.

[5] R. Choksi and M.A. Peletier, *Small volume fraction limit of the diblock copolymer problem II: Diffuse interface functional*, SIAM J. Math. Anal., **43** (2011), 739–763.

[6] R. Choksi, M.A. Peletier and J.F. Williams, *On the phase diagram for microphase separation of diblock copolymers: An approach via a nonlocal Cahn–Hilliard functional*, SIAM J. Appl. Math., **69** (2009), 1712–1738.

[7] R. Choksi and P. Sternberg, *On the first and second variations of a nonlocal isoperimetric problem*, J. Reine Angew. Math., **611** (2005), 75–108.

[8] M.P. do Carmo, "Differential Geometry of Curves and Surfaces," Prentice Hall, New Jersey, 1976.

[9] E. Giusti, *The equilibrium configuration of liquid drops*, J. Reine Angew. Math., **321** (1981), 53–63.

[10] E. Giusti, "Minimal Surfaces and Functions of Bounded Variation," Birkhäuser, Boston, 1984.

[11] K. Grosse-Brauckmann, *Stable constant mean curvature surfaces minimize area*, Pacific J. Math., **175** (1996), 527–534.

[12] J. Jost, "Riemannian Geometry and Geometric Analysis," Springer-Verlag, Berlin, 2005.

[13] N.S. Landkof, "Boundations of Modern Potential Theory," Springer-Verlag, Berlin, 1972.

[14] U. Massari, *Esistenza e regolarità delle ipersuperfici di curvatura media assegnata in* $\mathbb{R}^n$, Arch. Rational Mech. Anal., **55** (1974), 357–382.

[15] J. Montero, P. Sternberg and W. Ziemer, *Local minimizers with vortices to the Ginzburg-Landau system in 3d*, Comm. Pure Appl. Math., **57** (2004), 99–125.

[16] F. Morgan, *Regularity of isoperimetric hypersurfaces in Riemannian manifolds*, Trans. Amer. Math. Soc., **355** (2003), 5041-5052.

[17] F. Morgan, M. Hutchings and H. Howards, *The isoperimetric problem on surfaces of revolution of decreasing Gauss curvature*, Trans. Amer. Math. Soc., **352** (2000), 4889–4909.

[18] F. Morgan and A. Ros, *Stable constant-mean-curvature hypersurfaces are area minimizing in small* L^1 *neighborhoods*, Interfaces Free Bound., **12** (2010), 151–155.

[19] C.B. Muratov, *Droplet phases in non-local Ginzburg-Landau models with Coulomb repulsion in two dimensions*, Comm. Math. Phys., **299** (2010), 45–87.

[20] T. Ohta and K. Kawasaki, *Equilibrium morphology of block copolymer melts*, Macromolecules, **19** (1986), 2621-2632.

[21] M.A. Peletier and M. Veneroni, *Stripe patterns in a model for block copolymers*, Math. Model. Meth. Appl. Sci., **20** (2010), 843–907.

[22] X. Ren and J. Wei, *On the multiplicity of two nonlocal variational problems*, SIAM J. Math. Anal., **31** (2000), 909–924.

[23] X. Ren and J. Wei, *On the spectra of three dimensional lamellar solutions of the diblock copolymer problem*, SIAM J. Math. Anal., **35** (2003), 1–32.

[24] X. Ren and J. Wei, *Wriggled lamellar solutions and their stability in the diblock copolymer problem*, SIAM J. Math. Anal., **37** (2005), 455–489.

[25] X. Ren and J. Wei, *Many droplet pattern in the cylindrical phase of diblock copolymer morphology*, Rev. Math. Phys., **19** (2007), 879–921.

[26] X. Ren and J. Wei, *Spherical solutions to a nonlocal free boundary problem from diblock copolymer morphology*, SIAM J. Math. Anal., **39** (2008), 1497–1535.

[27] X. Ren and J. Wei, *Oval shaped droplet solutions in the saturation process of some pattern formation problems*, SIAM J. Math. Anal., **70** (2009), 1120–1138.

[28] X. Ren and J. Wei, *A toroidal tube solution of a nonlocal geometric problem*, Interfaces Free Bound., **13** (2011), 127–154.

[29] M. Ritore, *Constant geodesic curvature curves and isoperimetric domains in rotationally symmetric surfaces*, Comm. Anal. Geom., **9** (2001), 1093–1138.
[30] P. Sternberg and I. Topaloglu, *On the global minimizers of a nonlocal isoperimetric problem in two dimensions*, Interfaces Free Bound., **13** (2011), 155–169.
[31] P. Sternberg and W. Ziemer, *Local minimizers of a three-phase partition problem with triple junctions*, Proc. Roy. Soc. Edin. Sect. A, **124** (1994), 1059-1073.
[32] I. Tamanini, *Boundaries of Caccioppoli sets with Hölder-continuous normal vector*, J. Reine Angew. Math., **334** (1982), 27–39.
[33] P. Tang, F. Qiu, H. Zhang and Y. Yang, *Phase separation patterns for diblock copolymers on spherical surfaces: A finite volume method*, Phys. Rev. E, **72** (2005), 016710.
[34] E.T. Whittaker and G.N. Watson, "A Course of Modern Analysis," Cambridge University Press, London, 1927.

Department of Mathematics, Indiana University, Bloomington, IN
E-mail address: itopalog@umail.iu.edu