

Joseph D. Masters*, University of Texas at Austin, Mathematics Department,
e-mail: masters@math.utexas.edu

THE PERIMETER-MINIMIZING ENCLOSURE OF TWO AREAS IN S^2

Abstract

It is shown that the “standard double bubble” is the unique least-perimeter way to enclose and separate two given areas in the surface of the round sphere.

1 Introduction

Foisy et al. [F] have proved that, among all shapes that enclose and separate two given areas in the plane, the “standard double bubble” shape of Figure 1.1a uniquely minimizes perimeter. This paper uses some new findings of Hutchings [H] to generalize this result to the sphere. The solution is pictured in Figure 1.1b.

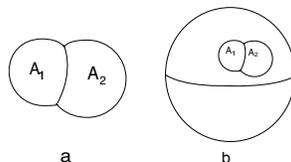


Figure 1.1: (a) The standard double bubble is the shortest enclosure of two prescribed areas in \mathbb{R}^2 . (b) We prove that the standard double bubble is the shortest enclosure of two prescribed areas in S^2 .

We have the following two main results:

Mathematical Reviews subject classification: 52A38

Received by the editors December 8, 1995

*This project was supported by the Williams College Bronfman Science Center and the National Science Foundation, and advised by Professor Frank Morgan. I would like to thank Professor Morgan for his invaluable help and encouragement. Special thanks also to Michael Hutchings, who suggested the outline of the proof. Finally, I would like to thank Professor Stewart Johnson and Professor Gary Lawlor for valuable suggestions.

Theorem 2.2 (Unique standard bubble). *Given areas $A_1, A_2 > 0$, with $A_1 + A_2 < 4\pi$, there exists a unique standard bubble, $D(r_1, r_2)$, in the unit round sphere enclosing and separating regions of area A_1 and A_2 .*

Theorem 2.10 (Standard bubble minimizing). *The least-perimeter way to enclose and separate two prescribed areas in S^2 is the standard double bubble.*

Incidentally, it was the results of Hutchings which made possible the recent announcement by Hass and Schlafly [HS], of a computer proof of the 3-dimensional Double Bubble Conjecture on the least-surface area way to enclose and separate two regions of equal volume in \mathbb{R}^3 (see also [P], [HHS]). There has been progress on the planar triple bubble too [CHK].

Idea of the proof

The generalization of the planar result to the sphere presents new difficulties. In particular, the proof of the planar case uses a powerful result about the monotonicity of the least-perimeter function which does not carry over to the sphere.

The major complication in dealing with minimal bubble clusters in general is the possibility of disconnected regions (see Figure 1.2). If these can be ruled out for the case of the double bubble in the sphere, then the regularity theorem [M3, 2.4] and our relatively easy uniqueness Theorem 2.2 will immediately imply the desired result.

If one of the three areas partitioned by the double bubble is at least twice as big as another, then a theorem in [H] implies connectivity. Otherwise, we follow a suggestion of Hutchings, and consider a “decomposition inequality”, which gives a lower bound for the perimeter of a double bubble with disconnected regions, as a function of the two prescribed areas. We need to show that, if no area is at least twice as big as another, then the perimeter of the standard double bubble is always less than this lower bound.

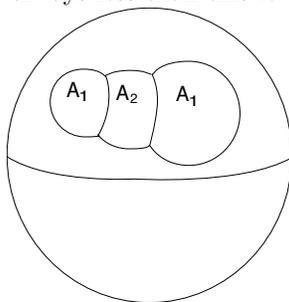


Figure 1.2: A double bubble with disconnected regions.

We provide a numerical proof. The idea is to plot the lower bound and the perimeter of the standard double bubble over a grid of closely spaced points. We show that the perimeter of the double bubble with disconnected regions is greater than the perimeter of the standard double bubble for all the points we plot. We then obtain a crude bound on their gradients, and by making the points close enough together, we show that within any of the small intervals, the functions cannot change enough to make up the difference. Therefore the regions are connected and the double bubble is standard.

Definitions

A *double bubble* in the unit round sphere consists of two disjoint open sets, B_1 and B_2 , known as *regions*. Let B_0 denote the complement of their closures, called the *exterior*. The *perimeter* of a double bubble is the one-dimensional Hausdorff measure of the topological boundaries of the bubble. Let A_i denote the area of B_i . A *standard double bubble* is made up of three arcs of constant curvature meeting at two vertices at 120° angles.

For further background and context see [M2] or the new chapter on soap bubbles in [M1].

2 The double bubble in S^2

In this section we prove that the standard double bubble uniquely minimizes perimeter in S^2 .

The following theorem follows immediately from [M3, 2.4]

Theorem 2.1 (Existence and regularity). *Given $A_1, A_2 > 0$, with $A_1 + A_2 < 4\pi$, there exists a perimeter-minimizing double bubble in the unit round sphere, made up of arcs of constant geodesic curvature meeting in threes at angles of 120° , enclosing and separating regions of area A_1 and A_2 .*

The proof of the following existence and uniqueness theorem is due to Frank Morgan.

Theorem 2.2 (Unique standard bubble). *Given areas $A_1, A_2 > 0$, with $A_1 + A_2 < 4\pi$, there exists a unique standard double bubble, $D(r_1, r_2)$, in the unit round sphere enclosing and separating regions of area A_1 and A_2 .*

PROOF. First note that for any two geodesic radii r_1 and r_2 , we can construct a standard double bubble $D(r_1, r_2)$ as follows (see Figure 2.1):

Take counterclockwise circular arcs C_1 and C_2 of radii r_1, r_2 emanating from a fixed point O at an angle of 120° (oriented as in Figure 2.1) until they

meet again (at 120° by symmetry). Complete D with a third circular arc, C_3 , meeting the other two at 120° angles at both points. Note that the centers of the three circles all lie on the same great circle. It is clear that every standard double bubble can be obtained in this way.

Let A_1 be the area enclosed by C_1 and C_2 ; A_2 the area enclosed by C_2 and C_3 and A_0 the area enclosed by C_3 and C_1 . A_0, A_1, A_2 are smooth functions of r_1, r_2 . As $r_2 \rightarrow 0$ (r_1 fixed), D degenerates to a single circle C_1 with $A_2 = 0$ (see Figure 2.1a). Similarly, as $r_1 \rightarrow \pi$ (r_2 fixed), D degenerates to a single circle C_2 with $A_0 = 0$. As $r_2 \rightarrow \pi$, D degenerates to a point with $A_0 = A_1 = 0$. As $r_1 \rightarrow 0$, D degenerates to a point with $A_1 = A_2 = 0$ (see Figure 2.1b).

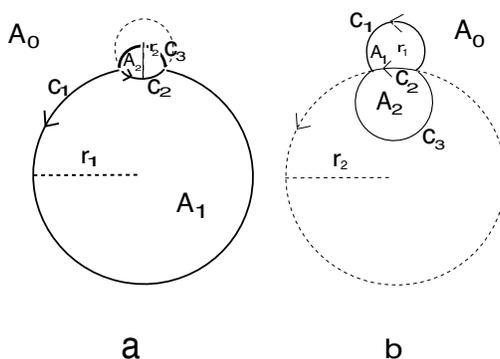


Figure 2.1: The construction of the standard double bubble in the unit round sphere.

Now consider the mapping F from R , the rectangle $0 \leq r_1, r_2 \leq \pi$ to T , the triangle $A_1, A_2 \geq 0, A_1 + A_2 \leq 4\pi$ (see Figure 2.2). The only discontinuity of F is at the upper left corner, $r_1 = 0, r_2 = \pi$. To avoid this discontinuity, we remove a neighborhood of the vertex along a short diagonal d , as shown in Figure 2.2, to form a new region, R' . Let T' be the region bounded by the curve $F(d)$, the segment $A_1 = 0$, and the segment $A_1 + A_2 = 4\pi$. F maps the interior of R_1 smoothly to the interior of T' and the boundary of R' continuously to the boundary of T' . Specifically, F maps the left side to the vertex $(0, 0)$, the bottom injectively onto the base, the right side injectively onto the hypotenuse, the top to the vertex $A_2 = 4\pi$, and d injectively onto $F(d)$. Therefore F has degree 1 and maps the interior of R' onto the interior of T' . As d gets small, $F(d)$ approaches the segment $A_1 = 0$, so F maps the interior of R onto the interior of T . So when A_1 and A_2 are non-zero, there exists a standard double bubble enclosing and separating them. This is sufficient, because when one of the areas is zero, the existence problem is

trivial. To prove uniqueness, it suffices to show that F is injective on the interior. This follows, because for fixed r_1 , A_1 is strictly decreasing in r_2 and A_2 is strictly increasing in r_2 ; similarly for fixed r_2 , A_1 and A_2 are both strictly increasing in r_1 . Hence for any change in (r_1, r_2) , there is a net change in either A_1 or A_2 .

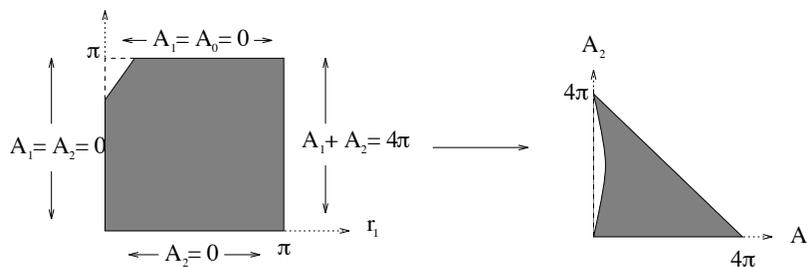


Figure 2.2: F maps the interior of R' bijectively onto the interior of T' . □

Lemma 2.3. *If one of the three regions partitioned by a perimeter-minimizing double bubble in the unit round sphere is connected, then all three are.*

PROOF. Assume that one of these regions is connected. The proof of Lemma 2.4 in [F] carries over to show that connectivity of the exterior implies connectivity of the other two regions. Since the problem of enclosing and separating A_1 and A_2 is equivalent to the problem of enclosing and separating any two of the A_i , the connected region may be viewed as the exterior, implying the connectivity of the other two regions. □

The following two results, generalized to n -dimensional spheres, appear in Theorem 3.7 and Corollary 3.8, respectively, in [H].

Proposition 2.4. *Let $L(A_1, A_2)$ be the function which gives the length of a perimeter-minimizing double bubble enclosing A_1 and A_2 . Then this function is strictly concave on every line in the triangular region $A_0 + A_1 + A_2 = 4\pi$; $A_0, A_1, A_2 \geq 0$ (see Figure 2.3a).*

Lemma 2.5. *If A_0 is at least twice as big as A_1 or A_2 , then the exterior is connected.*

We now have an immediate solution for the following case:

Corollary 2.6. *If any one of the three areas A_0, A_1, A_2 is at least twice as big as another, then all three regions are connected (see Figure 2.3b).*

PROOF. Assume one of the three areas is at least twice as big as another. We are free to view any of the three regions as the exterior, so Lemma 2.5 implies that one of the three regions is connected. Then by Lemma 2.3, all three are. \square

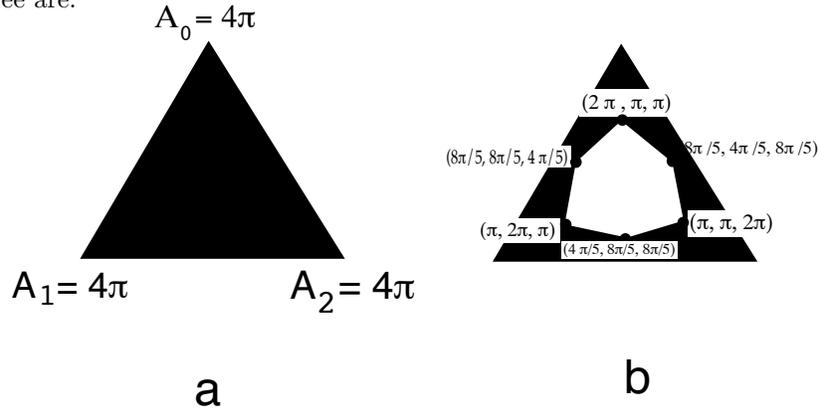


Figure 2.3: (a) The least-perimeter function is concave along any line in the triangular region. (b) We have an immediate solution for the case where one of the three areas is at least twice as big as another.

The following proposition is a trivial generalization of Lemma 4.1 in [H]. Although we will not present a proof, the key idea is contained in Figure 2.4.

Proposition 2.7. *If $B(y, z)$ is a perimeter-minimizing enclosure of regions R_1 , of area y , and R_2 , of area z , and if R_2 is disconnected with a connected component of area xz , $0 < x < 1$, then the following “decomposition inequality” holds:*

$$2L(y, z) \geq L(xz) + L(y, (1-x)z) + L(xz + y, (1-x)z). \tag{1}$$

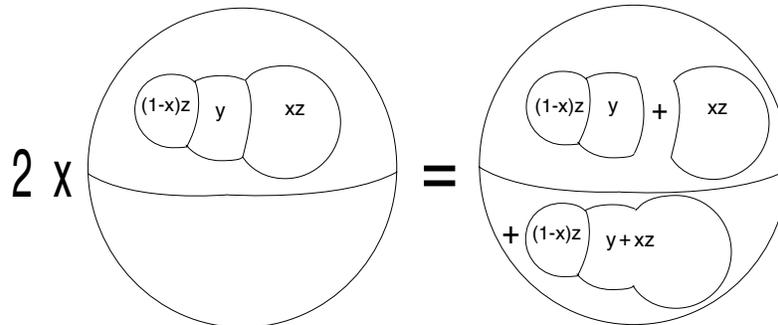


Figure 2.4: The key idea for the proof of Proposition 2.7.

We now derive another inequality:

Proposition 2.8. *Under the same conditions as in Proposition 2.7, the following inequality holds*

$$L(y, z) \geq \frac{L(xz)}{2x} + \frac{L(y)}{2} + \frac{L(z+y)}{2}. \tag{2}$$

PROOF. By concavity we have

$$L(y, (1-x)z) = L((1-x)(y, z) + x(y, 0)) \geq (1-x)L(y, z) + xL(y).$$

Similarly:

$$L(xz+y, (1-x)z) = L(x(z+y, 0) + (1-x)(y, z)) \geq xL(z+y) + (1-x)L(y, z).$$

By substitution into (1), we have

$$2L(y, z) \geq L(xz) + 2(1-x)L(y, z) + xL(y) + xL(z+y);$$

$$L(y, z) \geq \frac{L(xz)}{2x} + \frac{L(y)}{2} + \frac{L(z+y)}{2}.$$

Except for the last one, the following formulas are analogous to those which appear as Proposition 2.1 in [F]. □

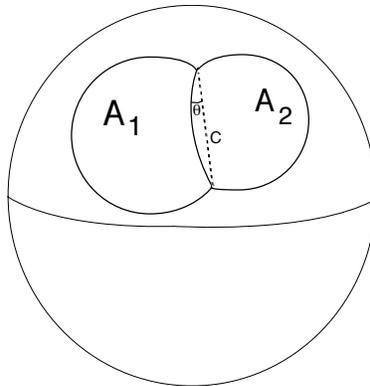


Figure 2.5: We can derive an explicit formula for the perimeter of a standard double variable.

Proposition 2.9. *The following formulas hold on the surface of the unit sphere (see Figure 2.5):*

- Length of a circular arc making an angle θ with a great circle arc of length C :

$$L(\theta, C) = 2 \sin \left[\arctan \left[\frac{\tan \left(\frac{C}{2} \right)}{\sin(\theta)} \right] \right] \arccos \left[\cos \left(\frac{C}{2} \right) \cos(\theta) \right].$$

- Area enclosed by a circular arc and a great circle arc:

$$A(\theta, C) = -2 \cos \left[\arctan \left[\frac{\tan \left(\frac{C}{2} \right)}{\sin(\theta)} \right] \right] \arccos \left[\cos \left(\frac{C}{2} \right) \cos(\theta) \right] + 2\theta.$$

- Length of a circle enclosing area A :

$$L(A) = 2\pi \sin \left[\arccos \left[1 - \frac{A}{2\pi} \right] \right].$$

- Given area, in terms of θ and C :

$$A_1(\theta, C) = A(\theta, C) + A \left(\frac{2\pi}{3} - \theta, C \right)$$

$$A_2(\theta, C) = A \left(\frac{2\pi}{3} + \theta, C \right) - A \left(\frac{2\pi}{3} - \theta, C \right).$$

- Perimeter of the standard double bubble:

$$L_{Stan}(\theta, C) = L(\theta, C) + L \left(\frac{2\pi}{3} - \theta, C \right) + L \left(\frac{2\pi}{3} + \theta, C \right).$$

- Lower bound for the perimeter of a disconnected bubble (under the same condition as Proposition 2.7):

$$L_{Disc}(\theta, C) \geq \frac{L(xA_1(\theta, C))}{2x} + \frac{L(A_2(\theta, C))}{2} + \frac{L(A_1(\theta, C) + A_2(\theta, C))}{2}$$

PROOF. The final inequality follows from Proposition 2.8. The other derivations are all straightforward. \square

We can now prove the main theorem:

Theorem 2.10 (Standard bubble minimizing). *The least-perimeter way to enclose and separate two prescribed areas in S^2 is the standard double bubble.*

PROOF. We are given (A_0, A_1, A_2) . By Lemma 2.5, if one of the areas is twice as big as another, then all three regions are connected. We want to show that for all remaining cases $\Lambda_{\text{Disc}} \geq \Lambda_{\text{Stan}}$.

By Proposition 2.9, if none of the three areas is twice as big as another, then θ and C are within the following range:

$$C < .77\pi$$

$$\theta < \frac{\pi}{3}.$$

By differentiating, we see that the lower bound on Λ_{Disc} obtained in Proposition 2.9 decreases with x . So to minimize this bound, we set $x = .5$. We plot Λ_{Stan} and the bound for Λ_{Disc} on Mathematica, using step sizes of $\Delta\theta = \Delta C = .012$. The program tells us that over this domain, the minimum difference of the two functions, to an accuracy of one decimal place, is $.9$.

It is a straightforward matter to obtain crude bounds:

$$|\Delta\Lambda_{\text{Disc}}(\theta, C)| < 22.48$$

$$|\Delta\Lambda_{\text{Stan}}(\theta, C)| < 13.167.$$

The maximum change along any interval is:

$$\Delta(|\Lambda_{\text{Disc}} - \Lambda_{\text{Stan}}|) < (2^{-5})(.012)(13.167 + 22.48) < .7 < .9.$$

Therefore, when no area is twice as large as another, $\Lambda_{\text{Disc}} > \Lambda_{\text{Stan}}$, so the regions must be connected. Then by Theorem 2.1, the double bubble on the unit sphere must be standard. By scaling, all of these results carry over to a round sphere of any radius. \square

References

- [CHK] C. Cox, L. Harrison, M. Hutchings, S. Kim, J. Light, A. Mauer, M. Tilton, *The shortest enclosure of three connected areas in \mathbb{R}^2* , Real Analysis Exchange, **20** (1994/95), 313–335.
- [F] J. Foisy, M. Alfaro, J. Brock, N. Hodges, J. Zimba, *The standard double bubble in \mathbb{R}^2 uniquely minimizes perimeter*, Pacific J. Math., **159** (1993), 47–59. Featured in 1994 AMS *What's Happening in the Mathematical Sciences*.
- [H] M. Hutchings, *The structure of area-minimizing double bubbles*, J. Geom. Anal. (to appear).

- [HHS] J. Hass, M. Hutchings, R. Schlafly, *The double bubble conjecture*, Elec. Res. Ann. AMS, **1** (1995), 98–102.
- [HS] J. Hass, R. Schlafly, *Double Bubbles Minimize*, preprint (1995).
- [M1] F. Morgan, *Geometric Measure Theory: a Beginner's Guide*, Academic Press, 2nd ed., 1995.
- [M2] F. Morgan, *Mathematicians, including undergraduates, look at soap bubbles*, Amer. Math. Monthly, **101** (1994), 343–351.
- [M3] F. Morgan, *Soap bubbles in \mathbb{R}^2 and in surfaces*, Pac. J. Math., **165** (1994), 347–361.
- [P] I. Peterson, *Toil and trouble over double bubbles*, Science News, **148**, August 12, 1995, 101.