Minimal connected partitions of the sphere

Edward Newkirk

Williams College, edward.newkirk@googlemail.com

Follow this and additional works at: https://scholar.rose-hulman.edu/rhumj

Recommended Citation
Available at: https://scholar.rose-hulman.edu/rhumj/vol11/iss2/1
Minimal connected partitions of the sphere

Edward Newkirk\textsuperscript{a}

Volume 11, No. 2, Fall, 2010

\textsuperscript{a}Williams College, Williamstown, MA
Minimal connected partitions of the sphere

Edward Newkirk

Abstract. We consider the soap bubble problem on the sphere $S^2$, which seeks a perimeter-minimizing partition into $n$ regions of given areas. For $n = 4$, it is conjectured that a tetrahedral partition is minimizing. We prove that there exists a unique tetrahedral partition into given areas, and that this partition has less perimeter than any other partition dividing the sphere into the same four connected areas.
1 Introduction

There is a legend - probably apocryphal, but popular among mathematicians - that Queen Dido of Carthage, when founding her city, was promised that she could have as much land as she could encompass with a single oxhide. According to the legend, Dido cut the oxhide into fine strips and laid it out in a circle, thereby maximizing the land enclosed by the hide. This legend is popular among mathematicians because it provides a perfect introduction to the soap bubble problem, which seeks the least-perimeter way to partition a surface into \( n \) given areas - in this case two given areas, an interior and an exterior. It was proved over a century ago that if one takes the surface of the Earth to be a sphere, a circle is indeed the most efficient boundary [B] (even before that the circle was known to be optimal in \( R^2 \), but outside of small areas the plane is a suboptimal model for the entire Earth). After Dido’s solution was validated, however, the field of sphere partitioning stalled. It was only in 1994 that a standard double bubble was proved to be minimizing for \( n = 3 \), and research is still ongoing for the \( n = 4 \) case.

One of the main reasons the problem remains unsolved is the difficulty of proving that each region of an optimal partition must be connected. For instance, a region might have three components each with the same area. Intuitively, it seems as though such a partition should be suboptimal, and it is generally conjectured (see, for instance, [Q, Conjecture 2.23]) that optimal partitions do have connected regions; however, a partition with some disconnected region may have shorter perimeter than a partition with connected regions but long or erratic boundaries, and there is no simple proof that the minimizer has connected regions. If this conjecture were to be true, it would follow that for \( n = 4 \) the minimizer is always the partition with graph structure of a regular tetrahedron (see Figure 1). Engelstein [E] has proved the tetrahedral partition minimizing for the particular case where all four areas are equal; we will consider the general problem for \( n = 4 \).

Figure 1: The conjectured minimizer for \( n = 4 \), for equal and unequal areas. The equal-areas image is based on an image from [AT], copyright 1976 Scientific American.

In the plane, Wichiramala [W] proved in 2002 the standard triple bubble (which corre-
sponds to the tetrahedral sphere partition) is minimizing. His proof resulted from a PhD thesis, with a great number of cases to consider, and it is not known whether the same arguments would work on the sphere.

In this paper, after some discussion of the general soap bubble problem on the sphere, we focus on the specific set of partitions in which all four regions are connected. We prove that among this class of partitions, the tetrahedral partition is indeed minimizing:

**Main Result.** There is a unique tetrahedral partition of the sphere into any given areas, up to isometries of the sphere. This partition is perimeter-minimizing.

This result implies that there exists a tetrahedral partition with any given combination of areas; as it turns out, our Proposition 3.2 proves that there exists a unique equilibrium tetrahedral partition enclosing given areas, up to isometries of the sphere. Our proof involves using stereographic projection from a particular vertex of a tetrahedral partition to map the partition onto the plane, scaling and translating it until its image back on the sphere has appropriate areas, and then showing that any equilibrium tetrahedral partition with appropriate areas has to map to the same planar partition.

1.1 Overview of the Proof of the Main Result.

There are two separate parts to the Main Result. First, Proposition 3.2 proves that there exists a unique equilibrium tetrahedral partition enclosing given areas, up to isometries of the sphere. Our proof involves using stereographic projection from a particular vertex of a tetrahedral partition to map the partition onto the plane, scaling and translating it until its image back on the sphere has appropriate areas, and then showing that any equilibrium tetrahedral partition with appropriate areas has to map to the same planar partition.

The second part of the Main Result is Theorem 6.2, which states that the minimal connected partition of the sphere is a tetrahedron (which must be the unique equilibrium tetrahedron). Our proof of this result largely follows the proof of the corresponding planar result by the 1992 SMALL Geometry Group [CHH]. Like them, we begin by discussing the soap bubble problem under the additional constraint that every region be connected. Although this constraint makes case-by-case analysis easier, as the number of components is now known, the requirement that regions be connected weakens regularity so that minimal partitions may in principal have some edges overlapping one another. Lemma 4.5 proves that certain curves do not bump themselves, eventually allowing us to show that an equilibrium tetrahedral partition does not have overlapping edges. Lemma 4.3 proves that every region of a minimal connected partition has at least two edges and Proposition 4.6 determines the overall number of edges. Lemma 4.7 combines these results to reduce the possible connected minimizers to a tetrahedron and a two-digon, two-quadrilateral partition.
In order to eliminate the two-digon partition, we then consider the soap bubble problem when curves and areas are allowed to overlap. This is where the greatest difficulty arises in adapting the arguments of [CHH] to the sphere. In the plane, for any bubble cluster enclosing finite area there must be a point on the exterior with multiplicity 0, which allows us to easily determine the multiplicity of overlapping regions. On the sphere, it is entirely possible for a bubble cluster to enclose the whole sphere (even multiple times) while still having finite area, so it is impossible to determine the multiplicity of any given region by looking only at the curve enclosing it. To solve this problem, we determine the multiplicity of each region when considering the vector of areas \((A_1, \ldots, A_n)\) modulo \(4\pi\). Definition 5.1 covers this procedure, as well as defining the soap bubble problem under these very different circumstances. We then prove in Proposition 5.4 that minimizing overlapping bubble clusters still behave somewhat analogously to non-overlapping minimizing clusters. Lemma 5.2 shows that the standard circle is still minimizing for given area among overlapping bubbles. The bulk of Section 5, however, is devoted to proving Proposition 5.6, which states that a non-degenerate partition with two digons and two quadrilaterals cannot be length-minimizing for its combinatorial type.

Section 6 contains the actual proof of Theorem 6.2. Having already reduced the possible candidates to two, we use Proposition 5.6 to eliminate the two-quadrilateral partition. Meanwhile, Lemma 4.5 allows us to prove Lemma 6.1 which states that tetrahedral partitions satisfying the connected-regions regularity conditions do not bump themselves.

1.2 Other Surfaces

The soap bubble problem has been considered on a variety of other surfaces over the years. In the plane, as mentioned above, the answer is known for \(n = 2, 3, 4\). The case \(n = 2\), which seeks to enclose one region of given area, has been solved for many surfaces, surveyed in [HHM], circular cylinders (a small circle for smaller areas, an annular band for larger areas), circular cones (a horizontal circle), flat tori or Klein bottles (a circle or a band), hyperbolic surfaces (partial solution), and the plane with various densities (see [DDN, Introduction and Theorem 3.16]).

The case \(n = 3\) has been solved in Gaussian space for nearly equal areas, and for general areas on flat two-tori (where there are five possible minimizing configurations, depending on the torus and the given areas). See [Mor2, Chapter 19] for a survey.

1.3 Open Questions

The primary open question relating to the soap bubble problem for \(n = 4\) is whether the tetrahedral partition really is minimizing for general areas (Conjecture 3.1). Theorem 6.2 proves that it is minimizing among partitions with connected regions, but the question of how it compares to partitions with disconnected regions is still largely untouched. We looked
at the proof of connectedness for \( n = 3 \), but could not see a way to generalize it to \( n = 4 \). The equal-areas proof for \( n = 4 \) also fails to generalize, as it relies on a result of Quinn [Q, Theorem 5.2] which is specific to the case where the highest-pressure region has area \( \pi \). One obvious course is to try adapting Wichiramala’s arguments [W] to the sphere.

For the \( n = 6 \) equal-areas case, there is a clear conjecture (the geodesic cube), but no proof. One possible line of attack might be to prove that if the high-pressure region is connected, then the partition is a cube, or even that if all the regions are connected then the solution is a cube; a similar result due to Quinn [Q, Theorem 5.2] played a key role in the solution to the equal-areas case for \( n = 4 \). Combinatorial arguments along the general lines of those used to prove Lemma 4.7 can reduce the connected candidates to the cube and a doubly truncated tetrahedron, but there is no known way to eliminate the tetrahedron. Simon Cox has run computer simulations suggesting candidate equal-area minimizers for values of \( n \) up beyond 30 [C], but there is no proof that any of them are minimizing. Finally, in the connected-regions version of the soap bubble problem, it is still not clear whether there exist minimizing partitions with bumping edges.

1.4 Acknowledgements

The author would like to thank his undergraduate adviser Frank Morgan for his patience and helpful suggestions throughout this process. Some of the work was done during the 2008 SMALL summer undergraduate research program; the author thanks fellow Geometry Group members Jon Dahlberg, Alex Dubbs, Quinn Maurmann, and Hung Tran for their help during the summer, and Williams College and the National Science Foundation for their support of SMALL. The author also thanks Conor Quinn [Q], whose Williams senior thesis inspired much of that early work. The majority of the work, however, was done as the author’s senior thesis [N], and the author thanks the Williams College math department for the opportunity to do this kind of research as an undergraduate. The author thanks Simon Cox and Max Engelstein for their helpful and thought-provoking correspondence, as well as for sharing some of their related work with him. The author thanks Satyan Devadoss for his helpful suggestions as a second reader.

2 Soap Bubbles and Sphere Partitions

The first thing we need in order to discuss the soap bubble problem is a clear definition of the problem. There are two definitions commonly in use. Both start with a smooth Riemannian surface \( M \) with area \( A \), and some given areas \( A_1, \ldots, A_n \) summing to \( A \). The first way of stating the soap bubble problem seeks the least perimeter-way to partition \( M \) into regions of area \( A_1, \ldots, A_n \). This definition is natural when intending to treat all the \( A_i \) the same way. However, if it is more convenient to have an exterior region which is treated differently from the rest (perhaps because \( M \) is some manifold like \( \mathbb{R}^2 \) which has infinite area, so one of the \( A_i \) must be infinite in size), then the natural way of phrasing the problem is to seek the
least-perimeter way to enclose and separate areas $A_1, \ldots, A_{n-1}$. Since $A_1, \ldots, A_n$ sum to $A$, the exterior of such a bubble cluster will always have area $A_n$, and the two formulations of the problem are therefore fundamentally equivalent. We will use both formulations of the problem as seems appropriate.

For the two smallest values of $n$, the problem has already been solved:

**Proposition 2.1.** (Bernstein, [B]) For given area $0 < A < 4\pi$, a curve enclosing area $A$ on the sphere has perimeter $P \geq \sqrt{A(4\pi - A)}$, with equality only for a single circle.

**Proposition 2.2.** (Masters, [Ma, Thm. 2.2, Thm. 2.10]) Given areas $A_1, A_2$, with $A_1 + A_2 < 4\pi$, the least-perimeter bubble cluster on the sphere enclosing and separating regions of area $A_1, A_2$ is the unique standard double bubble with appropriate areas.

For higher $n$, the problem has not been solved, but there are still some things we know to be true about its solution, not least among them the fact that said solution exists:

**Proposition 2.3.** [Mor1, Thm 2.3, Cor. 3.3] Given a smooth compact Riemannian surface $M$ and positive areas $A_i$ summing to the total area of $M$, there is a least-perimeter partition of $M$ into regions of area $A_i$. It is given by finitely many constant-curvature curves meeting in threes at $120^\circ$ at finitely many points.

Although the full proof of Proposition 2.3 would be redundant, there is one idea involved which is sufficiently significant to this paper to merit an explanation. In order to prove that curves in a minimal partition meet in threes, Morgan shows that given two curves meeting at an angle less than $120^\circ$ it is always possible to reduce perimeter by moving the two curves in the direction of the sum of their tangent vectors at the meeting point, then adding another edge back to the original meeting point. Since a point where four or more curves meet must necessarily have two curves meeting at less than $120^\circ$, it follows that curves must meet in threes.

Now that we have a general idea of what the solution to the soap bubble problem looks like, we can begin to define the elements of a potentially minimizing sphere partition.

**Definition 2.4.** A bubble cluster is a set of finitely many $C^1$ curves which partition the sphere into several distinct regions. The curves may overlap, but the interior regions they enclose must be disjoint (meaning that although the curves may share common points, or even entire segments, they cannot completely cross over one another). If two curves share a common segment, the length of that segment is counted twice when determining the perimeter; the region between the two curves is also considered to occupy that segment, which as a one-dimensional object has area 0.

**Definition 2.5.** A point where two curves come together to overlap is called a merge.
Figure 2: In [Mor1, Thm 2.3, Cor. 3.3] Morgan proves that it is possible to reduce perimeter by splitting a point where four curves meet in two and then connecting the two new points.

**Remark 2.6.** With the current version of the soap bubble problem, it is clear that the curves of a minimal bubble cluster will not overlap; any merge would effectively be a vertex of degree four, contradicting regularity. However, later in the paper we will discuss a variant of the soap bubble problem where overlapping is not so easy to exclude, so we permit overlapping here in order to allow more broadly applicable definitions and results.

**Definition 2.7.** A region may consist of multiple connected components. For instance, if the mainland United States of America are considered as a region of North America, Alaska constitutes one component and the main 48 states constitute another.

**Definition 2.8.** If the cluster consists of constant-curvature curves meeting in threes as 120°, as in Prop. 2.3, we say it is a regular bubble cluster.

**Definition 2.9.** A component bounded by exactly two edges is called a digon. In a regular bubble cluster, both points where the edges of a digon meet will have one other edge not bordering the digon; this third edge is called an incident edge of the digon.

**Definition 2.10.** Two bubble clusters share a combinatorial type if one can be continuously deformed to another without edges crossing or shrinking to a point (although overlapping edges can be separated).

**Remark 2.11.** Since a bubble cluster is a finite set of curves embedded in the sphere, it can also be represented by a planar graph in which the curves correspond to edges, curves’ meeting points correspond to vertices, and components correspond to faces. The one slightly tricky question is how to handle circles, which do not meet any other edges and would therefore correspond to a loop with no vertex; we resolve this by assigning them a single arbitrarily located vertex. We also have the pleasant result that, since curves in a regular
bubble cluster meet in threes, every vertex of the corresponding graph will have degree three and so regular clusters correspond to regular graphs. This paper will only need one graph theory result of Euler’s which will be quoted when used, but the terminology is often convenient (particularly ‘vertices’ for ‘meeting points’) and will be used frequently.

**Definition 2.12.** Two bubble clusters share a *combinatorial type* if they correspond to the same planar graph.

One very obvious characteristic of a minimizing partition is that if you vary the shape of the minimal bubble cluster while holding area constant, perimeter does not increase. In fact, it turns out the instantaneous change of perimeter must be zero, for if perimeter increased instantly then it would be possible to create an equal and opposite variation, whose change in area and perimeter would be the opposite of the original variation’s: zero and negative, respectively, contradicting minimality. To put these concepts in more rigorous terms,

**Definition 2.13.** We define a *variation* of a bubble cluster $B$ to be the set of images $B_t$ of $B$ under a smooth family of diffeomorphisms $f_t : S^2 \times [0,a) \to S^2$ such that $B_0 = B$. $B$ is in *equilibrium* if for every variation $B_t$ of $B$ in which $B_t$ encloses the same areas as $B$ for each $t$, we have

$$\frac{d\ell(B_t)}{dt} |_{t=0} = 0.$$

(1)

It follows immediately that a length-minimizing partition is in equilibrium.

**Remark 2.14.** Although the minimizing partition is in equilibrium, it is far from true that equilibrium partitions must be minimizing. For instance, a bubble cluster consisting of concentric circles is in equilibrium; however, by translating one of the circles until it collides with another, we can create a vertex of degree four, contradicting the regularity conditions that any minimizer is known to meet.

In order to determine when a partition is in equilibrium, it is useful to have an explicit formula for the change in perimeter caused by a variation:

**Proposition 2.15.** [HMR, Lemma 3.1] Consider a bubble cluster $B$ in a 2-manifold and a variation $B_t$. Let $u$ be the vector field determined by the initial velocity of $B_t$. Let $u_{ij}$ be the component of $u$ normal to the interface $\Sigma_{ij}$ between regions $R_i$ and $R_j$. Let $\Sigma_{ij}$ have curvature $\kappa_i$, and let $T_i(p)$ be the unit tangent vector to the $i$th incident curve at a point $p$. Then for the first variation of perimeter we have

$$\frac{d\ell(B_t)}{dt} |_{t=0} = -\sum_{0<i<j} \int_{\Sigma_{ij}} \kappa_{ij} u_{ij} - \sum_p u \cdot (T_1(p) + T_2(p) + T_3(p)).$$

(2)

**Lemma 2.16.** (after the planar case in [CHH, Lemma 4.1]) A bubble cluster $B$ in equilibrium must satisfy the following conditions:

1. The edges of $B$ have constant curvature, except that they may change curvature at merges.
2. At each vertex, the sum of the unit tangent vectors of the incident edges is zero. In particular, the edges incident to a degree three vertex meet at 120° angles.

3. For every path starting and ending in the same region that crosses edges transversely and never crosses merges or vertices, the sum with multiplicities of the oriented curvatures of the edges crossed is zero.

Proof. In addition to being similar to [CHH, Lemma 4.1], (1) and (2) follow from standard variational arguments, as in [Mor1, Thm. 3.2], so this paper will merely sketch an outline rather than attempt a fully rigorous proof. The core idea behind (1) is that If an equilibrium partition had an edge with non-constant curvature, then a variation which pushed that edge in one direction at a point with curvature \( \kappa_a \) and the same in the other direction at a point with curvature \( \kappa_b \) while leaving the rest of the partition intact would, by the first variation formula in Proposition 2.15, cause a non-zero change in perimeter, contradicting the definition of equilibrium. Similarly, since the sum of the unit tangent vectors at each vertex provides another part of the first variation formula, a properly constructed variation can prove (2) by contradiction.

To prove (3), we use the same fundamental idea as the proof of (1). Suppose a path \( \alpha \) goes through regions \( R_0, R_1, \ldots, R_n = R_0 \) crossing edges \( e_1, \ldots, e_n \) with oriented curvatures \( \kappa_1, \ldots, \kappa_n \) at points \( p_1, \ldots, p_n \), in that order. We can set the orientation of \( \alpha \) so that the sum of the oriented curvatures is nonpositive. For small \( t \), let \( B_t \) be the bubble cluster obtained by adjusting each \( e_i \) near \( p_i \) so that area \( t \) is transferred from \( R_i \) to \( R_{i+1} \). (We take care to only adjust constant-curvature segments of a given edge, avoiding merges. If the path crosses several overlapping edges, we adjust them all the same way.) Then we have

\[
\left. \frac{df(B_t)}{dt} \right|_{t=0} = \sum_{i=1}^{n} \kappa_i. \tag{3}
\]

Since \( B \) is in equilibrium, \( \sum \kappa_i = 0 \).

Equilibrium bubble clusters have many nice properties. One of the most significant to soap bubble problems as a whole is that, as a result of item (3) above,

**Proposition 2.17.** For a bubble cluster in equilibrium, each region has a pressure, defined up to addition of a constant, so that the sum of the curvatures crossed by a path between two regions is the difference of their pressures. We denote the pressure of a region \( R_i \) by \( pr(R_i) \).

Proof. It suffices to show that any two paths \( \alpha, \gamma \) from some region \( R_i \) to another region \( R_j \) (not necessarily distinct) must cross the same oriented curvatures. Let \( \beta \) be a path from the endpoint of \( \alpha \) to the endpoint of \( \gamma \). By Lemma 2.16 (3), the sum of the oriented curvatures crossed by \( \alpha + \beta - \gamma \) will be zero, as it is a path starting and ending in \( R_i \). Further, the sum of the oriented curvatures crossed by \( \beta \) must be zero as \( \beta \) begins and ends in \( R_j \). The result follows immediately.
In the plane, the exterior region is traditionally taken to have pressure 0. On the sphere, where there is no exterior region, we generally take the lowest-pressure region as having pressure 0. The great advantage of the concept of pressure is that it provides some information as to the shape of a bubble cluster’s edges, with edges bulging from a higher-pressure region into a lower-pressure region proportionally to the difference between the two pressures. This can be useful for numeric calculations (see [E] for an example of this kind of argument), or simply to tell which face an edge is curving into (which is what this paper will use pressure for).

In fact, we can define the pressure of a region for any bubble cluster satisfying property (3) of Lemma 2.16. Doing so allows us to prove that the three properties described in Lemma 2.16 are sufficient qualities for equilibrium.

**Lemma 2.18.** Let $B$ be a bubble cluster with the properties described in Lemma 2.16. Then $B$ is in equilibrium.

**Proof.** (after [Q, Cor. A.4]) Assume these three conditions hold. Note that Prop. 2.17 must hold as a direct consequence of condition 3. Let $B_t$ be a variation of $B$. Let $\vec{F}$ be the vector field determined by the initial velocity of $B_t$. Let $\Sigma_{ij}$ be the interface between $R_i$ and $R_j$. Let $F_{ij}$ be the component of $\vec{F}$ normal to $\Sigma_{ij}$. Let $\kappa_{ij}$ be the curvature of $H_{ij}$, oriented from $R_j$ into $R_i$; since Prop. 2.17 holds, we have $\kappa_{ij} = pr(R_i) - pr(R_j)$. By the first variation formula in Prop. 2.15, in which the tangents at each vertex cancel by 2.16(2), one can show that

$$\frac{d\ell(B_t)}{dt}|_{t=0} = \sum_{0<i<j} \int_{\Sigma_{ij}} \kappa_{ij} F_{ij}$$  \hspace{1cm} (4)

$$= \sum_{0<i<j} \int_{\Sigma_{ij}} pr(R_i) F_{ij} - pr(R_j) F_{ij}$$  \hspace{1cm} (5)

$$= \sum_{0<i<j} \int_{\Sigma_{ij}} pr(R_i) F_{ij} + pr(R_j) F_{ji}$$  \hspace{1cm} (6)

$$= \sum_{i=1}^{n} \sum_{j \neq i} pr(R_i) \int_{\Sigma_{ij}} F_{ij}$$  \hspace{1cm} (7)

$$= \sum_{i=1}^{n} pr(R_i) \sum_{j \neq i} \int_{\Sigma_{ij}} F_{ij}$$  \hspace{1cm} (8)

$$= \sum_{i=1}^{n} pr(R_i) \frac{dA_i}{dt}|_{t=0}$$  \hspace{1cm} (9)

as each $\sum_{j \neq i} \int_{\Sigma_{ij}} F_{ij}$ is the flux of $\vec{F}$ over the boundary of $R_i$. If $B_t$ is area-preserving, this clearly simplifies to the definition of equilibrium. \qed
We are now beginning to assemble specific tools that will be needed for our main argument. Another virtue of equilibrium partitions is that they behave very nicely when mapped onto the plane by stereographic projection:

**Lemma 2.19.** Stereographic projection maps regular equilibrium bubble clusters in the plane to regular equilibrium bubble clusters on the sphere and vice versa.

**Proof.** Suppose we have a regular equilibrium bubble cluster $B$ in the sphere being projected to a bubble cluster $B'$ in the plane. Stereographic projection is conformal, so preserves angles. It also maps constant-curvature edges to constant-curvature edges. Thus $B'$ must satisfy equilibrium conditions (1) and (2), as well as the regularity conditions. To prove $B'$ satisfies equilibrium condition (3), it suffices to show that it satisfies condition (3) in the immediate vicinity of its vertices, as we can decompose any path $\phi$ into a union of loops around vertices and pairs of opposed edges which cancel out (since all edges have constant curvature, we can move the crossing point or add consecutive opposed crossings without changing the total curvature crossed by the path). However, near the vertices $B'$ consists of constant-curvature edges meeting in threes at $120^\circ$, so is locally identical to the standard planar double bubble, which is known to be perimeter-minimizing [FAB] and hence in equilibrium. The proof in the other direction is identical, as by Proposition 2.2 the double bubble is also minimizing on the sphere. 

Our final general lemma is a result which is far easier to state using the second formulation of the soap bubble problem (seeking the minimal partition enclosing and separating areas $A_1, \ldots, A_{n-1}$):

**Lemma 2.20.** (identical to the planar case in [CHH, Lemma 3.3]) If $A$ and $B$ are length-minimizing for their areas and combinatorial types, then $A \cup B$ is length-minimizing for its areas and combinatorial type. Further, if $A$ and $B$ are uniquely length-minimizing, then any length-minimizing bubble cluster with the same combinatorial type and enclosed areas as $A \cup B$ is the union of a copy of $A$ and a copy of $B$.

**Proof.** Let $C$ be any bubble cluster with the same combinatorial type as $A \cup B$, enclosing the same areas. $C$ is the union of bubble clusters $A'$ and $B'$ which have the same combinatorial type and enclosed areas as $A$ and $B$, respectively. Since $\ell(A') \geq \ell(A)$ and $\ell(B') \geq \ell(B)$, $\ell(C) \geq \ell(A \cup B)$, so $A \cup B$ is length-minimizing. Uniqueness follows by a similar argument.

### 3 The Soap Bubble Problem For $n = 4$

With the $n = 2$ and $n = 3$ cases solved (see Propositions 2.1 and 2.2), the next open question in sphere partitioning is what happens when $n = 4$. In the special case when $A_1 = A_2 = A_3 = A_4 = \pi$, Engelstein [E] shows that the perimeter-minimizing bubble cluster is a geodesic tetrahedron. It is conjectured that regular tetrahedral partitions will be minimizing for any given areas; however, this result is not yet proven.
Conjecture 3.1. Given areas $A_1, A_2, A_3, A_4$ summing to the area of the sphere, the solution to the soap bubble problem for $A_1, A_2, A_3, A_4$ is a regular tetrahedral partition.

One obvious question about this conjecture is whether such a partition actually exists. Is there a regular tetrahedral partition of the sphere enclosing any given areas? It turns out that not only can we find such a partition for any given areas, it is unique.

Proposition 3.2. Given any areas $A_1, A_2, A_3, A_4$ such that $A_1 + A_2 + A_3 + A_4 = 4\pi$, there exists a regular tetrahedral partition of the sphere with areas $A_1, A_2, A_3, A_4$. This partition is unique up to isometries of the sphere.

Figure 3: By varying the location and scale of a regular partition of the plane, we can vary the areas resulting from projection onto the sphere.

Proof. By rotating and flipping the sphere, we can reorder the areas of any tetrahedral partition as we see fit. We therefore assume without loss of generality that $A_1 \geq A_2 \geq A_3 \geq A_4$. Consider a partition of the plane with one triangle and three infinite regions separated by constant-curvature edges meeting at 120°, as in Figure 3. Such a partition can be found by stereographic projection of a geodesic tetrahedron on the sphere, using one of the vertices as the projection point. Label the triangle $R_4'$ and label the infinite regions $R_1', R_2', R_3'$. Label the regions derived by taking the same stereographic projection back onto the sphere $R_1, R_2, R_3, R_4$ and denote their areas by $|R_i|$.

To begin, scale $R_4'$ until $|R_4| = A_4$. $R_1, R_2, R_3$ are still congruent, so $|R_1| = |R_2| = |R_3| = (4\pi - A_4)/3$. Thus if $A_1 = A_2 = A_3$ then $R_1, R_2, R_3, R_4$ will form the desired tetrahedral partition. Otherwise, we translate the planar partition in the direction of $R_3'$ along a vector making an angle $\theta \in [0, \pi/3]$ with the line between $R_2'$ and $R_3'$, scaling as we go to preserve $|R_4|$ (see Figure 3). The translation will move the center of the projection away from $R_3'$.
and the scaling will add area to $R'_4$ at the expense of the other three regions, so this process must decrease $|R_3|$. Continue until $|R_3| = A_3$.

Now, if $\theta = 0$, the origin lies on the line of symmetry between $R'_2$ and $R'_3$, so $|R_2|$ will be the same as $|R_3| = A_3$. $A_3$ must be less than $|R_1|$ as $|(R_1 + R_2 + R_3)|$ is $4\pi - A_4 = A_1 + A_2 + A_3 > 3A_3$ since we are in the case where $A_1 \neq A_3$. As $\theta$ increases, the ratio of $|R_2|$ and $|R_1|$ will increase until when $\theta = \pi/3$ the origin lies on the line of symmetry between $R_1$ and $R_2$ and $|R_1| = |R_2|$. Since $|R_1 + R_2| = |R_1| + |R_2|$ is fixed $(4\pi - A_4 - A_3)$, increasing the ratio between them corresponds to increasing $|R_2|$. Since $A_3 \leq A_2 \leq (4\pi - A_4 - A_3)/2$, there must be some value of $\theta$ for which $|R_3| = A_2$. Further, we can see that there will be only one value of $\theta$ for which this is the case. $|R_1|$ will then be $4\pi - A_4 - A_3 - A_2 = A_1$, so projecting back onto the sphere gives us a tetrahedral partition with appropriate areas. Since nothing we have done in the plane would affect angles and constant curvature, and stereographic projection preserves regularity by Lemma 2.19, our new sphere partition will be regular as desired.

To prove uniqueness, we first note that stereographic projection of a regular tetrahedral partition from one of its vertices must produce an equilateral triangle with three infinite regions adjacent. Projecting from one of the vertices guarantees three infinite regions separated by straight lines, and since the sum of the curvatures at each vertex must be zero it follows that all three sides of the triangle must have equal curvature. Since all three sides also meet at $120^\circ$, the triangle must be equilateral.

Now suppose we have two tetrahedral partitions $P_1$, $P_2$ enclosing areas $A_1 \geq A_2 \geq A_3 \geq A_4$ as above. Project each partition onto the plane from the point where the three largest regions meet. Label the regions of the planar partition $R'_1$, $R'_2$, $R'_3$, $R'_4$ as above. We can always get the regions to appear in the order shown above through isometries of the sphere (in particular rotations and reflections). Then by the above note $R'_4$ will be an equilateral triangle and $R'_1$, $R'_2$, $R'_3$ will be infinite regions. Further, since $|R_1| \geq |R_2| \geq |R_3|$, the origin (the south pole of the projection) will be moved away from the center of $R'_4$ towards $R'_4$ and at least as close to $R'_2$ as to $R'_3$. This is equivalent to what would be produced by taking the partition with the origin at the center of $R'_4$ and translating the partition towards the $R'_2$ side of $R'_4$. In other words, the projected partition must be of the kind generated by the process described in the existence part of the proof. However, that process determines a unique partition for given $A_1, A_2, A_3, A_4$. Hence stereographic projection from corresponding points maps $P_1$ and $P_2$ to the same partition in the plane. This is possible only if $P_1$ and $P_2$ are identical up to isometries of the sphere.

It would also be possible to prove the existence of a regular tetrahedral partition by showing that such a partition is the solution to some variant of the soap bubble problem for which we know a solution exists. In fact, this paper will include such an existence proof later on in Section 6. However, the great advantage of using a construction argument like the one above is that we also get uniqueness; arguing that a solution exists and that solution
must be a tetrahedral partition says nothing about whether there might be other regular tetrahedral partitions, enclosing the same area with higher perimeter.

Besides area, the other numeric quantity generally associated with regions of an equilibrium partition is pressure. After proving that there is a unique regular tetrahedral partition for any given areas, an interesting followup question is whether the same holds true for any given combination of pressures. This question is irrelevant to our main theorem; however, it is answerable by techniques which exploit an interesting relationship between bubbles in $\mathbb{R}^3$ and bubbles on the sphere:

**Proposition 3.3.** There exists exactly one regular tetrahedral partition of the sphere with given pressures up to isometries of the sphere.

**Proof.** We assume for convenience that the lowest-pressure region has pressure zero. Then the pressures of the other three regions will simply be the inverses of the curvatures of the edges separating them from the lowest-pressure region. Constant-curvature edges on the sphere are circles, which are also how pairs of spheres intersect, so we can represent these three edges as (part of) the orthogonal intersections of the unit sphere with three other spheres $S_1, S_2, S_3$ of radii $R_1, R_2, R_3$ respectively. Note that since the spheres intersect the unit sphere orthogonally, $R_1, R_2, R_3$ are determined by the curvatures of the edges. In fact, if $S_1$ intersects the unit sphere at a circle of radius $r_1$ then we have $r_1 = R_1/\sqrt{1 - r_1^2}$ or equivalently $r_1 = R_1/\sqrt{1 + R_1^2}$. Now, the three edges meet pairwise at $120^\circ$ by regularity. Combined with the fact that $S_1, S_2, S_3$ are orthogonal to the unit sphere, this means that $S_1, S_2, S_3$ must meet one another at $120^\circ$. We can therefore regard $S_1, S_2, S_3$ as the external boundary of a standard triple bubble in $\mathbb{R}^3$, as shown in Figure 4. Conversely, given a standard triple bubble in $\mathbb{R}^3$ that intersects a unit sphere orthogonally, we can define a tetrahedral partition of the sphere by looking at where the exterior of the triple bubble intersects the sphere, and that partition will have pressures determined by the radii of the triple bubble. By [Mon, Prop. 3.2] there is exactly one triple bubble in $\mathbb{R}^3$ with any $R_1, R_2, R_3$, so there is exactly one set of spheres $S_1, S_2, S_3$ for any given set of pressures.

Thus proving that there is a unique tetrahedral partition for any given pressures is equivalent to proving that any given set of $S_i$ corresponds to exactly one tetrahedral partition. To do this, we begin by finding a unit sphere that intersects an arbitrary set of $S_i$ orthogonally. Begin by taking a unit sphere orthogonal to $S_1$ and $S_2$, then taking the planar slice through the centers of $S_1, S_2$, and the unit sphere. We then have a triple bubble with fixed pressures in $\mathbb{R}^2$ and a unit circle orthogonal to two of the bubbles. Fixing the location of the unit circle and the two orthogonal bubbles, there will be two possible spots for the center of the third bubble - one on each side of the unit circle. Now use circle inversion to map the unit circle to a line. The triple bubble will be mapped to another triple bubble, with the line passing through the centers of two of the bubbles. Clearly, the third bubble will form an angle $> 90$ with the line on the bubble’s side of the line. Circle inversion doesn’t change which side
of the unit circle the third bubble is on, and it preserves angles, so the third bubble’s two possible planar positions must make an obtuse angle with the unit circle on different sides of the unit circle. In other words, one planar position will see the third bubble make an angle > 90° with the unit circle and the other will see it make an angle < 90° with the unit circle.

Now return to \( R^3 \). Here we have a triple bubble with a unit sphere orthogonal to \( S_1 \) and \( S_2 \). If the radii of all three bubbles are fixed, the center of \( S_3 \) can lie anywhere on a circle determined by fixed distance from the centers of each of the other two bubbles. If we move the center of \( S_3 \) in a continuous fashion along that circle, the angle between \( S_3 \) and the unit sphere will also vary continuously. By our analysis of the planar case, there is one point on the circle where the angle is obtuse and another where it is acute. Thus, by the Intermediate Value Theorem, there is a possible center for \( S_3 \) which would leave \( S_3 \) orthogonal to the unit sphere, meaning that there is, as desired, a unit sphere orthogonal to all three \( S_i \). In addition, the angle between \( S_3 \) and the unit sphere varies monotonically as we move the center of \( S_3 \) between the two extreme positions, so there are only two possible centers for \( S_3 \): one on either side of the plane through the centers of \( S_1, S_2 \), and the unit sphere. Reflecting across that plane will not change the partition of the unit sphere generated by the \( S_i \), so there is exactly one tetrahedral partition corresponding to \( S_1, S_2, S_3 \).

\[ \square \]

**Remark 3.4.** Analogous arguments work to show that, in general, there exists a unique \((n + 1)\)-simplex partition of \( S^n \) with given pressures.

## 4 Minimal Connected Partitions

One of the major difficulties with solving soap bubble problems is that since one region can potentially have multiple components (consider, for instance, a triangular prism in which the two triangles both belong to the same region), there are an infinite number of combinatorial types which might be minimizing. Intuitively, it seems that having multiple components would be an inefficient use of perimeter, and there are currently no known or conjectured minimizers on the sphere with multiple components; however, converting that

![Figure 4: The standard triple bubble in \( R^3 \); image from [Mon, Figure 3].](image)
intuition into a proof has proved extremely difficult. Still, given the conjecture that the minimal partition has regions of only one component (i.e. its regions are connected), there is a natural question: what is the shortest partition dividing the sphere into connected regions of given area?

Figure 5: Given two overlapping edges, it is possible to reduce perimeter by treating the point where they overlap as the intersection of four curves and following the procedure from Fig. 2; however, this procedure has the effect of splitting the region between the overlapping edges into two separate components, potentially disconnecting it.

**Remark 4.1.** Unfortunately, requiring that regions be connected causes problems with several arguments that were very useful when discussing the general version of the problem. For instance, in Remark 2.6, we pointed out that a minimal partition could not contain merges or overlapping edges because the merges could be treated as vertices of degree four in which case we would have two curves meeting at an angle less than $120^\circ$, contradicting regularity. However, if we require that regions be connected, this technique is not always legal. In Figure 5, the colored region is currently connected since both parts are enclosed on the same sides of the same four edges (two of which overlap). If we treat one of the merges as a vertex where four distinct edges meet, then suddenly the colored region is disconnected; in fact, the procedure which reduces perimeter when given an angle less than $120^\circ$ will wind up taking its two components further away from one another. Thus we cannot automatically eliminate the possibility of overlapping edges.

Other useful arguments from the general case also fail. In the general version of the problem it can be shown that there are no circles (components with only one edge) by translating hypothetical circles until they overlap another edge, creating a contradiction; we used this argument earlier in Remark 2.14 to suggest why concentric circles are an inefficient way to partition anything. A similar argument shows that there are no digons (see Definition 2.9) by lengthening one of the digon’s incident edges and shortening the other - effectively sliding it along the incident edges - until one of the digon’s edges overlaps another edge, creating the usual contradiction, or one of the incident edges has length zero, creating a vertex of degree four (this argument does need the incident edges to be distinct; for full details of either of the general soap bubble problem results discussed here, see [Q, Lemmas 2.7, 2.10, 2.11]). If the digon slides until an incident edge has length zero this argument
still works, as no regions will be disconnected by splitting the degree four vertex. In the former case, however, we can run into the situation in Figure 6, where the overlap cannot be eliminated without disconnecting a region and sliding the digon any further will result in the overlapping edges crossing one another so that their interior regions overlap illegally. The argument eliminating circles from a minimal partition will be salvaged later in this section; eliminating digons is, in general, much harder to do.

Since we can no longer eliminate the possibility that a minimal partition may have overlapping edges, the connected soap bubble problem has a slightly different existence and regularity result.

**Proposition 4.2.** [Mor1, Cor. 3.3] Given areas $A_1, A_2, \ldots, A_n$ such that $A_1 + A_2 + \ldots + A_n = 4\pi$, there is a shortest graph $G$ on the surface of the unit sphere with areas $A_1, A_2, \ldots, A_n$. (The edges, but not the faces, of $G$ are allowed to overlap. For overlapping edges, count length with multiplicity.) $G$ consists of disjoint or coincident curves of constant curvature meeting

1. in threes at $120^\circ$ angles at vertices of $G$

2. at other isolated points where the edges remain $C^{1,1}$.

Note that points of type (2) will be merges (see Definition 2.5). If two edges overlap for an extended period in this kind of intersection, we will say that they squeeze one another and the region between them. This period of squeezing must begin and end at a merge. Also note that edges must come together at $0^\circ$ in order to remain $C^1$ at merges. Finally, note that any partition satisfying these conditions without overlapping itself will be regular by the standards of the general soap bubble problem.
As Remark 4.1 suggested, it is possible to salvage the result that a minimal partition contains no circular regions for the connected-regions version of the problem. In fact, we can prove a more general result. We say that a partition has connected boundary if the edges of the partition form a connected set counting only intersections of type (1) above. Edges which merely overlap one another do not count when determining connectedness of the boundary. Note that any partition containing a circle will necessarily not have connected boundary unless that circle is the only edge in the partition.

**Lemma 4.3.** The boundary of any region of a minimal connected partition $P$ of the sphere is connected. Furthermore, the partition itself has connected boundary.

*Proof.* Suppose there is some nonempty set of edges $E$ in $P$ none of which meet any of the other edges in $P$ (i.e. the edges not in $E$). Then, as in Quinn’s proof of the general case, $E$ can be slid until it encounters some other edge. It must encounter this edge either at a vertex of valence greater than 3 or along an overlapping edge. If they meet at a vertex, we can then replace the vertex with two or more vertices of valence 3, decreasing perimeter while preserving area. If they meet along an edge, there will be at least one merge, which we can treat as a vertex of valence greater than 3, allowing for the same procedure as before. In either case, we claim the new partition created by this change must still have connected regions. The only region whose topological type is affected by the change is the region between $E$ and the rest of $P$, which passed through the vertex of valence greater than three or the overlapping edges but will not pass through the new vertices-and-edge set. However, the region between $E$ and the rest of $P$ had previously completely encircled the region(s) enclosed by $E$, so it will still be connected going the other way around $E$. Thus we have a new partition with connected regions and smaller perimeter than the original minimal connected partition, a contradiction. 

Since we can now have merges and overlapping edges in equilibrium partitions, we need to think about how the curvature of overlapping edges behaves, in particular with respect to the pressures of the regions in question.

**Lemma 4.4.** (after the result for the plane in [CHH, Lemma 5.1]): Let $B$ be a bubble cluster in equilibrium. Suppose edges $e_1, \ldots, e_n$ bump at $p$, and let $R_0, \ldots, R_n$ be the regions they separate, in that order. Then for $0 \leq i \leq n$,

$$Pr(R_i) \leq \frac{n-i}{n} pr(R_0) + \frac{i}{n} pr(R_n).$$

(10)

*Proof.* Let $\kappa_i$ be the curvature of $e_i$ at $p$, oriented in the direction of increasing indices. (If the curvature of $e_i$ is not defined at $p$, carry out the subsequent argument near $p$.) Let

$$q_i = pr(R_0) + \sum_{j=1}^{i} \kappa_j.$$  

(11)
Clearly $\kappa_1 \leq \kappa_2 \leq \ldots \leq \kappa_n$, and since $q_n = pr(R_n)$ it follows that

$$q_i \leq \frac{n-i}{n}pr(R_0) + \frac{i}{n}pr(R_n).$$

(12)

Given $0 \leq i \leq n$, let $\phi$ be a path whose endpoints lie in $R_i$ and $R_0$. Consider a variation of $B$ in which $B_i$ is obtained by first pulling edges $e_1, \ldots, e_i$ away from $e_{i+1}, \ldots, e_n$, so as to transfer area $q$ from $R_0$ to $R_i$, and then transferring area $t$ from $R_i$ to $R_0$ along $\phi$, as in the proof of Lemma 2.16(3). Since $B$ is in equilibrium,

$$0 \leq \frac{dl(B_t)}{dt}|_{t=0} = \sum_{j=1}^i \kappa_j + pr(R_0) - pr(R_i) = q_i - pr(R_i),$$

(13)

so $pr(R_i) \leq q_i$. \hfill \Box

Although we cannot eliminate the possibility of merges in general, we can prove that particular edges do not merge under some specific circumstances; this will be useful when we try to prove that specific partitions (such as an equilibrium tetrahedral partition) do not have overlapping edges.

**Lemma 4.5.** (after the result for the plane in [CHH, Lemma 5.3]): Let $B$ be a bubble cluster in equilibrium whose vertices all have degree three. Let $\gamma$ be a closed curve composed of edges in $B$; we suppose that $\gamma$ does not cross but may overlap itself. Suppose the regions $U$ and $V$ into which $\gamma$ divides the sphere are such that the pressure of every region of $B$ in $U$ is greater than or equal to the pressure of every region in $V$. Suppose also that $\gamma$ has no more than three $60^\circ$ exterior angles (where $V$ is the "interior"). Then $\gamma$ does not overlap itself.

**Proof.** Suppose $\gamma$ overlaps itself. We claim that there is a segment $\gamma'$ of $\gamma$ that starts and ends at a merge $p$ and bounds a region $V' \subset V$. This is evidently true when there are two points of $\gamma$ that merge with $V$ between them. However, there must be two such points, because if not, then there are two points of $\gamma$ which merge with $U$ between them, but do not merge with anything else; however, near this merge, one of the two portions of $\gamma$ containing the two points must bulge outward from $V$, contradicting our pressure assumptions.

So $\gamma'$ and $V'$ exist. Now assign an orientation to $\gamma'$. Observe that wherever segments of $\gamma'$ bump, they alternate orientation.

Let $\kappa$ denote the curvature of $\gamma'$ (oriented inward towards $V'$) and let $\theta_1, \ldots, \theta_n$ be the exterior angles of $\gamma'$ (where $V'$ is the "interior") at vertices of $B$. Since the two ends of $\gamma'$ are tangent with opposite orientations, they create an exterior angle of $\pi$, and thus

$$\int_{\gamma'} \kappa + \sum \theta_i = \pi.$$  

(14)
We claim that $\int_{\gamma'} \kappa \leq 0$. If an even number of segments of $\gamma'$ overlap, then their curvatures cancel in the integral, since their orientations alternate. Suppose that $2k + 1$ segments of $\gamma'$ overlap (possibly with other segments from $\gamma - \gamma'$). Let $R$ and $S$ be the two outermost regions separated by the overlapping edges. One of these regions, say $R$, is a subset of $V'$, while the other is not. If $S \subset U$, then $pr(R) \leq pr(S)$ and $k + 1$ of the segments of $\gamma'$ will have $\kappa$ zero or negative, so the curvature integral over the $2k + 1$ segments will be zero or negative. If $S \subset V - V'$, then some region in $U$ is squeezed between the overlapping edges. Since this region must have pressure at least as great as $R$ and $S$, Lemma 4.4 gives $pr(R) = pr(S)$. Hence all $2k + 1$ segments have curvature zero. (Note that when $k = 0$ the segment has non-positive curvature by our pressure hypotheses).

This proves that $\int_{\gamma'} \kappa \leq 0$, or equivalently $\sum \theta_i \geq \pi$. It follows that $\gamma'$ must have at least three $60^\circ$ exterior angles. Suppose there are exactly three such angles. Then $\sum \theta_i = \pi$, so $\int_{\gamma'} \kappa \leq 0$. Near one of its endpoints, $\gamma'$ must bulge outward from $V'$. This outward bulging segment must be one of an even number of overlapping segments of $\gamma'$, since otherwise $\int_{\gamma'} \kappa < 0$. But these segments squeeze a region in $U$ between regions in $V'$, so the segments are straight by Lemma 4.4 as before. This is a contradiction, so $\gamma'$ must have at least four $60^\circ$ exterior angles.

Note that Lemmas 2.16, 2.18, and 4.4 will in fact hold for connected partitions on any surface with similar definitions of equilibrium and pressure.

Given the amount of trouble we are having eliminating the possibility of overlapping, which was trivial to dispose of in the general version of the problem, the question may arise of why the connected version of the soap bubble problem is seen as a promising line of attack. The answer is that knowing how many components there are in a partition (one per region, since all the regions are connected) is incredibly powerful:

**Proposition 4.6.** Given an integer $n$, a minimal connected $n$-partition of $S^2$ has $2n - 4$ vertices and $3n - 6$ edges.

*Proof.* Such a partition has $n$ faces. Since its edges meet in threes,

$$e = \frac{3v}{2}. \quad (15)$$

Since the Euler characteristic of a planar graph is 2, and sphere partitions can be seen as planar graphs as in Remark 2.11,

$$2 = v - e + f = v - \frac{3v}{2} + n \quad (16)$$

The result follows immediately. \qed
For a concrete example of how useful this result can be, consider the case where \( n = 4 \). We have conjectured (Conjecture 3.1) that a regular tetrahedral partition is the solution to the general soap bubble problem for any given areas; if this is the case, it will certainly be the minimizer among partitions with connected regions. We already have enough information to show that there is only one possible alternative:

**Lemma 4.7.** A minimal connected 4-partition of the sphere is either a tetrahedron or a pair of digons sharing the same incident edges (as in Figure 6).

**Proof.** By Proposition 4.6, a minimal connected partition of the sphere \( B \) has six edges. By Lemma 4.3, every region has at least two sides. If every region has at least three sides, then every region must have exactly three sides, so the partition is a tetrahedron.

Suppose one region has exactly two sides. Then the two incident edges \( \alpha_1, \alpha_2 \) must be distinct as the whole partition has connected boundary, hence four of the six edges are accounted for. The two edges branching off from the far endpoint of \( \alpha_1 \) must therefore be the same as the two edges branching off from the far endpoint of \( \alpha_2 \), meaning that \( B \) is a two-digon, two-quadrilateral partition of the kind shown in Figure 6.

\[\square\]

## 5 Overlapping Soap Bubbles

The obvious step, at this point, is to investigate the two-digon sphere partition more closely in order to eliminate it as a possible connected minimizer. To do this, we will consider yet another version of the connected-regions problem where curves and regions are allowed to overlap one another, and prove that a partition with two digons and two quadrilaterals cannot be minimizing under these conditions. This new form of the problem requires some care with definitions. First, since the total area enclosed by overlapping regions may vary from the area of the sphere, we exclusively use the definition of the soap bubble problem which seeks to separate the sphere into three regions of given area and one exterior region. Second, since regions are allowed to overlap one another, it is possible that boundary curves may cross one another, invalidating our previous definition of a bubble cluster. We resolve this issue by expanding and formalizing the graph theory definition of a bubble cluster originally sketched in Remark 2.11:

**Definition 5.1.** In order to define the concept of area for overlapping bubble clusters, we will use a somewhat different definition of bubble clusters. An overlapping bubble cluster consists of an embedded graph \( G \subset \mathbb{R}^2 \) and a \( C^1 \) map \( f : G \to S^2 \). The type of \((G, f)\) is the isomorphism class of \( G \). If one of the edges of \( G \) is mapped to a single point, we say that \((G, f)\) is degenerate (but still of the same type). The length of \((G, f)\) is the sum of the lengths of its edges.

The area of a face of \((G, F)\) is determined by counting the areas with multiplicity of individual non-overlapping components of the face. We can find the multiplicity with which
any given component $C$ is enclosed, assuming we know the multiplicity of some point $P$ on the sphere, by looking at how many times a path from $P$ to $C$ crosses the boundary of the face going from left to right. The orientation of the boundary is determined by going along the boundary of the face in $G$ so that the interior of the face is always on the left. The multiplicity of a component is path-independent since loops around a vertex add nothing to the tally (crossing each edge twice in opposite directions) and by adding such loops we can continuously deform any two paths into one another.

Since the multiplicity of the exterior of the face is arbitrary, this procedure only determines the area of the face modulo $4\pi$. Thus for the purposes of overlapping bubbles we define $I(A)$ to be the perimeter of the single circle enclosing area $A \mod 4\pi$ as given by Proposition 2.1. In general, if a bubble cluster has $n$ regions, we consider a vector of areas $(A_1, A_2, \ldots, A_n)$ modulo $4\pi$, as once we know the multiplicity of one bubble in the cluster we can determine the multiplicities of the other bubbles.

The areas enclosed by a bubble cluster will always sum to an integer multiple of $4\pi$, as a point enclosed with higher than usual multiplicity in one bubble will be enclosed with lower than usual multiplicity by adjacent bubbles whose boundaries are oriented in the opposite direction.

Our first step is to show that this new definition of $I(A)$ does indeed give the minimum perimeter required to enclose area $A$ (or, in other words, the standard isoperimetric result still holds for $0 < A < 4\pi$).

**Lemma 5.2.** The least-perimeter way to enclose given area $A$ on the surface of the sphere, with overlapping allowed, is a circle of length $I(A)$.

**Proof.** Suppose we have a perimeter-minimizing curve $\gamma$ enclosing total area $A$. If $\gamma$ does not overlap itself, then it must be a solution to the regular isoperimetric problem on the sphere for area $A \mod 4\pi$, hence is a circle.

Suppose $\gamma$ overlaps itself. Then we can decompose $\gamma$ into curves $\gamma_i$ which may touch one another but do not overlap or cross. If there is some $\gamma_i$ enclosing area with absolute value $\geq 4\pi$, we split off a region consisting of the whole sphere with no boundary and whatever multiplicity is necessary to ensure $\gamma_i$ encloses area of absolute value less than $4\pi$. The $\gamma_i$ will now enclose regions $R_i$ of area $0 < A_i < 4\pi$ with multiplicity $\pm 1$. Since $\gamma$ overlaps itself, we have $i \geq 2$. Thus we must have at least two regions of positive area or one region of positive area and one region of negative area.

If we have two regions of positive area $R_1, R_2$, then we have $\ell(\gamma_1) \geq I(A_1)$ and $\ell(\gamma_2) \geq I(A_2)$ by the standard isoperimetric result. Then since the isoperimetric function on the surface of a sphere is strictly concave, $I(A_1) + I(A_2) > I(A_1 + A_2)$, hence $\ell(\gamma_1) + \ell(\gamma_2) > I(A_1 + A_2)$. This means that we could reduce the perimeter of $\gamma$ by replacing $\gamma_1$ and $\gamma_2$ with
a circle enclosing area \( A_1 + A_2 \), contradicting \( \gamma \) minimal.

If we have one region \( R_1 \) of positive area and another region \( R_2 \) of negative area, then again \( \ell(\gamma_1) \geq I(A_1) \) and \( \ell(\gamma_2) \geq I(A_2) \). Now, since \( I(A) = I(4\pi - A) \) for all \( A \), using strict concavity as above gives us

\[
I(A_1 - A_2) = I(4\pi - A_1 + A_2) \\
< I(4\pi - A_1) + I(A_2) \\
= I(A_1) + I(A_2) \leq \ell(\gamma_1) + \ell(\gamma_2).
\]

This allows us to reduce the perimeter of \( \gamma \) by replacing \( \gamma_1 \) and \( \gamma_2 \) with a circle enclosing area \( A_1 - A_2 \mod 4\pi \), again contradicting \( \gamma \) minimal. Hence \( \gamma \) cannot overlap itself. \( \square \)

We can also use this result to show that any individual component of a minimal partition must be an arc of a circle or a geodesic arc:

**Corollary 5.3.** Given an oriented geodesic arc \( \vec{PQ} \) and a real number \( r \), the unique shortest curve \( \alpha \) from \( Q \) to \( P \) such that \( \vec{PQ} + \alpha \) encloses area is an arc of a circle or a geodesic arc.

**Proof.** A better competitor, combined with the rest of the circle (possibly crossing it), would enclose the area of the circle more efficiently and contradict Lemma 5.2. \( \square \)

We can now show that, given any areas and any combinatorial type, there exists an overlapping bubble minimizing perimeter for that combinatorial type as long as we are willing to tolerate possible degeneracies. For the rest of this section, many of the proofs will be lengthy, technical, and highly similar to the planar case in [CHH]. We will therefore simply list the results here and provide the rigorous proofs in Section 7 for readers who are interested in the details of our argument.

**Proposition 5.4.** (after [CHH, Prop. 7.4]): Let \( G \subset S^2 \) be an embedded graph with bounded faces numbered 1, \ldots, \( n \), and let \( A_1, \ldots, A_n \) be given. Then there exists an overlapping bubble of type \( G \) (which may be degenerate) such that region \( i \) has area \( A_i \), which minimizes length among all such overlapping bubbles. Any such minimal overlapping bubble satisfies:

1. All the edges are arcs of circles or geodesic arcs.
2. At any vertex, the sum of the unit tangent vectors of the incident edges is zero.
3. At any vertex, the sum of the oriented curvatures of the incident edges is zero.

**Lemma 5.5.** Suppose a length-minimizing overlapping bubble cluster, as given by Proposition 5.4, has a digon \( R_i \) with vertices \( P \) and \( Q \) of degree 3. Let \( \alpha_1 \) be the edge incident to \( P \) and not part of the digon, and \( \alpha_2 \) be the corresponding edge incident to \( Q \). Then \( \alpha_1 \) and \( \alpha_2 \) lie on the same constant-curvature arc.
Proposition 5.6. Let four areas $A_1, A_2, A_3, A_4$ be given. Let $B$ be a length-minimizing overlapping bubble cluster consisting of two digons and two quadrilaterals, as given by Proposition 5.4 and shown (in the plane) in Figure 6. Then:

1. If $B$ is nondegenerate, then its edges do not intersect except at the common endpoints.

2. If all such $B$ are degenerate, then the only minimizer given by Proposition 5.4 is a pair of intersecting circles.

6 The Connected-Regions Soap Bubble Problem for $n = 4$

We are now ready to prove that a non-overlapping equilibrium tetrahedral partition of the sphere is the solution to the connected-regions soap bubble problem for $n = 4$. We start by proving that an equilibrium tetrahedral partition is necessarily non-overlapping

Proposition 6.1. (identical to the planar case in [CHH, Lemma 6.1]) Edges do not overlap in an equilibrium tetrahedral partition of the sphere.

Proof. Let $B$ be an equilibrium tetrahedral partition with regions $R_1, R_2, R_3, R_4$. We may assume $pr(R_1) \geq pr(R_2) \geq pr(R_3) \geq pr(R_4)$. Let $e_{ij}$ denote the edge separating $R_i$ and $R_j$; let $v_{ijk}$ be the vertex at which $R_i, R_j$, and $R_k$ meet.

By Lemma 4.5, the curves $\partial(R_1), \partial(R_1 + R_2)$, and $\partial(R_1 + R_2 + R_3)$ do not overlap themselves. It follows that no edge overlaps itself, and the only pairs of edges that can possibly overlap are $(e_{12}, e_{23}), (e_{12}, e_{24}), (e_{13}, e_{34})$, and $(e_{23}, e_{34})$.

Suppose $e_{12}$ and $e_{24}$ overlap. Starting from $v_{124}$, let $p$ be the first point on $e_{24}$ which is also on $e_{12}$. Let $e'_{12}$ and $e''_{24}$ be the segments of $e_{12}$ and $e_{24}$, respectively, between $v_{124}$ and $p$. Let $e'_{12}$ and $e''_{24}$ be the segments of $e_{12}$ and $e_{24}$ between $p$ and $v_{123}$ and $v_{234}$ respectively. Since $e_{24}$ can only overlap $e_{12}$, $e''_{24}$ does not overlap any edges. It follows that $e'_{12}$ overlaps something, as otherwise $e'_{12}$ and $e''_{24}$ will have constant curvature and therefore merge at $120^\circ$, which is impossible. The only edges that $e'_{12}$ can possibly overlap are $e_{23}$ and $e''_{24}$. But the curve consisting of $e'_{24}, e''_{12}, e_{13}$, and $e_{14}$ separates $e_{12}$ from both $e_{23}$ and $e''_{24}$ (see Figure 7). This is a contradiction, so $e_{12}$ and $e_{24}$ cannot overlap. A nearly identical argument shows that $e_{13}$ and $e_{34}$ do not overlap.

Suppose $e_{23}$ and $e_{34}$ overlap. Starting from $v_{234}$, let $p$ be the first point on $e_{34}$ which is also on $e_{23}$; let $q$ be the last such point on $e_{34}$. Note that $e_{23}$ never bulges out from $R_3$, since the only regions that can be on the other side of $e_{23}$ from $R_3$ are $R_2$ and $R_1$, both of which have higher or equal pressure. From this, one shows that the segment of $e_{34}$ from $v_{234}$ to $p$ has total curvature greater than $4\pi/3$ and the segment of $e_{34}$ from $q$ to $v_{134}$ has total curvature greater than $2\pi/3$. Since $e_{34}$ always bulges out from $R_3$ with curvature at least as
large as the curvature of the segments from $v_{234}$ to $p$ and $q$ to $v_{134}$, it follows that $e_{23}$ cannot possibly overlap $e_{34}$ on the correct side.

Now the only pair of edges that can possibly overlap is $e_{12}$ and $e_{23}$, but these cannot overlap since they would have to merge at a 120° angle. \hfill \Box

**Theorem 6.2.** A minimal connected partition of the sphere into exactly four regions of prescribed area is a non-overlapping tetrahedron.

**Proof.** By Lemma 4.7, a minimal connected partition of the sphere $B$ is either a tetrahedron or a two-digon partition (see Figure 8 for a planar example of the latter). If it is a tetrahedron, it does not overlap by Lemma 6.1. Suppose $B$ is instead a two-digon partition. By Proposition 5.6, a length-minimizing overlapping bubble of the two-digon type is always non-overlapping, so $B$ is a minimizing overlapping bubble (for the areas of the two digons and either quadrilateral). By Theorem 4.2, $B$ is not degenerate. We can therefore lengthen one of the two edges shared by the two quadrilaterals ($\alpha_1$ in Figure 8) and shorten the other ($\alpha_2$) until we have a partition $B'$ with the same perimeter and areas as $B$ and a vertex of degree four. $B'$ is not of allowable type, hence cannot be length-minimizing; thus $B$ is not minimizing either. \hfill \Box

Note that this minimizer must in fact be the unique equilibrium tetrahedral partition given by Theorem 3.2. We also now have the tools for an alternative proof of the existence part of Theorem 3.2; however, it unfortunately does not imply uniqueness.

**Corollary 6.3.** Given any areas $A_1, A_2, A_3, A_4$ such that $A_1 + A_2 + A_3 + A_4 = 4\pi$, there exists a tetrahedral partition of the sphere with areas $A_1, A_2, A_3, A_4$ and constant-curvature edges meeting at 120°.

**Proof.** By Theorem 4.2, there exists a partition of the sphere into connected regions with the appropriate areas whose edges have constant curvature and meet either in threes at
120° angles or at other isolated points where the edges remain $C^{1,1}$. By Theorem 6.2, that partition is a tetrahedron. By Lemma 6.1, the edges of the tetrahedron do not bump. \qed 

7 Appendix

Here are the detailed proofs of Propositions 5.4 and 5.6, as well as Lemma 5.5.

Proposition 7.1. (after [CHH, Prop. 7.4]; identical to Prop. 5.4): Let $G \subset S^2$ be an embedded graph with bounded faces numbered 1,\ldots, n, and let $A_1,\ldots, A_n$ be given. Then there exists an overlapping bubble of type $G$ (which may be degenerate) such that region $i$ has area $A_i$, which minimizes length among all such overlapping bubbles. Any such minimal overlapping bubble satisfies:

1. All the edges are arcs of circles or geodesic arcs.
2. At any vertex, the sum of the unit tangent vectors of the incident edges is zero.
3. At any vertex, the sum of the oriented curvatures of the incident edges is zero.

Proof. By induction on the faces of $G$, one can construct an overlapping bubble of type $G$ and areas $A_1,\ldots, A_n$ all of whose edges are arcs of circles or geodesic arcs. The set of all such bubbles can be parameterized with finitely many variables, and by a standard compactness argument, there is a length-minimizing overlapping bubble $B_0$ in this set. We claim $B_0$ minimizes length among all overlapping bubbles of type $G$ with appropriate areas. This is true because, given a bubble $B$ of type $G$ with at least one non-constant-curvature edge, Corollary 5.3 allows us to reduce perimeter by replacing the non-constant-curvature edges with arcs of circles or geodesic arcs. This procedure gives a new bubble $B'$ whose length will be less than that of $B$, but greater than or equal to that of $B_0$ by definition of $B_0$. This proves existence and condition (1).

To prove (2), first observe that at any vertex, the unit tangent vectors $v_1,\ldots, v_m$ of the incident edges must form a length-minimizing network (in its type) connecting the $m$ points at their heads. If not, then there exists some $\alpha > 0$ such that for small $r$ the edges can be adjusted inside a ball of radius $r$ about the vertex (possibly changing areas) with a length decrease of at least $\alpha r$. We can restore the areas by adjusting the edges elsewhere; since the area changes were at most on the order of $r^2$, this increases length by at most $\beta r^2$ for some constant $\beta$ (cf. equation (9) in Lemma 2.18). For sufficiently small $r$, we have $\alpha r - \beta r^2 > 0$, contradicting minimality. Thus the vectors $v_1,\ldots, v_m$ form a minimizing network. If we move the center of the network in a direction $u$, then the initial change in length is $\sum_i v_i \cdot u$, but this must be zero for every vector $u$, so $\sum v_i = 0$.

The proof of (3) is essentially the same as the proof of Lemma 2.16(3). Suppose $\alpha_1,\ldots, \alpha_m$ are the edges entering a vertex. Let $\kappa_i$ denote the curvature of $\alpha_i$, oriented with respect to
a clockwise path around the vertex. Create a new variation $B_t$ by shifting area $t$ from one region to the next along this clockwise path. This will enclose the same areas as $B_0$, since each region gains area $t$ along one edge and loses area $t$ along another. Since $B_0$ is in equilibrium,

$$0 = \frac{dP(B_t)}{dt}|_{t=0} = \sum \kappa_i.$$  

(20)

**Lemma 7.2.** (identical to Lemma 5.5) Suppose a length-minimizing overlapping bubble cluster, as given by Proposition 5.4, has a digon $R_i$ with vertices $P$ and $Q$ of degree 3. Let $\alpha_1$ be the edge incident to $P$ and not part of the digon, and $\alpha_2$ be the corresponding edge incident to $Q$. Then $\alpha_1$ and $\alpha_2$ lie on the same constant-curvature arc.

**Proof.** Locally, $R_i$ is identical to a digon in a standard double bubble. The incident edges of a digon in a double bubble are both the same constant-curvature edge, hence clearly lie on the same constant-curvature arc. \hfill \Box

**Proposition 7.3.** (identical to Prop. 5.6) Let four areas $A_1, A_2, A_3, A_4$ be given. Let $B$ be a length-minimizing overlapping bubble cluster consisting of two digons and two quadrilaterals, as given by Proposition 5.4 and shown (in the plane) in Figure 8. Then:

1. If $B$ is nondegenerate, then its edges do not intersect except at the common endpoints.

2. If all such $B$ are degenerate, then the only minimizer given by Proposition 5.4 is a pair of intersecting circles.

Figure 8: The type of bubble cluster discussed in Proposition 5.6; shown here in the plane, image from [CHH, Figure 11]
Proof. (1) Suppose B is nondegenerate. Label the edges of B as in Figure 8. We want to show that no two of these edges intersect, except at common endpoints.

We first show that $\alpha_1$ and $\alpha_2$ do not intersect. Suppose that they do. By Lemma 5.5, $\alpha_1$ and $\alpha_2$ are both arcs of the same circle. There are three ways that two arcs of the same circle can intersect; these are cases (a), (b), and (c) of Figure 9. In each case, angle requirements determine the locations of the other four edges up to labeling and orientation. Consider case (a). Without loss of generality, $\alpha_1$ is oriented upward, as shown in the figure (otherwise we can rotate the sphere $180^\circ$). Then the left inner edge must be $\beta_1$, not $\gamma_1$, in order for the area enclosed by $(\beta_1 - \gamma_1)$ to be positive. But then the area enclosed by $\alpha_1 + \beta_1 + \gamma_1 + \alpha_2$ is negative, a contradiction. A similar argument eliminates case (b). For case (c), we can rotate the left-hand region counterclockwise around the circle on which $\alpha_1$ and $\alpha_2$ lie, lengthening $\alpha_1$ and shortening $\alpha_2$, until $\alpha_1$ overlaps itself. This procedure does not change perimeter, so the resulting overlapping bubble is still length-minimizing, but $\alpha_1$ is no longer an arc of a circle, contradicting Proposition 5.4(1).

Any pair of circles or circular arc on the sphere intersects at most twice; $\gamma_i$ meets $\beta_i$ and the $\alpha$-circle twice (once at each endpoint), so cannot intersect them anywhere else. Similarly, $\beta_i$ cannot overlap $\beta_i$ or either $\alpha$ edge except at common endpoints. Since the $\beta_i$ and the $\gamma_i$ are on opposite sides of the $\alpha$-circle, they cannot intersect except at their endpoints where they meet the $\alpha$-circle.

It only remains to show that $\gamma_1$ and $\gamma_2$ ($\beta_1$ and $\beta_2$) do not intersect. Label $P, Q_i, R_i$ as in Figure 10. Then $\gamma_1$ is completely enclosed by $PQ_1 + \delta_1 + R_1P$, and $\gamma_2$ is completely enclosed by $PQ_2 + \delta_2 + R_2P$. These two triangles do not intersect except at $P$, so $\gamma_1$ and $\gamma_2$ do not meet. An identical argument looking at the other side of the $\alpha$-circle shows that $\beta_1$ and $\beta_2$ do not intersect.
Figure 10: The two triangles $PQ_1 + \delta_1 + R_1P$ and $PQ_2 + \delta_2 + R_2P$ completely separate $\gamma_1$ and $\gamma_2$, image from [CHH, Figure 13]

Figure 11: The first few possible degeneracies of the two-digon partition, image from [CHH, Figure 14]

(2) Suppose all length-minimizers are degenerate. If a length-minimizer has the type labeled (d) in Figure 11, one can show that the top vertex is on the same circle as the bottom edge by an argument similar to the proof of Lemma 5.5. Hence, by rotating the two digons downwards along this circle and stretching the top vertex into an arc of the circle, we obtain a non-degenerate bubble with the same length and areas, contradicting the assumption that all length-minimizers are degenerate.

A length-minimizer of type (e) will be a pair of overlapping circles by Lemma 2.20 (which works identically for overlapping bubbles).

None of the other four degeneracies in Figure 11, or any degeneracies thereof, can be length-minimizers. To see this, observe that each of these four types is the union of a
standard double bubble type and a circle. It is immediate from Proposition 5.4 that the unique length-minimizing bubble cluster of the standard double bubble type is in fact a standard double bubble. So the shortest overlapping bubble of one of these four types is the union of a standard double bubble and a circle. But if a standard double bubble and a circle are joined together as in the bubble of type (f) in Figure 11, one can split the degree four vertex vertically into two vertices of degree three to obtain a shorter bubble (still of the two-digon type) since the top and bottom angles of the degree four vertex are less than \(120^\circ\).

\[\Box\]

References


[C] Simon Cox, Calculations of the minimal perimeter for \(N\) deformable bubbles of equal area confined to the surface of a sphere, preprint, accessed online at http://hdl.handle.net/2160/2317 on 23 May 2009.

[CHH] Christopher Cox, Lisa Harrison, Michael Hutchings, Susan Kim, Janette Light, Andrew Mauer, and Meg Tilton, The shortest enclosure of three connected areas in \(\mathbb{R}^2\), Real Analysis Exchange 20, 313-335, 1994/5.


