Real Analysis Exchange
Vol. 20(1), 1994/95, pp. 313–335

Christopher Cox, Lisa Harrison, Michael Hutchings, Susan Kim, Janette Light, Andrew Mauer and Meg Tilton.
c/o Prof. Frank Morgan, Department of Mathematics, Williams College, Williamstown, MA 01267, (email: Frank.Morgan@williams.edu)

THE SHORTEST ENCLOSURE OF THREE CONNECTED AREAS IN $\mathbb{R}^2$

Abstract

We show that the "standard triple bubble" is the shortest way to enclose and separate three areas in $\mathbb{R}^2$, assuming that the enclosed regions and the exterior region are connected.

1 Introduction

In nature, soap bubbles tend to enclose fixed volumes of air using the least possible surface area. The analogous problem in 2 dimensions is to find the shortest way to enclose and separate areas $A_1, \ldots, A_m$ in the plane. If $m = 1$, the solution is a circle. Joel Foisy, Manuel Alfaro, Jeffrey Brock, Nickelous Hodges, and Jason Zimba [F2] have proved that for $m = 2$, the unique solution is the "standard double bubble" shown in Figure 1.

In general, a troublesome possibility is that the regions enclosed may be disconnected, as shown in Figure 2(a). The exterior region might also be disconnected. We do not believe that either of these situations occurs, but we do not know how to prove this for $m > 2$.

In this paper, we seek the shortest enclosure of three areas, assuming that all regions are connected. It was proved only recently that a solution to this problem exists [M1]. When we insist that the regions be connected, we have to allow for the possibility that the curves that separate the regions may bump up against each other and change curvature, so that regions may be connected only by "infinitesimal strips", as shown in Figure 2(b). The precise statement of this existence and regularity theorem is given in §2.

Key Words: soap bubble clusters, isoperimetric problems
Mathematical Reviews subject classification: Primary: 49Q10, 52A38, 28A75, Secondary: 49Q20, 52A40
Received by the editors October 26, 1993

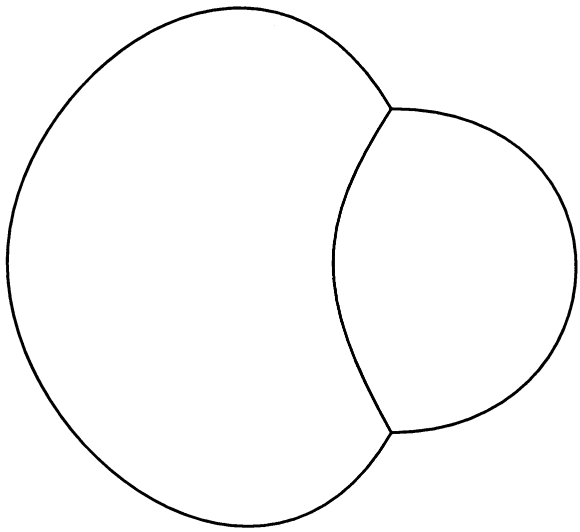

Figure 1: The standard double bubble in $\mathbb{R}^2$ is uniquely length-minimizing.

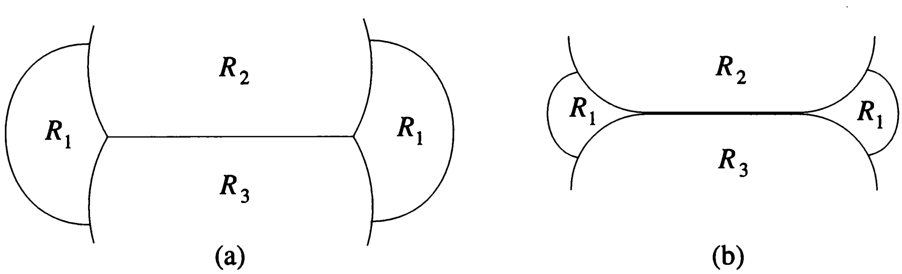

Figure 2: (a) In general, regions may be disconnected. (b) In the connected regions problem the regions may be connected only by "infinitesimal strips".

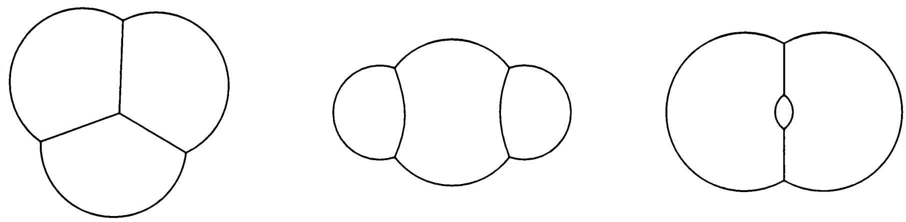

Figure 3: We first show that any length-minimizing enclosure of three connected areas has one of these three combinatorial types.

In §3, we show that there are only three combinatorial types that a solution to this problem can have. These are shown in Figure 3. Our main result is the following:

Main Theorem. *Given $A_1, A_2, A_3 > 0$, the shortest way to enclose and separate areas A_1, A_2, A_3 in $\mathbb{R}^2$ is a bubble of type (a) with no bumping.*

That there is a unique such triple bubble was proved by J. Foisy [F1, Thm. 4.6] for equal areas and by A. Montesinos Amilibia [M] in general.

To prove the Main Theorem we will find the shortest bubbles of types (b) and (c) and show that they are inferior to type (a). The main difficulty is that curves may change curvature where they bump. To get by this complication, we consider in §7 a different problem in which the regions are allowed to overlap. When the curves are allowed to move about freely in this way, they revert to circular arcs, as shown in Figure 4. It is then relatively easy to determine the length-minimizing "overlapping bubbles" of types (b) and (c), and we find in §8 that they do not have any overlapping regions. Hence we have found the shortest bubbles of types (b) and (c) for the original problem, and their curves do not change curvature. A simple argument (Theorem 8.5) then shows that these are inferior to type (a).

To complete the proof of the Main Theorem, we need to show that there is no bumping in length-minimizing bubbles of type (a). This would follow if we could prove that regions in a length-minimizing overlapping bubble of type (a) do not overlap. However, this seems to be harder than for types (b) and (c). Instead we give a direct but more complicated argument in §4–§6, using curvature considerations ("pressure"), to rule out bumping. In §4, we also present a nice formula relating area, length, and pressure for bubbles in equilibrium.

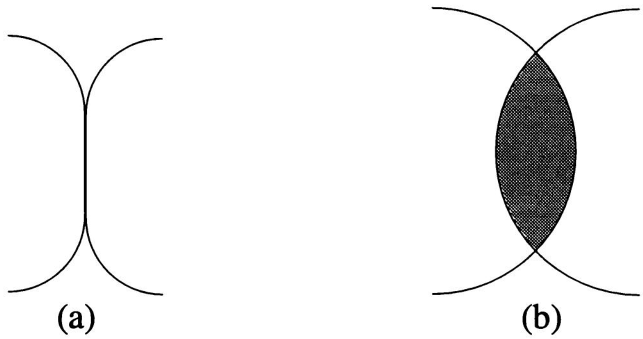

Figure 4: When we allow regions to overlap, bumped-up curves revert to constant-curvature arcs.

1.1 Acknowledgements

This paper is the work of the Geometry Group of the SMALL Undergraduate Research Project, Summer 1992, a National Science Foundation site for Research Experiences for Undergraduates (REU). For a period of nine weeks, each of twenty-three students worked in one or two of the seven groups that comprised the project. Professor Frank Morgan advised the Geometry Group.

Support for the project was provided by grants from the National Science Foundation Research Experiences for Undergraduates Program, NECUSE, and the Bronfman Science Center at Williams College.

We are indebted to Joe Masters for help with editing.

2 Existence and Regularity

We define a *bubble* to be a finite collection of immersed C^1 curves $\gamma_i : [0,1] \to \mathbb{R}^2$, called *edges*, which may overlap ("bump") but not cross. A common endpoint of one or more edges is called a *vertex*. We assume that the degree of each vertex is at least three. To handle the case of single bubbles, we also allow "edges" to be smooth maps $S^1 \to \mathbb{R}^2$, with no vertices. If two edges bump along a maximal segment (possibly of length zero), we call the endpoints of this segment *merges*. Note that edges make $0°$ angles at merges, in order to remain C^1.

The *length* of a bubble B, which we denote by $\ell(B)$, is the sum of the lengths of the edges of B. Length counts with multiplicity where edges bump.

Given a bubble B with edges $\{\gamma_i\}$, we say that two points $p, q \in \mathbb{R}^2 - \bigcup \gamma_i$

are in the same *region* if and only if there is a path $\varphi : [0,1] \to \mathbb{R}^2$, which may overlap, but not cross, the γ_i's, with $\varphi(0) = p$ and $\varphi(1) = q$. We refer to the unbounded region as the *exterior.*

The following theorem of Morgan provides the existence and regularity of length-minimizing bubbles with connected regions.

Theorem 2.1 **([M1])** *Given prescribed areas $A_1, \ldots, A_m$, there exists a length-minimizing bubble B enclosing exactly m bounded regions of areas $A_1, \ldots, A_m$. Any such B has the following properties:*

(1) The edges of B have constant curvature, except that they may change curvature at merges.

(2) Every vertex of B has degree 3, and incident edges meet at $120°$ angles.

3 Combinatorial Types

We say that two bubbles have the same *combinatorial type* if each can be continuously deformed into the other, without allowing edges to cross or shrink to points (but allowing bumping edges to be separated). If we deform a bubble B, without ever allowing edges to cross, in such a way that the lengths of some of the edges shrink to zero (but no regions disappear), and if we then fuse the endpoints of zero-length edges into single vertices, we call the type of the resulting bubble a *degeneracy* of the type of B.

An analogue of Theorem 2.1 in [M1] states that for any combinatorial type and areas, there is a length-minimizing bubble enclosing the given areas with the given combinatorial type, or a degeneracy thereof. However, we will not need this result.

3.1 Allowable Types

A combinatorial type represented by bubble B is *allowable* if B has the following properties:

(1) Every vertex of B has degree three.

(2) Every edge of B is part of the boundaries of two distinct regions; that is, B has no *cut edges*.

(3) The edges and vertices of B form a connected graph.

Lemma 3.1 *The combinatorial type of any solution to the connected regions problem, as given by Theorem 2.1, must be allowable.*

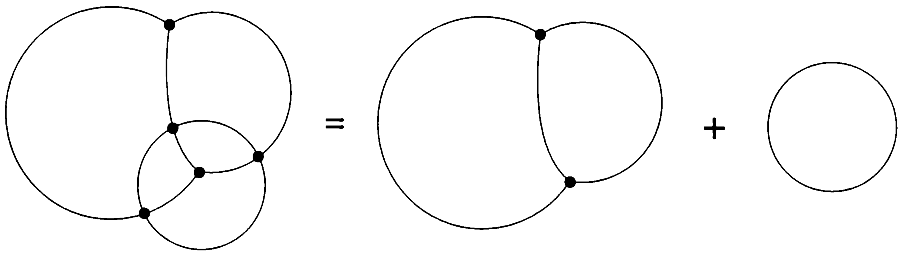

Figure 5: It is sometimes possible to place lower bounds on the length of a bubble by decomposing it into simpler pieces. The bubble on the left may be decomposed into a standard double bubble and a single bubble.

PROOF. Let B be a length-minimizing enclosure of areas $A_1, \ldots, A_m$ with connected regions. By Theorem 2.1(2), every vertex of B has degree three. If B had a cut edge, then we could delete this edge to obtain a shorter bubble enclosing the same areas, so B would not be length-minimizing.

Suppose B is disconnected. Let B_1 be a connected subset of B, and let $B_2 = B - B_1$. We can translate B_1 (possibly through distance zero) until some edges of B_1 bump against some edges of B_2, without crossing. At one of the points of contact, we can add a vertex, without disconnecting any regions, to obtain a new bubble B'. The new vertex has degree at least four, so B' is not length-minimizing. However, B' has the same length and encloses the same areas as B, so B is not length-minimizing. □

We easily deduce:

Corollary 3.2 *Any length-minimizing enclosure of three connected areas must have one of the three combinatorial types shown in Figure 3.*

3.2 Decomposition

It is sometimes possible to place lower bounds on the length of a bubble by decomposing it into simpler pieces, as illustrated in Figure 5.

The most basic example of this technique is the following:

Lemma 3.3 *If A and B are length-minimizing bubbles for their combinatorial types and the areas they enclose, then $A \cup B$ is length-minimizing for its combinatorial type and the areas it encloses. Furthermore, if A and B are uniquely length-minimizing, then any length-minimizing bubble 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 with the same combinatorial type as $A \cup B$, enclosing the same areas. C is the union of bubbles 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. The uniqueness part is proved similarly. □

Example 3.4 A bubble consisting of two intersecting circles is uniquely length-minimizing for its combinatorial type and the areas it encloses. (We will need this fact in §7.)

4 Pressure and Equilibrium

In this section we discuss the notion of pressure, which gives us useful information on the curvatures of edges. This allows us to rule out bumping in length-minimizing bubbles of type (a). (See §5 and §6.) In fact, pressure considerations apply not just to length-minimizing bubbles, but to any bubbles in equilibrium.

4.1 Equilibrium

We define a *variation* of a bubble B to be a C^1 family of bubbles $\{B_t \mid t \in (-a, a)\}$ of the same type as B, with $B_0 = B$. More precisely, there must be C^1 functions $\gamma_1, \ldots, \gamma_n : [0,1]\, x(-a,a) \to \mathbb{R}^2$ such that for each t, the edges of B_t are $\gamma_1(\cdot, t), \ldots, \gamma_n(\cdot, t)$. Also, vertices are preserved, i.e., if $s_i, s_j \in \{0, 1\}$ and $\gamma_i(s_i, 0) = \gamma_j(s_j, 0)$, then $\gamma_i(s_i, t) = \gamma_j(s_j, t)$ for all t. (There are separate provisions for the occasional circular edges.)

We sometimes need to consider perturbations of a bubble that can only go in one direction; for example, we may wish to pull bumped up edges away from each other. Thus we define a *half-variation* of a bubble B to be the same as a variation, except that we consider instead a family of bubbles parametrized by $t \in [0, a)$.

A bubble B is in *equilibrium* if for every half-variation $\{B_t\}$ of B in which B_t encloses the same areas as B for each t, we have

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

This implies that for every variation $\{B_t\}$ of B in which B_t encloses the same areas as B for each t,

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

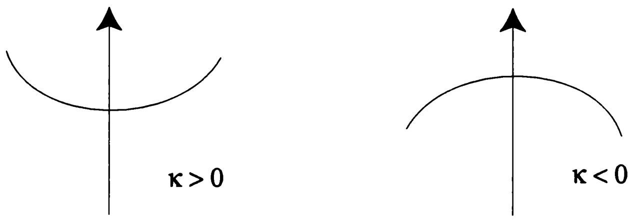

Figure 6: Sign conventions for the oriented curvature of an edge crossed by a path.

For example, the length-minimizing bubbles given by Theorem 2.1 are in equilibrium.

Lemma 4.1 *If a bubble B is in equilibrium, then:*

(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 closed path that only crosses edges transversely and never crosses merges, the sum of the oriented curvatures of the edges crossed is zero.

(We use the sign conventions for oriented curvature shown in Figure 6.)
PROOF. One can prove (1) and (2) with standard variational arguments, as in [M1, Thm. 3.2]. To prove (3), suppose a path φ goes through regions $R_0, R_1, \ldots, R_n = R_0$ and crosses edges $e_1, \ldots, e_n$ with oriented curvatures $\kappa_1, \ldots, \kappa_n$ at points $p_1, \ldots, p_n$, in that order. For $|t|$ small, let B_t be the bubble obtained by adjusting each e_i near p_i so that area t is transferred from R_i to R_{i+1}. (If the path crosses an edge whose curvature is not constant, we only adjust a constant-curvature segment of the edge. If the path crosses several bumping edges, then we adjust all the bumped edges the same way, along some constant-curvature segment containing no merges.)

We have
$$\left.\frac{d\ell(B_t)}{dt}\right|_{t=0} = \sum_{i=1}^{n} \kappa_i.$$
(See, for instance, [M2, p. 6].) Since B is in equilibrium, $\sum \kappa_i = 0$. □

4.2 Pressure

Proposition 4.2 *Let B be a bubble in equilibrium enclosing areas $A_1, \ldots, A_m$. Then there exist constants $p_1, \ldots, p_m$ such that for any variation $\{B_t\}$ of B,*

$$\left.\frac{d\ell(B_t)}{dt}\right|_{t=0} = \sum_{i=1}^{m} p_i \left.\frac{dA_i(t)}{dt}\right|_{t=0}$$

where $A_i(t)$ denotes the area of the i^{th} bounded region of B_t.

PROOF. For each i, choose a variation with initial velocity vector field v_i which changes the area of the i^{th} region at a unit rate while leaving the areas of the other regions fixed. Let p_i be the initial change in perimeter resulting from this variation.

Let $\{B_t\}$ be an arbitrary variation with initial velocity vector field v, and let $c_i = dA_i/dt|_{t=0}$. Let $\{B'_t\}$ be a variation with initial velocity vector field $v - \sum c_i v_i$, so that $\{B'_t\}$ encloses the same areas as B for each t. By the definition of equilibrium,

$$0 = \left.\frac{d\ell(B'_t)}{dt}\right|_{t=0} = \left.\frac{d\ell(B_t)}{dt}\right|_{t=0} - \sum_{i=1}^{m} c_i p_i. \qquad \square$$

We call the constant p_i the *pressure* of the the i^{th} region. Given two regions R_i and R_j, the difference in pressures, $p_i - p_j$, represents the initial rate of change of length of the bubble as area is transferred between the two regions at a unit rate. We define the pressure of the exterior to be zero. We denote the pressure of a region R by $\text{pr}(R)$.

To find $\text{pr}(R)$, draw a path that begins in the exterior, ends in R, only crosses edges transversely, and never crosses a merge. By an argument similar to the proof of Lemma 4.1(3), we find that $\text{pr}(R)$ is the sum of the oriented curvatures of the edges crossed by this path. For example, the curvature of an edge that does not bump against any other edges is equal to the difference between the pressures of the two regions it separates. (The edge bulges outward from the region with higher pressure.) Where n edges bump along a curve, the curvature of the common arc is equal to $1/n$ times the pressure difference between the two outermost regions that the n edges separate.

We have the following nice relation between area, length, and pressure for bubbles in equilibrium. (This will not be needed for the Main Theorem).

Proposition 4.3 *Let B be a bubble in equilibrium whose bounded regions have areas A_i and pressures p_i. Then*

$$\frac{\ell(B)}{2} = \sum p_i A_i.$$

PROOF. Fix $q \in \mathbb{R}^2$, and for $|t| < 1$, let B_t be the bubble obtained by scaling B about center q with ratio $(1+t)$. Then $\ell(B_t) = (1+t)\ell(B)$, and $A_i(t) = (1+t)^2 A_i$. Now apply Proposition 4.2. □

Remark 4.4 The previous two results generalize easily to bubbles comprised of hypersurfaces in $\mathbb{R}^n$. To find the pressure of a region of such a bubble, take the description of pressure for bubbles in $\mathbb{R}^2$ and replace "curvature" with "mean curvature". In $\mathbb{R}^n$, the factor of $1/2$ in Proposition 4.3 is replaced by $(n-1)/n$.

5 Bumping Lemmas

In this section we prove two lemmas which we use in the next section to show that edges in equilibrium bubbles of type (a) do not bump.

Lemma 5.1 *Let B be a bubble in equilibrium. Suppose edges $e_1, \ldots, e_n$ of B bump at p, and let $R_0, \ldots, R_n$ be the regions they separate, in that order. (The "order" of bumping edges makes sense if we view B as a limit of bubbles in which edges do not bump.) Then for $0 \le i \le n$,*

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

PROOF. Let κ_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, we can carry out the subsequent argument near p.) Let

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

Clearly $\kappa_1 \le \kappa_2 \le \ldots \le \kappa_n$, and since $q_n = \mathrm{pr}(R_n)$, it follows that

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

Given $0 \le i \le n$, let φ be a path whose endpoints lie in R_i and R_0. Consider a half-variation of B in which B_t is obtained by first pulling edges $e_1, \ldots, e_i$ away from $e_{i+1}, \ldots, e_n$ so as to transfer area t from R_0 to R_i, and then transferring area t from R_i to R_0 along φ, as in the proof of Lemma 4.1(3). Since B is in equilibrium,

$$0 \le \left.\frac{d\ell(B_t)}{dt}\right|_{t=0} = \sum_{j=1}^{i} \kappa_j + \mathrm{pr}(R_0) - \mathrm{pr}(R_i) = q_i - \mathrm{pr}(R_i),$$

so $\mathrm{pr}(R_i) \le q_i$. □

Remark 5.2 It turns out that conditions (1) through (3) of Lemma 4.1, along with the conclusion of Lemma 5.1, are sufficient conditions for equilibrium. (Assume these four conditions hold. By 4.1(3), it makes sense to define the pressure of a region as the sum of the oriented curvatures of the edges crossed by a path starting in the exterior and ending in the region. By a calculation using Stokes' theorem, in which vertex terms cancel by 4.1(2), one can show that for any half-variation $\{B_t\}$,

$$\left.\frac{d\ell(B_t)}{dt}\right|_{t=0} \geq \sum p_i \left.\frac{dA_i}{dt}\right|_{t=0}$$

which implies equilibrium.)

Lemma 5.3 *Let B be a bubble in equilibrium whose vertices all have degree three. Let γ be a closed curve composed of edges in B; we suppose that γ does not cross but may bump itself. Let U and V be the two regions into which γ divides $\mathbb{R}^2 - \gamma$, in the sense of §2. Suppose 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 γ has no more than three $60°$ exterior angles (where V is the "interior"). Then γ does not bump itself.*

PROOF. Suppose γ bumps itself. We claim that there is a segment γ' of γ 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 γ that merge with V between them. (See Figure 7.) But there must be two such points, because if not, then there are two points of γ which merge with U between them, but do not merge with anything else; however, near this merge, one of the two portions of γ containing the two points must bulge outward from V, contradicting our pressure assumptions.

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

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

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

We claim that $\int_{\gamma'} \kappa \leq 0$. If an even number of segments of γ' bump, then their curvatures cancel in the integral, since their orientations alternate. Suppose that $2k+1$ segments of γ' bump (possibly with other segments from $\gamma - \gamma'$). Let R and S be the two outermost regions separated by the bumping

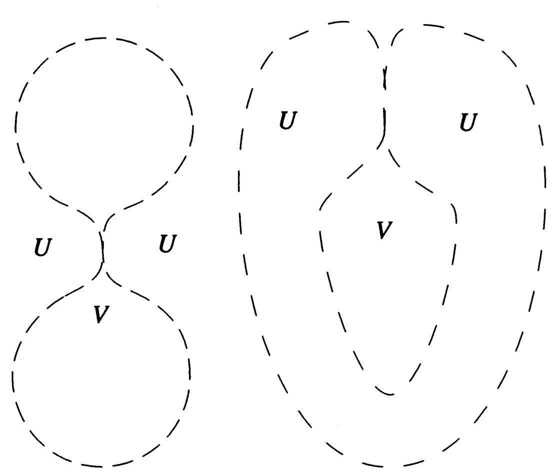

Figure 7: The two possibilities when two points of γ merge with V between them.

edges. One of these regions, say R, is a subset of V', while the other is not. If $S \subset U$, then $\mathrm{pr}(R) \leq \mathrm{pr}(S)$, and $k+1$ of the segments of γ' will have κ 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 bumping edges. By Lemma 5.1, $\mathrm{pr}(R) = \mathrm{pr}(S)$, so 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 $\int_{\gamma'} \kappa \leq 0$, and hence $\sum \theta_i \geq \pi$. It follows that γ' must have at least three 60° exterior angles. We claim that there must in fact be at least four such angles. Suppose there are only three. Then $\sum \theta_i = \pi$, so $\int_{\gamma'} \kappa = 0$. Near one of its endpoints, γ' must bulge outward from V'. This outward bulging segment must be one of an even number of bumping segments of γ', 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 5.1, as before. This is a contradiction. $\square$

Remark 5.4 Similar arguments show that if a bubble encloses four or fewer regions, then each of the enclosed regions has positive pressure. In particular, this implies that the smallest length required to enclose four or fewer prescribed connected areas is a strictly increasing function of the areas. We do not know if this is true for the more general problem in which regions are allowed to be disconnected (except for one or two areas), but if it is true, then it follows that the exterior of a length-minimizing bubble is connected.

6 Bubbles of Type (a)

Proposition 6.1 *Edges do not bump in equlibrium (e.g. length-minimizing) bubbles of type (a).*

PROOF. Let B be an equilibrium bubble of type (a). Let $R_1, \ldots, R_4$ be the regions determined by B; assume that $\mathrm{pr}(R_1) \geq \mathrm{pr}(R_2) \geq \mathrm{pr}(R_3) \geq \mathrm{pr}(R_4)$. 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 5.3, the curves $\partial(R_1)$, $\partial(R_1 + R_2)$, and $\partial(R_1 + R_2 + R_3)$ do not bump themselves. It follows that no edge bumps itself, and the only pairs of edges that can possibly bump are (e_{12}, e_{23}), (e_{12}, e_{24}), (e_{13}, e_{34}), (e_{23}, e_{34}), and (e_{12}, e_{34}). It is enough to show that the first four of these pairs cannot bump, since e_{12} and e_{34} can bump only if e_{23} bumps e_{12} and e_{34}.

Suppose e_{12} and e_{24} bump. 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} be the segment of e_{12} between p and v_{123}. Since e_{24} can only bump e_{12}, e'_{24} does not bump any edges. It follows that e'_{12} bumps something; otherwise e'_{12} and e'_{24} will have constant curvature and hence will merge at p at a $120°$ angle, which is impossible. The only edge that e'_{12} can possibly bump is e_{23}. But the curve consisting of e'_{24}, e''_{12}, e_{13}, and e_{14} separates e'_{12} from e_{23}. (See Figure 8.) This is a contradiction; therefore e_{12} and e_{24} do not bump. A nearly identical argument shows that e_{13} and e_{34} do not bump.

Suppose e_{23} and e_{34} bump. 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, 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 bump e_{34} on the correct side. (See Figure 9.)

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

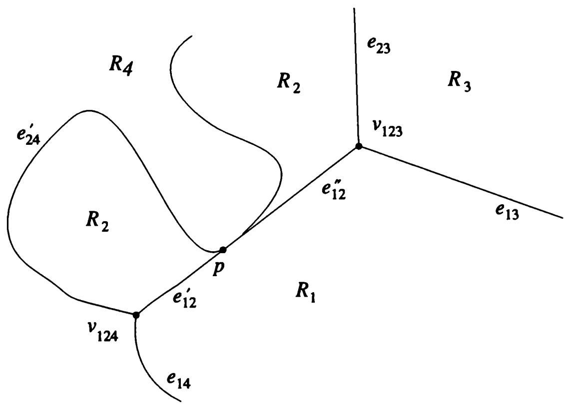

Figure 8: If e_{12} and e_{24} bump, then e'_{12} cannot bump e_{23}.

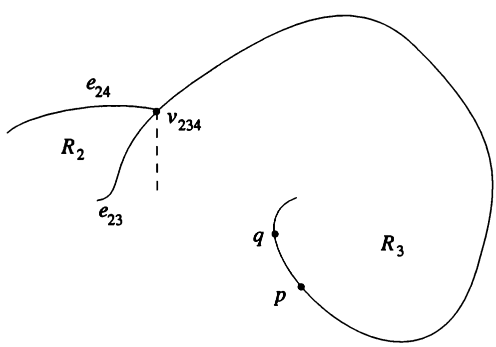

Figure 9: e_{23} and e_{34} cannot bump.

7 Overlapping Bubbles

In this section we define overlapping bubbles and we prove existence and regularity of length-minimizers. In the next section we use overlapping bubbles to find the length-minimizing bubbles of types (b) and (c) and show that these are inferior to type (a).

7.1 Definitions

If φ is a piecewise smooth 1-cycle in $\mathbb{R}^2$, we define the area enclosed by φ, which we denote by $A(\varphi)$, to be

$$A(\varphi) = \int_\varphi \frac{1}{2}(x\,dy - y\,dx).$$

By Stokes' theorem, this definition agrees with the usual definition of area for simple closed curves, up to orientation.

An *overlapping bubble* consists of an embedded graph $G \subset \mathbb{R}^2$ and a C^1 map $f : G \longrightarrow \mathbb{R}^2$. The *type* of (G, f) is the homotopy class of G. If one of the edges of G is mapped by f to a single point, we say that (G, f) is *degenerate* (but still of the same type, unlike the degeneracies of ordinary bubbles defined in §3). Number the bounded faces of G with $1, \ldots, n$. Let φ_i be a cycle which consists of one copy of each of the curves in the boundary of the i^{th} face of G, oriented so that the face is on the left when one goes around the curve. Then the area of the i^{th} region enclosed by (G, f) is defined to be $A(f_* \varphi_i)$. An edge of (G, f) is the push-forward by f of an edge in G. The length of (G, f), which we denote by $\ell((G, f))$, is the sum of the arc-lengths of its edges.

7.2 Existence and Regularity of Length-Minimizing Overlapping Bubbles

First we observe that the classical isoperimetric inequality extends to curves that may cross themselves. A proof of this can be found in, for instance, [O, pp. 1183-4]. We give below a different argument which also works in higher dimensions.

Lemma 7.1 *If γ is a piecewise smooth closed curve in $\mathbb{R}^2$, then $\ell(\gamma)^2 \geq 4\pi A(\gamma)$, with equality if and only if γ is a circle with the right orientation.*

PROOF. We may assume $A > 0$. As in Figure 10, γ can be decomposed into rectifiable boundaries γ_i of regions R_i of area $A_i \in \mathbb{R}$ such that $A = \sum A_i$, $A_i \geq 0$ if $i \geq 1$, and $A_i \leq 0$ if $i \leq 0$. (See [M3, p. 98]; $R_i = M_i$ for $i \geq 1$, and

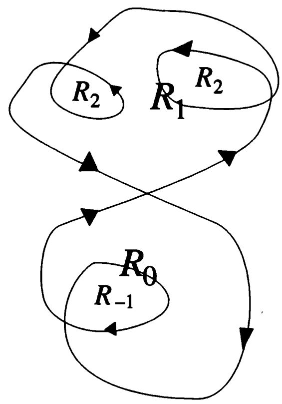

Figure 10: γ can be decomposed into boundaries of regions R_i.

$R_i = -(\mathbb{R}^2 - M_i)$ for $i \leq 0$.) Then

$$\ell(\gamma)^2 = \left(\sum_{i\in\mathbb{Z}} \ell(\gamma_i)\right)^2 \geq \sum_{i\in\mathbb{Z}} \ell(\gamma_i)^2$$

$$\geq \sum_{i\geq 1} \ell(\gamma_i)^2 \geq \sum_{i\geq 1} 4\pi A_i \geq 4\pi A(\gamma),$$

with equality only if γ is a single γ_i ($i \geq 1$) and γ is a circle (with the correct orientation). □

Corollary 7.2 *Given an oriented line segment $\overrightarrow{PQ}$ and a real number r, the unique shortest curve α from Q to P such that $A(\overrightarrow{PQ} + \alpha) = r$ is an arc of a circle or a line segment.*

PROOF. A better competitor, combined with the rest of the circle (possibly crossing it), would contradict Lemma 7.1. □

Proposition 7.3 *Let $G \subset \mathbb{R}^2$ be an embedded graph with bounded faces numbered $1, \ldots, n$, and let $A_1, \ldots, A_n > 0$ be given. Then there exists an overlapping bubble of type G (which may be degenerate) such that the i^{th} region 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 line segments.

(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 number of 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 line segments. The set of all such overlapping bubbles can be parametrized with finitely many variables, and by a standard compactness argument, there is a length-minimizing overlapping bubble B_0 in this set. We claim that B_0 minimizes length among all overlapping bubbles of type G with areas $A_1, \ldots, A_n$. For suppose B is such an overlapping bubble, some of whose edges are not arcs of circles or line segments. By Corollary 7.2, we can replace each such edge with an arc of a circle or a line segment, without affecting the areas of the regions, to get a shorter overlapping bubble B'. Then $\ell(B) > \ell(B') \geq \ell(B_0)$. This proves existence and regularity 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 combinatorial type) connecting the m points at their heads. If not, then there exists a constant $\alpha > 0$ 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 αr. We can restore the areas by adjusting the edges elsewhere; since the original area distortion was at most πr^2, the length increase will be at most βr^2 for some constant β (cf. Proposition 4.2). For sufficiently small r, we have $\alpha r - \beta r^2 > 0$, contradicting minimality. Thus the vectors $v_1, \ldots, v_m$ form a length-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 4.1(3). Suppose $\alpha_1, \ldots, \alpha_m$ are the edges entering a vertex. Let κ_i denote the curvature of α_i, oriented with respect to a clockwise path around the vertex. For every ε, define a new overlapping bubble B_ε by adjusting each α_i such that $A(\alpha_i)$ is increased by ε. This will enclose the same areas as B_0, since the area of a region adjacent to the vertex is given by an expression of the form $A(\alpha_i) - A(\alpha_j) + A(\beta)$. Since B_0 is in equilibrium,

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

Incidentally, this proposition also holds for negative areas, although we will never need to consider such cases.

8 Bubbles of Types (b) and (c)

The following lemma is based on a suggestion by John Sullivan.

Lemma 8.1 *Suppose a length-minimizing overlapping bubble, as given by Proposition 7.3, has a region with only two edges β and γ and two vertices P and Q. Suppose P and Q each have degree 3; let α_1 be the other edge incident to P, and let α_2 be the other edge incident to Q. Then β and γ lie on the same circle or line.*

PROOF. The angles at the vertices are $2\pi/3$, and the sum of the oriented curvatures of the edges at each vertex is zero. These conditions uniquely determine whether or not α_1 and α_2 lie on the same circle or line. But there exists a circle (or line) through P and Q which makes the same angles with β as γ and α_1 and α_2 do. We need only check that this curve has the right curvature. A short calculation shows that a constant-curvature arc through P and Q that makes an angle θ with $\overline{PQ}$ has curvature $2\sin\theta/l(\overline{PQ})$. Thus the curvature is correct, since $\sin(\theta)+\sin(\theta+\frac{2\pi}{3})+\sin(\theta+\frac{4\pi}{3})=0$ for every θ. □

Proposition 8.2 *Let three positive areas be given. Let B be a length-minimizing overlapping bubble of type (b) enclosing the three areas, as given by Proposition 7.3. Then:*

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

(2) If all such B are degenerate, then the only length-minimizer given by Proposition 7.3 is a pair of intersecting circles.

PROOF. (1) Suppose B is nondegenerate. Let $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2$ denote the edges of the B corresponding to the edges marked on the combinatorial type in Figure 11. We wish to show that no two of these edges intersect, except at common endpoints.

We first show that α_1 and α_2 do not intersect. Suppose they do. By Lemma 8.1, α_1 and α_2 are on the same circle or line. There are three ways that two arcs of the same circle or line can intersect; these are cases (a), (b), and (c) of Figure 12. In each case, angle requirements determine the placement of the other four edges, up to labeling and orientation. Consider case (a). Without loss of generality, α_1 is oriented upward, as indicated in the figure. (Otherwise we can rotate the picture 180°.) This determines the labeling and orientation of all the other edges. For instance, the left inner edge must be β_1, not γ_1, because $A(\beta_1-\gamma_1)>0$. But now $A(\alpha_1+\gamma_1+\alpha_2+\gamma_2)<0$, which is a contradiction. Case (b) works similarly.

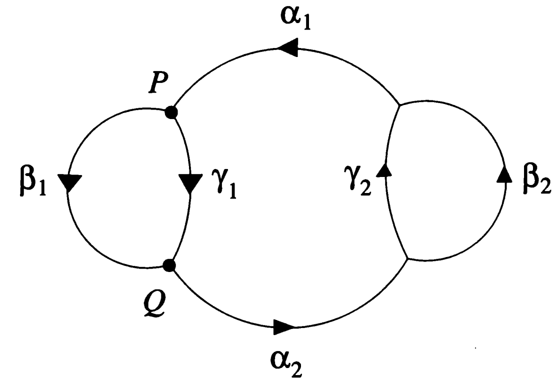

Figure 11: Lemma 8.1 shows that α_1 and α_2 lie on the same circle or line. These edge labelings are used in the proof of Proposition 8.2, where we show that in nondegenerate, length-minimizing, overlapping bubbles of type (b), edges do not overlap.

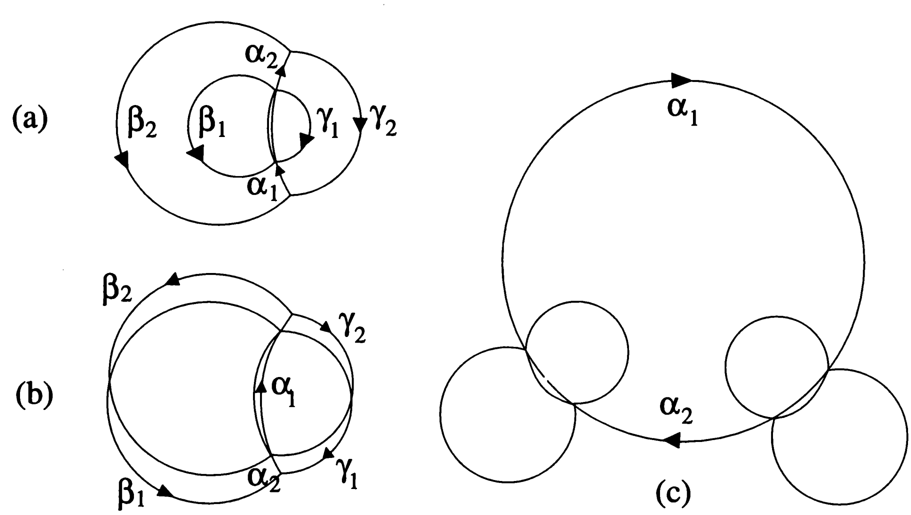

Figure 12: Different ways in which α_1 and α_2 might overlap.

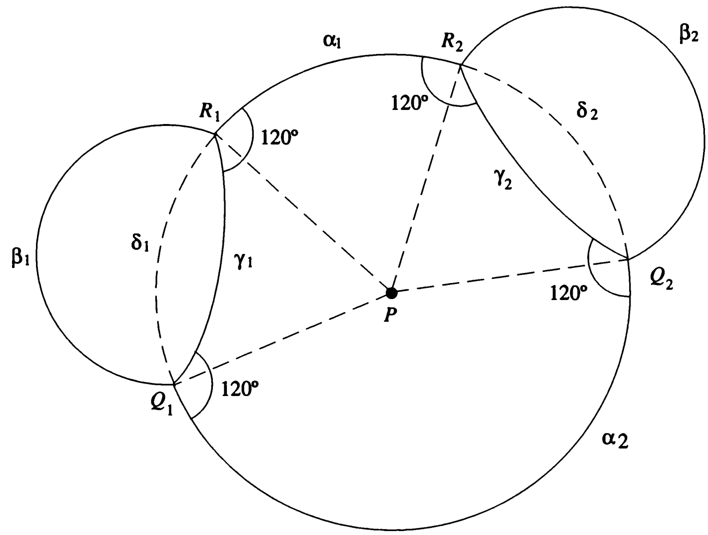

Figure 13: Edges γ_1 and γ_2 of bubble B cannot intersect because they are separated by the two dotted triangles.

For case (c), we can rotate the left-hand region counterclockwise around the circle on which α_1 and α_2 lie, lengthening α_1 and shortening α_2, until α_1 overlaps itself. This process does not change the length of B or the areas enclosed, so the resulting overlapping bubble is still length-minimizing. But by Proposition 7.3(1), α_1 is an *arc* of a circle, which is a contradiction.

It is now simple to show that no other pair of edges intersects away from common endpoints. Let C denote the circle on which α_1 and α_2 lie, and let P be its center. (If α_1 and α_2 lie on a line instead of a circle, then one of the enclosed areas must be negative.) Let Q_i and R_i be the endpoints of β_i and γ_i. Let δ_i be the arc of C connecting Q_i to R_i that does not contain the α_i's. The bubble B is shown in Figure 13. (It is temporarily conceivable that we have labeled the β_i's and γ_i's incorrectly, but this is irrelevant to the argument.) Since two distinct circles or lines can intersect in at most two points, edges β_i and γ_i do not intersect except at their endpoints; also α_i does not intersect β_j or γ_j except at common endpoints. Since β_1 is outside circle C and γ_2 is inside, these two edges cannot intersect; likewise β_2 and γ_1 cannot intersect. To see that γ_1 and γ_2 do not intersect, observe that γ_i is contained in the region bounded by $\overline{PQ_i}, \overline{PR_i}$, and δ_i, and the two such regions are disjoint.

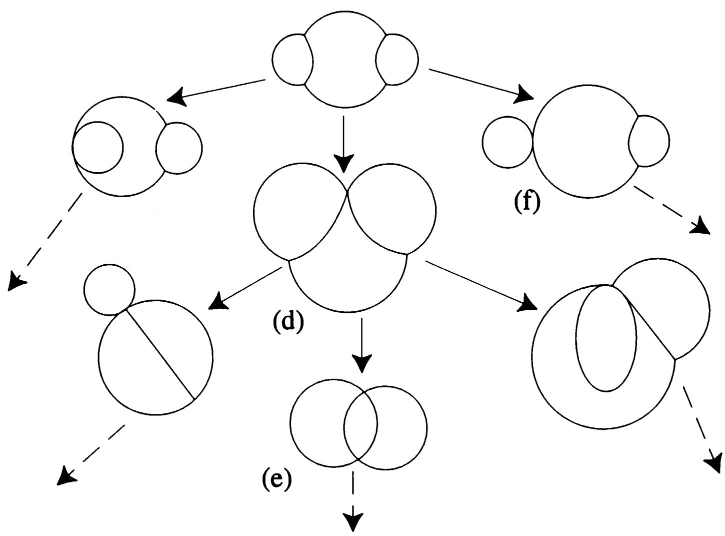

Figure 14: The first few degeneracies of type (b).

To see that β_1 and β_2 do not intersect, invert about circle C and repeat the proof that γ_1 and γ_2 do not intersect.

(2) Suppose all length-mfinimizers are degenerate. If a length-minimizer has the type labeled (d) in Figure 14, one can show that the top vertex is on the same circle as the bottom edge. Hence, by rotating the two small regions downward along this circle and stretching the top vertex into an arc of this circle, we obtain a nondegenerate bubble with the same length and areas. This contradicts the assumption that all length-minimizers are degenerate.

A length-minimizing overlapping bubble of the degenerate type labeled (e) in Figure 14 will be a pair of intersecting circles, by Lemma 3.3. (Lemma 3.3 works also for overlapping bubbles.)

None of the other four degenerates in Figure 14, or any degeneracies thereof, can be a length-minimizer. 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 7.3 that the unique shortest overlapping bubble 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 type labeled (f) in Figure 14, one can split the degree four vertex vertically into two degree three vertices to obtain a shorter bubble (which is still of type (b)) because the top and bottom angles at the degree four vertex are less than 120°. (Indeed, they are 0°.) □

Remark 8.3 The shortest overlapping bubble of type (b) will be a pair of intersecting circles if the two outer areas are large and the inner area is small.

Proposition 8.4 *The conclusions of Proposition 8.2 also hold for type (c).*

PROOF. This is more or less the same as the proof of Proposition 8.2, since types (b) and (c) are different embeddings of the same graph. □

Theorem 8.5 *If B is a length-minimizing enclosure of three positive connected areas, as given by Theorem 2.1, then B must be of type (a).*

PROOF. Suppose B has type (b). By Proposition 8.2, a length-minimizing overlapping bubble of type (b) is always non-overlapping. Hence B is the shortest overlapping bubble of type (b) enclosing the given areas.

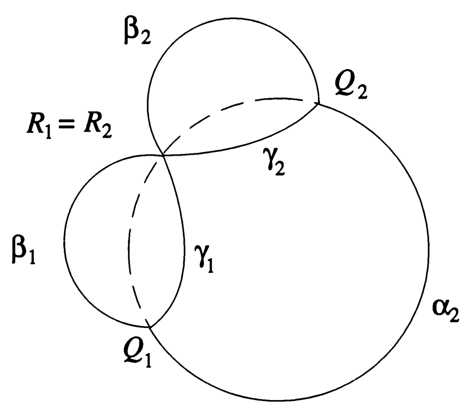

Figure 15: Bubble B' has the same length and encloses the same areas as bubble B of Figure 13, but is not of allowable type; hence B is not length-minimizing.

By Theorem 2.1, B is nondegenerate. Let α_1 and α_2 be as in the proof of Proposition 8.2; since α_1 and α_2 lie on the same circle, we can rotate one of the smaller regions around this circle, lengthening α_2 and shortening α_1, until α_1 shrinks to zero. The resulting bubble B' is shown in Figure 15. B'

has the same length and encloses the same areas as B, and Proposition 8.2 shows that B' is non-overlapping. But B' is not of allowable type; thus B is not length-minimizing.

Type (c) is ruled out similarly, using Proposition 8.4. By Corollary 3.2, B must have type (a). $\square$

Theorems 2.1 and 8.5, along with Proposition 6.1, imply the Main Theorem. The problem of finding the shortest enclosure of three areas in $\mathbb{R}^2$ with possibly disconnected regions remains open.

References

[F1] Joel Foisy, *Soap bubble clusters in* $\mathbb{R}^2$ *and* $\mathbb{R}^3$, undergraduate thesis, Williams College, 1991.

[F2] Joel Foisy, Manuel Alfaro, Jeffrey Brock, Nickelous Hodges, and Jason Zimba, The standard double soap bubble in $\mathbb{R}^2$ uniquely minimizes perimeter, *Pacific J. Math.* **159** (1993), p. 47-59. Featured in the 1994 AMS What's Happening in the Mathematical Sciences.

[M] A. Montesinos Amilibia, *The uniqueness of the triple bubble in the plane*, preprint, 1994.

[M1] F. Morgan, Soap bubbles in $\mathbb{R}^2$ and in surfaces, *Pacific J. Math.*, **165** (1994), p. 141–155.

[M2] F. Morgan, *Riemannian Geometry: a Beginner's Guide*, A.K. Peters, Wellesley, 1992.

[M3] F. Morgan, *Geometric Measure Theory: a Beginner's Guide*, Academic Press, 1988; second edition, 1995.

[O] R. Osserman, The isoperimetric inequality, *Bull. AMS* **84** (1978) p. 1182-1237.