

THE ISOPERIMETRIC PROBLEM ON SOME SINGULAR SURFACES

ANDREW COTTON, DAVID FREEMAN, ANDREI GNEPP, TING NG,
JOHN SPIVACK and CARA YODER

(Received 5 August 2002; revised 20 October 2003)

Communicated by C. D. Hodgson

Abstract

We characterize least-perimeter enclosures of prescribed area on some piecewise smooth manifolds, including certain polyhedra, double spherical caps, and cylindrical cans.

2000 *Mathematics subject classification*: primary 49Q10, 53A10.

1. Introduction

The classical isoperimetric problem seeks the least-perimeter enclosure of a prescribed area on a given surface. We consider this problem on *singular* closed surfaces, such as polyhedra, double discs (two round discs each of constant Gauss curvature, glued together along their boundaries), and a cylindrical can, and characterize all such minimizing curves for a few sample surfaces. The minimizers for the cube are illustrated in Figure 1; the other polyhedra we study are the regular tetrahedron, regular octahedron, and rectangular prism. We also consider the problem on double discs in higher dimensions.

Singularities make this problem interesting. Previous proofs of existence and regularity consider only smooth surfaces. On singular surfaces, not all minimizers look the same. On the cube, for instance, the vertex singularities cause minimizers to fall into distinct families based on how many vertices they enclose.

We classify all simple closed constant-curvature curves on several polyhedra, compute the lengths and areas of these curves, and show that one-component curves are

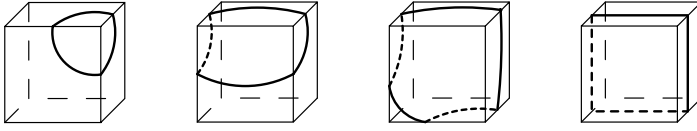


FIGURE 1. There are four types of least-perimeter enclosures on the cube.

best. For double discs and cylinders, we use spherical Schwarz symmetrization to limit complexity and then show that the minimizers have one component by creating illegal singularities in multi-component competitors. For general dimension double discs, we use Schwarz spherical symmetrization combined with the maximum principle to show that minimizers are spherical and not simply constant mean curvature.

Hugh Howards, Michael Hutchings, and Frank Morgan [9] provide a survey of least-perimeter enclosures. Some higher dimensional ambients with conical singularities are treated in [15] and [3].

2. Existence and regularity

We consider piecewise smooth (stratified) n -dimensional closed submanifolds M of \mathbb{R}^N , with a piecewise smooth, continuous Riemannian metric within a bounded factor of the induced metric, possibly undefined on strata of dimension less than $n - 1$. We do not allow M to have cusps. (Technically we require that M be a ‘compact Lipschitz neighbourhood retract’; see [12, 5.5].) Such manifolds include polyhedra and curvilinear polyhedra, cylindrical cans, truncated cones, and pairs of spherical caps attached along their boundaries. Altering the metric on such spherical caps produces flat and hyperbolic double discs.

On such manifolds, we seek regions R of prescribed volume and least perimeter. The boundary ∂R of such a region is called an isoperimetric surface.

PROPOSITION 2.1 (Existence and Regularity). *On a piecewise smooth, n -dimensional closed Riemannian manifold M as above, given $0 < V < \text{vol}(M)$, there exists a least-perimeter region R of volume V . Away from the singularities of M , the isoperimetric surface ∂R is smooth and has constant mean curvature, except for a set of Hausdorff dimension at most $n - 8$. Where the metric is Lipschitz, ∂R is C^1 ($C^{1,1}$ if $n = 2$) except for a set of Hausdorff dimension at most $n - 8$.*

PROOF. For the metric induced from \mathbb{R}^N , standard compactness arguments of geometric measure theory [12, 5.6] produce convergence of a minimizing subsequence R_i to a minimizer R . Since the prescribed metric remains within a factor of the induced metric, the subsequence still converges to R in the prescribed metric. Since

the metric is continuous off a negligible set, area is lower semicontinuous (see [12, 12.5] and [6, Theorem 5.1.5]), and R is perimeter-minimizing. Regularity away from the singularities is a standard result [12, 8.6], even if M is just $C^{1,1}$ [14, Corollaries 3.7 and 3.8]. \square

REMARK. On a 2-dimensional manifold, the constant mean curvature condition implies that at a singular point an isoperimetric curve consists of finitely many arcs (or else it would have infinite length).

PROPOSITION 2.2. *A least-area enclosure of given volume on a piecewise smooth manifold can coincide with a singular hypersurface on an open subset of the singular set only if $H_1 + H_2 \leq 0$, where H_1 and H_2 are the mean curvatures of the singular surface with respect to the normals N_1 and N_2 pointing into the two pieces bounded by the hypersurface.*

PROOF. Suppose that $H_1 + H_2 > 0$ at some point P on the singular hypersurface S , and that the isoperimetric surface coincides with S on some open ball around P . Within this open ball, perturb the surface slightly such that the amount of volume enclosed on each side of the singularity changes by a small amount Δv . This volume will be inside the enclosed region on one side of the singularity and outside the enclosed region on the other, so the total change in volume is zero. Since we are perturbing into each side, the total change in area, to first order in Δv , is $-(H_1 + H_2)\Delta v$. Thus for Δv sufficiently small, we can reduce area while maintaining volume, so the surface coinciding with the singularity is not a minimizer. \square

Note that if $H_1 + H_2 \leq 0$, then the minimizer can coincide with the singularity (see Theorem 5.4).

REMARK. An isoperimetric curve can cross a singular curve Γ infinitely many times. For example, consider the curve Γ on the side of a tall cylindrical can which encircles the can halfway up and locally looks like $e^{-1/x^2} \cdot \sin(1/x)$. Γ divides the can into two pieces, which fit together along this (removable) singular curve to form the cylindrical can. On this can the horizontal circle around the middle of the can, a geodesic isoperimetric curve enclosing half the can's area (see Theorem 6.3), crosses the singular curve infinitely many times.

Regularity on 2-manifolds When we consider only two-dimensional piecewise smooth manifolds, we can prove stronger regularity results.

LEMMA 2.3. *At an isolated singular point p of a piecewise smooth closed surface S , if two pieces of a least-perimeter enclosure meet at an angle of θ (where θ is the smaller of the two possible angles as measured within the manifold), then $\theta \geq 180^\circ$.*

PROOF. By Proposition 2.1, at p the isoperimetric curve consists of finitely many arcs of the same constant curvature. Consider two adjacent arcs, and suppose that the angle θ between these two arcs is interior to the region they enclose. Recall that curvature is the rate of change of length with respect to area. If the two pieces of the curve make an angle of less than 180° at the intersection point p , then we can shave off a small piece of area ΔA near p , resulting in a decrease in length by ΔL_p , such that $\Delta L_p/\Delta A$ is arbitrarily large. Adding a corresponding piece of the same area near a regular point p' with finite curvature has a corresponding ratio

$$\frac{\Delta L_{p'}}{\Delta A} \approx \left. \frac{dL}{dA} \right|_{p'} = \kappa(p') < \frac{\Delta L_p}{\Delta A}.$$

So the total change in length will be $\Delta L_{p'} - \Delta L_p < 0$, and our region cannot be a minimizer.

If the angle θ is exterior to the enclosed region, then we may add a small piece of area at p and remove a corresponding piece at p' to achieve the same result. \square

LEMMA 2.4. *On a piecewise smooth closed surface S , a least-perimeter enclosure of area A has multiplicity at most $\lfloor \phi/2\pi \rfloor$ at an isolated (vertex) singularity of S with total angle ϕ . (We use $\lfloor x \rfloor$ to denote the greatest integer less than or equal to x .)*

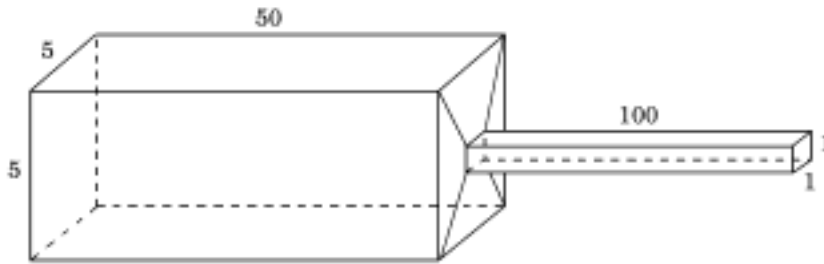


FIGURE 2. (Not to scale.) The least-perimeter enclosure of area 401, a curve of length 4 around the base of the protrusion, passes through vertices of this non-convex polyhedron.

PROOF. Let $p \in S$ have total angle $\phi = \phi(p)$. Suppose that our minimizer C has multiplicity $m > \phi/2\pi$ at p . By Proposition 2.1, m is finite. Since there are $2m$ tangents to C at p , some two adjacent tangents must meet at an angle of less than $\pi = 180^\circ$, contradicting Lemma 2.3. Hence, as claimed, C has multiplicity at most $\lfloor \phi/2\pi \rfloor$ at an isolated singularity of total angle ϕ . \square

COROLLARY 2.5. *On a convex polyhedron, a least-perimeter enclosure does not pass through any vertices.*

PROOF. A vertex of a convex polyhedron is an isolated singularity with total angle $\phi < 2\pi$. \square

For non-convex polyhedra, however, a minimizer can indeed pass through vertices, as in Figure 2. Morgan [13] treats higher-dimensional cases.

Least-perimeter enclosures will typically be one-component curves. To show this, we will use the technical result below.

PROPOSITION 2.6. *Let S be a piecewise smooth closed surface, and for $0 < A < \text{area}(S)$, let \mathcal{U} be the set of all discs $D \subset S$ such that D has constant boundary curvature. Suppose that*

$$\bar{\Omega}(A) \equiv \inf \left\{ \frac{\text{length}(\partial D)^2}{A} : D \in \mathcal{U}, \text{area}(D) = A \right\}$$

is a non-increasing function of A . Then every region of least perimeter is an element of \mathcal{U} .

PROOF. Let us denote by an (i, j) -region one that has at most i components, each of whose boundaries consists of at most j curves. (So a $(1, 1)$ -region is just a disc.) Let $\Omega_{i,j}(A)$ denote the infimum of L^2/A over all constant-curvature curves of length L and area A that enclose a (i, j) -region. Note that $\bar{\Omega}$ as defined above is equal to $\Omega_{(1,1)}$. In addition, let $\Omega(A)$ denote the least value of the isoperimetric ratio L^2/A , taken over all curves of area A and length L .

We begin by eliminating $(k, 1)$ -regions (that is, multiple components, where the boundary of each component consists of a single curve). Suppose that we have a region R consisting of $k > 1$ distinct components, of areas A_1, \dots, A_k and corresponding boundary lengths L_1, \dots, L_k . We compute the value of L^2/A for the whole region R :

$$\frac{\text{length}(R)^2}{\text{area}(R)} = \frac{(L_1 + \dots + L_k)^2}{A_1 + \dots + A_k} > \frac{L_1^2 + \dots + L_k^2}{A_1 + \dots + A_k} = w_1 \frac{L_1^2}{A_1} + \dots + w_k \frac{L_k^2}{A_k},$$

where $w_i = A_i/(A_1 + \dots + A_k)$. Since $w_1 + \dots + w_k = 1$, this is just a weighted average of the L_i^2/A_i , so is at least as big as the smallest value, say L_1^2/A_1 .

Now, one of the components of R is a $(1, 1)$ -region of length L_1 containing area A_1 , so we know that $L_1^2/A_1 \geq \Omega_{1,1}(A_1)$. Using the fact that $\Omega_{1,1}$ is non-increasing, we have

$$(1) \quad \frac{(L_1 + \dots + L_k)^2}{A_1 + \dots + A_k} > \frac{L_1^2}{A_1} \geq \Omega_{1,1}(A_1) \geq \Omega_{1,1}(A_1 + \dots + A_k).$$

Hence there exists some $(1, 1)$ -region M containing area $A_1 + \dots + A_k$ such that

$$\frac{(L_1 + \dots + L_k)^2}{A_1 + \dots + A_k} > \frac{\text{length}(M)^2}{A_1 + \dots + A_k};$$

that is, $\text{length}(R) > \text{length}(M)$.

It follows that $\Omega_{k,1}(A) = \Omega_{1,1}(A)$: allowing $(k, 1)$ -regions does not decrease the smallest possible length. Now, the complement of a $(k, 1)$ -region is a $(1, k)$ -region (that is, one component, with a boundary consisting of at most k curves). Letting $L_{i,j}(A)$ denote the infimum of the length of all (i, j) -regions of area A (that is, $\Omega_{i,j}(A) = L_{i,j}(A)^2/A$), we see that $L_{1,k}(A) = L_{k,1}(\text{area}(S) - A) = L_{1,1}(\text{area}(S) - A) = L_{1,1}(A)$. Thus $\Omega_{1,k}(A) = \Omega_{1,1}(A)$.

The reasoning given above to extend from $\Omega_{1,1}$ to $\Omega_{k,1}$ generalizes exactly to yield that $\Omega_{l,k} = \Omega_{1,k}$. But this is just $\Omega_{1,1}$! It follows that $\Omega = \Omega_{1,1}$. Indeed, clearly $\Omega(A) \leq \Omega_{1,1}(A)$; if $\Omega(A) < \Omega_{1,1}(A)$ then there exists some (l, k) -region R for which $\text{length}(R)^2/\text{area}(R) < \Omega_{1,1}(A)$, contradicting $\Omega_{l,k}(A) = \Omega_{1,1}(A)$.

To see that $\Omega(A)$ is attained by $(1, 1)$ -regions, we note that by existence and regularity, no region that is not of constant curvature can be a minimizer (that is, attain $\Omega(A)$). Furthermore, inequality (1) above is strict, so that a $(k, 1)$ -region with more than one component is strictly worse (that is, has longer perimeter) than some $(1, 1)$ -region of the same area. It follows that a $(1, k)$ -region with more than one boundary curve is also strictly worse than some $(1, 1)$ -region; and a (l, k) -region that is not a $(1, 1)$ -region is strictly worse than some $(1, 1)$ -region. Therefore, $\Omega(A)$ can only be attained by $(1, 1)$ -regions; since it is attained by some region, we know that it is attained by a $(1, 1)$ -region. \square

3. The Gauss-Bonnet theorem

The boundary of a piece of a piecewise smooth closed Riemannian 2-manifold S consists of finitely many smooth curves and finitely many singular points. Each such curve will be called an *edge*, and the singular points on an edge will be called *edge singularities*. The singular points will be called *vertex singularities*.

Since a vertex singularity $v \in S$ has a well-defined angle with respect to each piece of S that v bounds, we may define the *total angle* $\phi(v)$ to be the sum of these angles. For edge singularities and non-singular points $p \in S$, we let $\phi(p) = 2\pi$. We now define the *total contribution to Gaussian curvature* at a point $p \in S$ to be $G(p) = 2\pi - \phi(p)$. If $e \in S$ is an edge singularity bounding two pieces of S for which the edge containing e has curvatures κ_1 and κ_2 , then we define the *linear density of Gaussian curvature* to be $k(e) = \kappa_1(e) + \kappa_2(e)$. With these definitions, the Gauss-Bonnet Formula extends as follows (see for example [10, Chapter 9, Section 1.4]):

PROPOSITION 3.1. *Let S be a two-dimensional manifold with singularities, and let $R \subset S$. Suppose the boundary of R , ∂R , contains no vertices and only isolated edge singularities. Let E and V denote the sets of edge and vertex singularities of S ,*

respectively, and let $K(p)$ be the Gaussian curvature of R at a non-singular point p . Then

$$\int_R K \equiv \int_{R-E-V} K + \int_{R \cap E} k + \int_{R \cap V} G = 2\pi \chi(R) - \int_{\partial R} \kappa,$$

where $\chi(R)$ is the Euler characteristic and κ is the curvature of ∂R .

PROOF. It is well known that a singularity in ∂R where two smooth pieces meet at an angle of α needs to be dealt with by adding a term $\pi - \alpha$; our method of handling vertex and edge singularities is analogous. \square

Note also that one could use the Dirac delta function to write the Gaussian curvature as $K(p) = G(p)\delta^2(p)$ for vertex singularities and $K(e) = k(e)\delta^1(e)$ for edge singularities.

Typically, minimizers will also belong to a smooth family of curves. When this is the case, they are much easier to work with.

LEMMA 3.2. *Suppose that a family of constant-curvature, one-component curves is smoothly parameterized by the radius of curvature $r = 1/\kappa$ and let each curve enclose Gaussian curvature $G(r)$. Then the length L and area A of these curves satisfy $L(r) = (2\pi - G(r))r$ and $A(r) = \pi r^2 - r^2 G(r) + \int G(r)r dr$. If the enclosed Gaussian curvature and length are smooth functions of area, then $L(A)^2/2 = 2\pi A - \int G(A) dA$. If furthermore $G(A) \geq 0$ for all A (for example, if the surface itself has nonnegative Gaussian curvature everywhere), then these curves enclose area more efficiently than flat circles if and only if the curve of smallest area (largest curvature) does.*

PROOF. By the Gauss-Bonnet Formula, $2\pi = \int_R K + \int_{\partial R} \kappa = G(r) + L(r)/r$. Since we are working within a smooth family of curves, $L(r)$ and $A(r)$ are differentiable functions. Recalling that curvature $1/r$ is the rate of change of length with respect to area, we have

$$\begin{aligned} A(r) &= \int \frac{dA}{dL} \frac{dL}{dr} dr = \int \frac{1}{\kappa} (2\pi - G(r) + rG'(r)) dr \\ &= \pi r^2 - r^2 G(r) + \int G(r)r dr. \end{aligned}$$

Also, within this family of curves, $L(A)$ is a differentiable function of A , so the curvature $\kappa = L'(A)$. Integrating the Gauss-Bonnet formula $G(A) + L(A)L'(A) = 2\pi$ with respect to A , we get $\int G(A)dA + L(A)^2/2 = 2\pi A$, as desired.

Finally, note that for flat circles, $L(A)^2 = 4\pi A$. Hence these curves enclose area more efficiently than flat circles do if and only if $\int G(A)dA > 0$. Since

$$\frac{d}{dA} \left(\int G(A)dA \right) = G(A) \geq 0,$$

this inequality holds everywhere if and only if it holds for the smallest value of A achieved by curves in this family. \square

COROLLARY 3.3. *On a polyhedron, the family of curves containing a given set of vertices has lengths and areas satisfying $L(r) = 2cr$ and $A(r) = cr^2 - d$ for some constants c and d .*

PROOF. All of the curves in such a family contain the same Gaussian curvature $G = G(r)$. Hence $L = (2\pi - G)r$, and $A = (\pi - G)r^2 - Gr^2/2 + d$, where d is a constant of integration and $c = \pi - G/2$. Of course, $L^2 = 4c^2r^2 = 4cA + 4cd$. \square

We will see later that the constants have a geometric interpretation: an unfolded constant-curvature curve will form a fraction c/π of a (planar) circle, enclosing a region that is a fraction c/π of a circle with a hole of area $d\pi/c$.

CONJECTURE 3.4. *On a two-dimensional closed convex surface, as enclosed area increases, the Gaussian curvature G enclosed by minimizers is non-decreasing, the curvature of the minimizer is non-increasing, and the isoperimetric ratio L^2/A is non-increasing.*

By Theorems 4.6, 4.8, 4.4, and 5.4, this conjecture is true for regular tetrahedra and octahedra, rectangular prisms, and double discs consisting of two identical caps. See [7, Section 7] for further remarks.

4. Least-perimeter enclosures on polyhedra

In this section, we find all of the least-perimeter enclosures on a few polyhedra. We do this by finding all constant-curvature curves and calculating which is best; by Propositions 2.1 and 2.6 these will be the desired minimizers. The polyhedra we treat are the cube, rectangular prisms, regular tetrahedron, and regular octahedron. We give full proofs in the case of the cube; the other polyhedra are similar and generally easier. For full details, see [7, Section 4].

To find constant-curvature curves, we let $r = 1/\kappa$ denote the radius of curvature. We will unfold a constant-curvature curve into the plane by traveling along the curve and drawing faces as we traverse them (see Figures 4, 7, 9, and 11).

The cube We first consider least-perimeter enclosures on the surface of a cube. We begin by finding all simple closed geodesics.

LEMMA 4.1. *There are three (types of) simple closed geodesics on the surface of the cube. They make angles $\tan^{-1}(1) = 45^\circ$, $\tan^{-1}(2) \approx 63.4^\circ$, and $\tan^{-1}(\infty) = 90^\circ$ with the edges of the cube.*

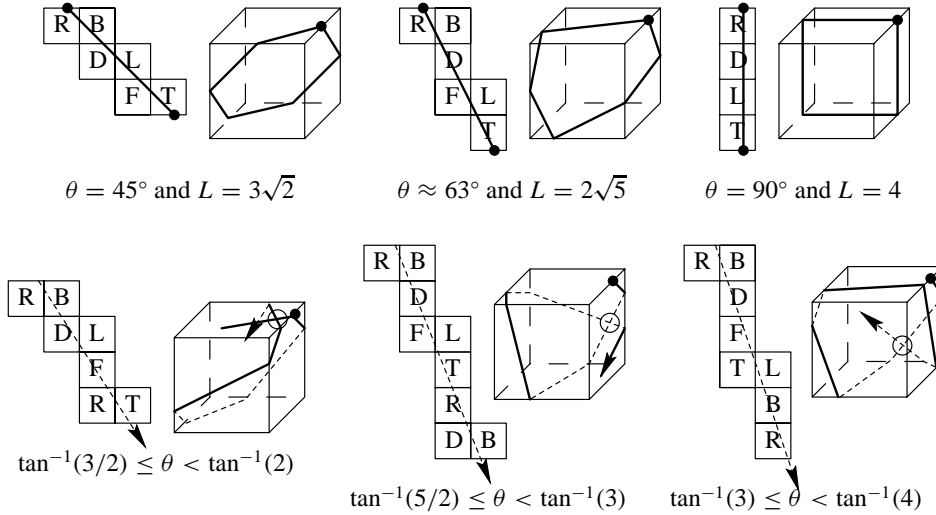


FIGURE 3. The three simple closed geodesics on the cube have angles $\tan^{-1}(1) = 45^\circ$, $\tan^{-1}(2) \approx 63.4^\circ$, and $\tan^{-1}(\infty) = 90^\circ$, and lengths $3\sqrt{2}$, $2\sqrt{5}$, and 4, respectively. By unfolding the cube, we can show that no other angles are possible: any other geodesics must cross themselves at the circled locations (bottom).

PROOF. A geodesic must intersect some edge of the cube; by symmetry it suffices to consider only intersection angles θ for which $45^\circ \leq \theta \leq 90^\circ$. Slide the geodesic parallel to itself until it crosses (say) the Top-Right edge at angle θ very near the Top-Right-Back vertex. After passing through the Right face for a very small distance, the geodesic enters the Bottom face, which we count as the first face it crosses.

- Suppose that $\tan^{-1}((n + 1)/n) \leq \theta < \tan^{-1}(n/(n - 1))$ for some $n \geq 2$. Then the geodesic must cross, in order, the faces B, D, L, F, T, R, at which point the cycle repeats. After crossing $2n$ faces, the geodesic exits this cycle at a Right-Down, Down-Front, or Front-Right edge (if n is 0, 1, or 2 mod 3, respectively), traverses another face (Down, Front, or Right, respectively), and then enters what would have been the $(2n + 1)$ st face in the cycle (Back, Left, or Top, respectively) at right angles to its previous crossings of this face. Hence the geodesic must cross itself on this face, which is impossible. (See Figure 3 for an illustration of the case $n = 2$.)

• Suppose that $\tan^{-1}((2n + 1)/n) \leq \theta < \tan^{-1}((2n - 1)/(n - 1))$ for some $n \geq 2$. Then the geodesic must cross, in order, the faces B, D, F, L, T, R, at which point the cycle repeats. After crossing $3n$ faces, the geodesic exits this cycle at a Front-Top or Right-Down edge (if n is odd or even, respectively), traverses another face (Top or Down, respectively), and then enters what would have been the $(3n + 1)$ st face in the cycle (Left or Back, respectively) at right angles to its previous crossings of this face. Hence the geodesic must cross itself on this face, which is impossible. (See Figure 3 for an illustration of the case $n = 2$.)

• Suppose that $\tan^{-1}(n) \leq \theta < \tan^{-1}(n + 1)$ for $n \geq 3$. Then the geodesic must cross, in order, the faces B, D, F, T, at which point the cycle repeats. After crossing $n + 1$ faces, the geodesic exits this cycle into the Left face, and enters the next face (which would have been the $(n + 2)$ nd face in the cycle) at right angles to its previous crossings of this face. Hence the geodesic must cross itself on this face, which is impossible. (See Figure 3 for an illustration of the case $n = 3$.)

It follows that there do not exist geodesics with angles other than $\tan^{-1}(1)$, $\tan^{-1}(2)$, or $\tan^{-1}(\infty)$. Of course, these three (types of) geodesics clearly do exist. \square

Now we can find all constant-curvature curves on the cube.

LEMMA 4.2. *On the surface of the unit cube, constant curvature curves of the following types exist. For a given curvature, types (2)–(6) are unique up to symmetries of the cube (types (7)–(9) may also be shifted parallel to themselves). Furthermore, these are all of the constant curvature curves on the cube.*

- (1) a circle containing no vertices;
- (2) a circle centered at one vertex;
- (3) a constant-curvature curve about two adjacent vertices;
- (4) a constant-curvature curve about two vertices a distance $\sqrt{2}$ apart;
- (5) a constant-curvature curve about three vertices sharing a face;
- (6) a constant-curvature curve about three vertices, exactly two of which are adjacent;
- (7) a straight line around four sides of the cube, meeting the edges at right angles;
- (8) a straight line around five sides of the cube, meeting the edges at an angle of $\tan^{-1}(2) \approx 63.4^\circ$; and
- (9) a straight line on all six sides of the cube, meeting the edges at 45° angles.

PROOF. We use an unfolding argument to find all constant-curvature curves; see Figure 4.

• If a region contains no vertices, then it is a planar circle, with perimeter $L = 2\pi r$, enclosed area $A = \pi r^2$, and $L^2/\pi = 4A$. There exists such a circle only for $r < 1/\sqrt{2}$, that is, $A < \pi/2$.

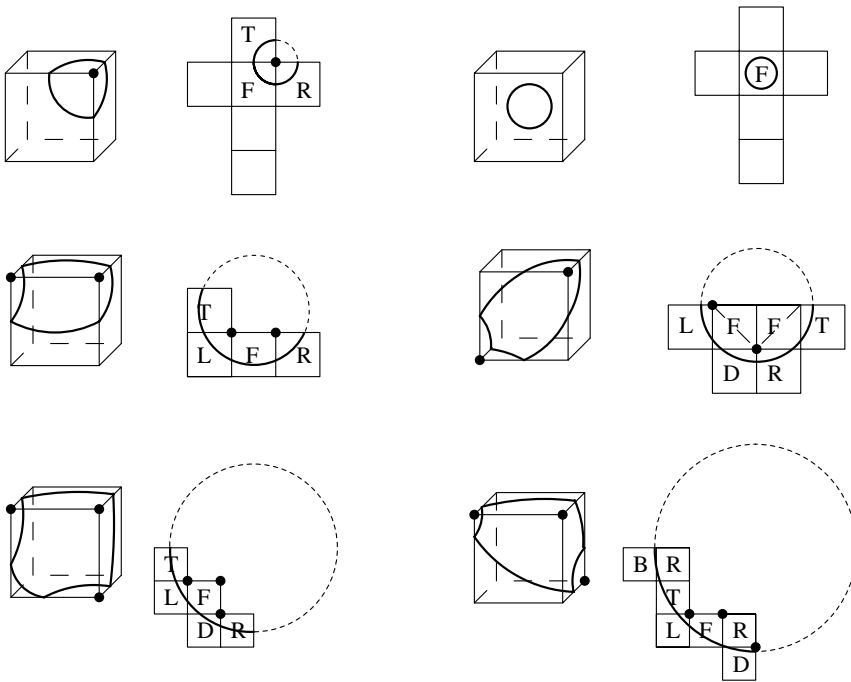


FIGURE 4. All constant-curvature curves on the cube (excluding geodesics) are pictured here along with their unfoldings in the plane. The three on the left can be least-perimeter enclosures. The faces of an unfolded cube are indicated as Front, Back, Top, Down, Right, and Left. Dots indicate enclosed vertices.

- If a region contains one vertex, then it can be unfolded in the plane to curve three-fourths of the way about the origin. Since the folded curve meets itself smoothly on the cube, rotating the unfolded curve 270° about the origin will extend it smoothly. Four copies of the unfolded curve form a closed, constant-curvature curve that wraps about the origin three times and has 270° rotational symmetry. Hence this curve must be a single circle (with multiplicity three), centered at the origin. Our original region on the cube is therefore a circle centered at the vertex, with $L = 3\pi r/2$, $A = 3\pi r^2/4$, and $L^2/\pi = 3A$. Such circles exist only for $r < 1$, that is, $A < 3\pi/4$.

- If a constant-curvature curve contains two vertices, there are three possibilities: the vertices can be adjacent, diagonally opposite along a face, or antipodal. If the two enclosed vertices share an edge, then when unfolded, the curve forms half of a circle, and encloses a region that is half of a circle with a one-face square hole (of area 1). It follows that the region has length $L = \pi r$, area $A = (\pi r^2 - 1)/2$, and $L^2/\pi = 2A + 1$. A curve about two adjacent vertices exists for $\sqrt{2}/2 < r < \sqrt{5}/2$, that is, for $\pi/4 - 1/2 < A < 5\pi/4 - 1/2$ ($0.29 \lesssim A \lesssim 3.43$).

If the two enclosed vertices are a distance of $\sqrt{2}$ apart, then when unfolded, the curve

encloses a region that is half of a circle with a two-face diamond-shaped hole (of area 2). So for these curves, $L = \pi r$ and $A = (\pi r^2 - 2)/2$, which is always worse than a curve around two adjacent vertices.

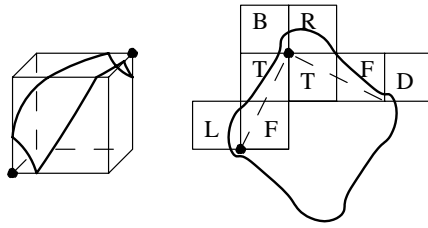


FIGURE 5. Two antipodal vertices cannot be enclosed by a constant-curvature curve.

The third case cannot occur; the unfolded curve would have to enclose a region that would be half of a circle with a diamond-shaped hole of area 5 (see Figure 5). But then to stay on the faces indicated, the radius would have to satisfy $\sqrt{10}/2 < r < \sqrt{10}/2$, since the L-T-F vertex needs to be outside the circle while the T-B-R vertex is inside it.

- If a region is to contain three vertices, there are again three possibilities: the vertices share a face; two vertices are adjacent to each other but not to the third; or no two vertices are adjacent. If the three vertices share a face, then when unfolded, the curve encloses a region that is a quarter of a circle with a plus-shaped hole (of area 12). Hence $L = \pi r/2$, $A = (\pi r^2 - 12)/4$, and $L^2/\pi = A + 3$. These curves exist for $\sqrt{5} < r < \sqrt{8}$ (since the F-D-R vertex must be inside the region and the L-F-D vertex outside it), or for $5\pi/4 - 3 < A < 2\pi - 3$ ($0.93 \lesssim A \lesssim 3.28$).

If only two of the vertices share an edge, then the unfolded curve encloses a region that is a quarter of a circle with a square hole of area 16. Hence $L = \pi r/2$ is the same as for three co-facial vertices, but $A = (\pi r^2 - 16)/4$ is less, and these curves do worse than ones containing three co-facial vertices.

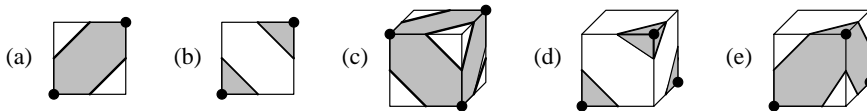


FIGURE 6. If a constant-curvature curve could enclose three vertices, no two adjacent, then a given face would look like either (a) or (b). Neither (c) nor (d) can happen with a single curve, so we just need to rule out (e). The drawings are schematic only, with curves replaced by straight lines.

The third case cannot occur. To see this, consider what happens on each of the three faces containing two enclosed vertices. The curve must have two components on such a face, since it intersects each of the four edges. The region can enclose either one (Figure 6 (a)) or two components (Figure 6 (b)) on the face. If the region has one

component on all three faces, its boundary would consist of two curves (Figure 6 (c)), and if the region has two components on two faces, then the region has (at least) two components (Figure 6 (d)). Since (for the moment) we are only considering single constant-curvature curves, we can assume that (say) the Front and Right faces have one component and the Down face has two (Figure 6 (e)). But now consider the Front-Right edge: the curve proceeds from this edge down to the Front-Down and Right-Down edges, so cannot be convex inwards to the region—a contradiction.

- If a region is to contain four vertices, it must be a (simple closed) geodesic, by the Gauss-Bonnet Formula (Proposition 3.1). By Lemma 4.1, there are only geodesics making angles of 90° , 45° , or $\tan^{-1}(2)$ with the edges of the cube. These have lengths $L = 4$, $L = 3\sqrt{2}$, and $L = 2\sqrt{5}$, respectively (see Figure 3, top).

- If a region contains $v > 4$ vertices, its complement must contain $8 - v < 4$ vertices. So we get no new constant-curvature curves.

Hence we have an exhaustive list of all possible combinations of vertices that can be enclosed by a constant-curvature curve, so we have explicitly found all constant-curvature curves on the cube. They are unique as described in the statement of Lemma 4.2 (that is, there is only one way to enclose a given set of vertices with a curve of given constant curvature) because they must meet themselves smoothly, as described above. \square

Having found all constant-curvature curves, we can now find all least-perimeter enclosures.

THEOREM 4.3. *On the surface of the unit cube, the following are all of the least-perimeter enclosures of given area A :*

- curves of type (2), for $0 < A \leq 1$;
- curves of type (3), for $1 \leq A \leq 2$;
- curves of type (5), for $2 \leq A \leq 16/\pi - 3$ (≈ 2.09); and
- curves of type (7), for $16/\pi - 3 \leq A \leq 3$.

Of course, for $A \geq 3$, the least-perimeter enclosure of an area of A is the complement of the least-perimeter enclosure for area $6 - A$.

PROOF. It follows from Propositions 2.1 and 2.5 that least-perimeter enclosures exist, have constant curvature, and do not pass through any vertices. Lemma 4.2 gives us a complete classification of constant-curvature curves, so we just need to determine, for given area A , which of types (1) – (9) has the least perimeter.

It is clear that, for given area that can be enclosed by both of the corresponding types, a curve of type (2) beats one of type (1), (3) beats (4), (5) beats (6), and (7) beats (8) and (9). (For instance, type (2) has $L^2 = 3\pi A$ while type (1) has $L^2 = 4\pi A$ so the former is better as long as it exists, that is, for $0 < A < 3\pi/4$.)

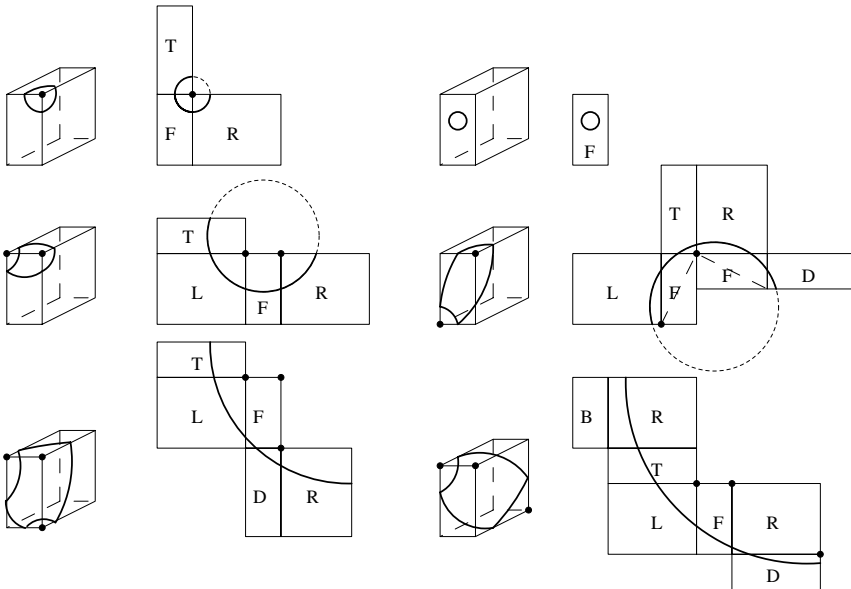


FIGURE 7. All constant-curvature curves on a rectangular prism (excluding geodesics) are pictured here along with their unfoldings in the plane. The three on the left can be least-perimeter enclosures. The faces of an unfolded prism are indicated as Front, Back, Top, Down, Right, and Left. Dots indicate enclosed vertices.

Now, for $3A < 2A + 1$ (that is, $A < 1$), type (2) beats (3); for $A < 2$, type (3) beats (5); and for $A < 16/\pi - 3$, type (5) beats (7). Since $A \leq 1$ are areas attainable by type (2), we see that type (2) will be a least-perimeter enclosure (among one-component constant-curvature curves) for areas $A \leq 1$. Similarly, we can identify which types do best between the other transition points.

Hence we know which constant-curvature curves do best. It is easy to check that L^2/A for these best constant-curvature discs is decreasing, and therefore by Proposition 2.6 they are indeed the least-perimeter enclosures. \square

The rectangular prism We next consider a rectangular prism with sides a , b , and c . This is the only non-regular polyhedron we study.

THEOREM 4.4. *Let $m = (4 - \pi)/(2\pi - 4)$ and $K = m + \sqrt{m^2 + m} \approx 1.10$. On the surface of a rectangular prism with sides $c \leq b \leq a$, the following are all of the least-perimeter enclosures of given area A :*

- a circle about one vertex, for $0 < A \leq c^2$;
- a constant-curvature curve about 2 adjacent vertices a distance of c apart, for $c^2 \leq A \leq b^2 + bc$;

- a constant-curvature curve about 3 vertices sharing a face with edges of length b and c , for $b^2 + bc \leq A \leq (4/\pi - 1)(b^2 + c^2 + 2bc) + bc$ if $b/c \leq K$;
- a geodesic meeting the edges of length a at right angles, for $(4/\pi - 1)(b^2 + c^2 + 2bc) + bc \leq A \leq ab + bc + ac$ if $b/c \leq K$, or for $b^2 + bc \leq A \leq ab + bc + ac$ if $b/c \geq K$.

Of course, for $A \geq ab + bc + ac$, the least-perimeter enclosure for area A is the complement of the least-perimeter enclosure for area $2(ab + bc + ac) - A$.

PROOF. As with the cube, one finds all constant-curvature curves (shown in Figure 7) and selects the best. For further details, see [7, Theorem 4.8]. The condition on the ratio b/c occurs because the transition points between types of curves depend on the shape of the prism; comparing these transition points shows that a curve enclosing three vertices can be most efficient if and only if the smallest face is nearly square. \square

The regular tetrahedron One feature of the regular tetrahedron not encountered in other polyhedra is that there are infinitely many types of simple closed geodesics, a fact mostly irrelevant to the present investigation but of independent interest. (For our purposes it suffices to show that the shortest simple closed geodesic has length 2; this can be done without the following lemma.)

LEMMA 4.5. *Consider a tiling of the plane with equilateral triangles that have one side parallel to the x -axis. Suppose a straight line connecting two vertices of this tiling makes an angle θ with the x -axis. Then there exist geodesics on the regular tetrahedron that make an angle of θ with some edges of the tetrahedron. Furthermore, these are simple closed geodesics, and all simple closed geodesics can be obtained in this manner.*

The angle $\theta = 60^\circ$ is the most obvious simple closed geodesic, and was all that was found in [2], although, as Heppes later pointed out [8], there is an infinitude of geodesics. [5] also found the next most obvious $\theta = 90^\circ$ geodesic, and it was their Figure 4 that led to our discovery of all of the simple closed geodesics.

PROOF. We begin by unfolding our surface as in Figure 8; the nice feature of the tetrahedron is that it unfolds to a regular tiling of the plane. A simple closed geodesic will unfold to become a straight line in the plane. Translate this line parallel to itself until it passes through a vertex; since our unfolding is a lattice, it will hit other vertices every time the lattice repeats. Conversely, given a line segment connecting two occurrences of the same vertex in our lattice, we can translate the segment parallel to itself by a small amount so that it no longer goes through a vertex, and then the line becomes a simple closed geodesic on the folded-up tetrahedron. (The geodesic does not intersect itself because any two occurrences of a given face of the tetrahedron in

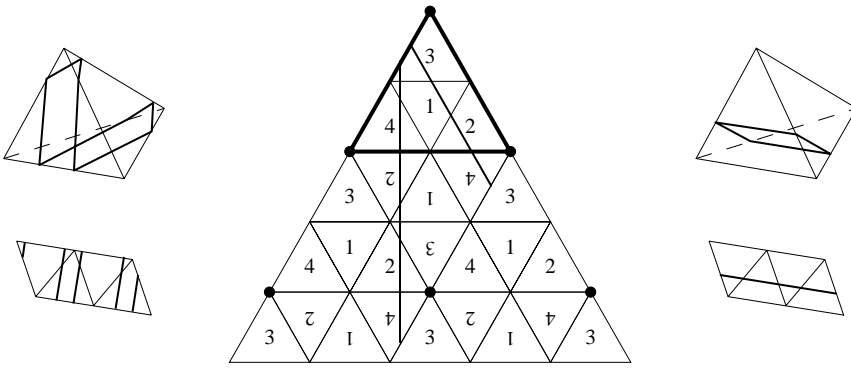


FIGURE 8. To find geodesics on the regular tetrahedron, we unfold it to a tiling of the plane. The fundamental domain—one copy of the tetrahedron—is outlined. Simple closed geodesics must be (translates of) line segments connecting two occurrences of the same vertex, here indicated by the large dots. The 60° and 90° geodesics (of lengths 2 and $2\sqrt{3}$, respectively) are shown in the tiling, on the tetrahedron, and on a single unfolded tetrahedron.

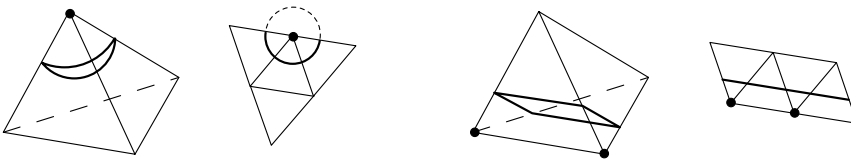


FIGURE 9. The least-perimeter enclosures on the regular tetrahedron are circles centered at a vertex and simple closed geodesics. Enclosed vertices are indicated by dots.

the lattice are either translates of each other, or 180° rotations, and thus all segments of the closed geodesic on that face are parallel to each other.)

Hence it suffices to find all lines connecting occurrences of the same vertex in our lattice. That is, the possible angles of simple closed geodesics are the possible angles of lines between two points of the lattice, as claimed. \square

THEOREM 4.6. *On the surface of the regular tetrahedron with edge 1, the following are all of the least-perimeter enclosures of given area A (see Figure 9):*

- a circle about a single vertex, for $0 < A \leq 2/\pi$ (≈ 0.637); and
- a simple closed geodesic making an angle of 60° with four edges of the tetrahedron, for $2/\pi \leq A \leq \sqrt{3}/2$ (≈ 0.866).

Of course, for $A \geq \sqrt{3}/2$, the least-perimeter enclosure for area A is the complement of the least-perimeter enclosure for area $\sqrt{3} - A$.

PROOF. Again, one finds all one-component constant-curvature curves and selects the best. For further details, see [7, Theorem 4.2]. \square

The regular octahedron Finally, we consider the regular octahedron. The simple closed geodesics are as follows:

LEMMA 4.7. *There are two (types of) simple closed geodesics on the surface of the regular octahedron. They make angles $\tan^{-1}(\sqrt{3}) = 60^\circ$ and $\tan^{-1}(\infty) = 90^\circ$ with some edges of the octahedron.*

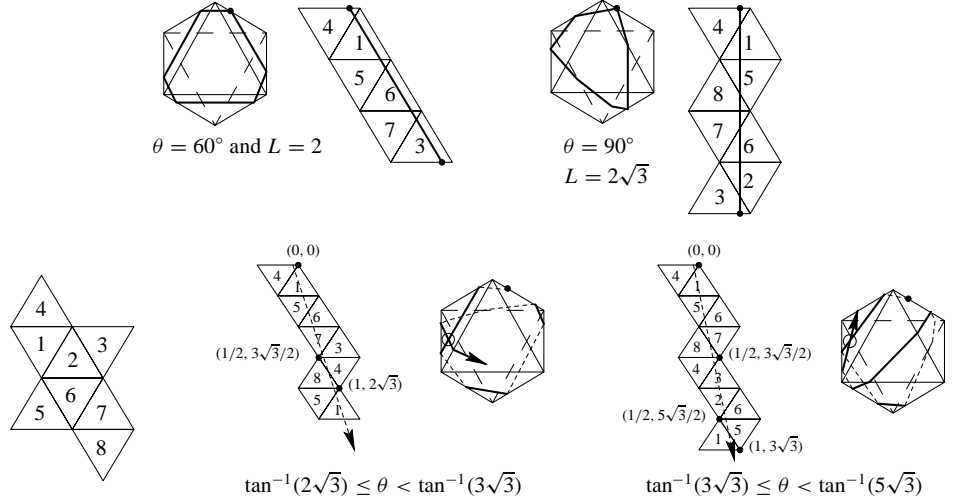


FIGURE 10. The two simple closed geodesics on the octahedron have angles $\tan^{-1}(\sqrt{3}) = 60^\circ$ and $\tan^{-1}(\infty) = 90^\circ$; and lengths 3 and $2\sqrt{3}$, respectively (top). By unfolding the octahedron (bottom left), we can show that no other angles are possible: any other geodesics must cross themselves at the circled locations (bottom).

PROOF. The proof uses an unfolding argument similar to that used for the cube (see Figure 10) to show that the geodesic can only intersect an edge at 60° or 90° . For further details, see [7, Lemma 4.3]. □

THEOREM 4.8. *On the surface of a regular octahedron with edge 1, the following are all of the least-perimeter enclosures of given area A (see Figure 11).*

- a circle about a single vertex, for $0 < A \leq \sqrt{3}/2$ (≈ 0.866);
- a constant-curvature curve about two adjacent vertices, for $\sqrt{3}/2 \leq A \leq 27/4\pi - \sqrt{3}/2$ (≈ 1.28); and
- a simple closed geodesic making an angle of 60° with six edges of the octahedron, for $27/4\pi - \sqrt{3}/2 \leq A \leq \sqrt{3}$ (≈ 1.73).

Of course, for $A \geq \sqrt{3}$, the least-perimeter enclosure for area A is the complement of the least-perimeter enclosure for area $2\sqrt{3} - A$.

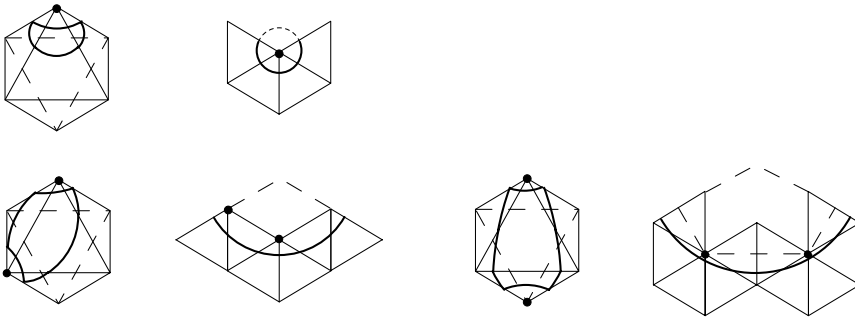


FIGURE 11. The least-perimeter enclosures on the regular octahedron are (1) circles centered at a vertex, (2) constant-curvature curves about two adjacent vertices, and (3) simple closed geodesics containing a face (Figure 10, upper left). Also shown, on the right, is a (non-minimizing) constant-curvature curve about two antipodal vertices. Enclosed vertices are indicated by dots.

PROOF. Again, one finds all one-component constant-curvature curves and selects the best. For further details, see [7, Theorem 4.4]. \square

5. Double discs

We now move away from polyhedra, which have only isolated vertex singularities, and consider manifolds with edge singularities. (The edges of polyhedra are not intrinsic singularities, as one can see by unfolding the polyhedron.) We first study manifolds that consist of two round, constant-curvature n -dimensional discs glued together along their spherical boundaries; see Figure 12 for some examples. These manifolds have one singular surface and are symmetric under rotation about a particular axis. If the two halves of the manifold are identical, then there is additional reflectional symmetry. Although we will usually draw these manifolds using spherical caps, we also allow the sectional curvature of one or both caps to be negative, in which case part or all of the manifold is hyperbolic. These manifolds are $C^{1,1}$ Riemannian manifolds everywhere (see Proposition 2.1), but we will not use this fact in proving our results.

PROPOSITION 5.1. *Consider an n -dimensional manifold M constructed by taking two round balls, each of constant sectional curvature, and gluing them together along their boundaries. (Of course, to do this their boundaries must be congruent.) On this manifold, a least-area enclosure of given volume is either a round spherical cap on each side of the singularity or a round sphere on one side (possibly the singular set).*

PROOF. We use spherical Schwarz symmetrization (see Figure 13 and [4, Chapter 2,

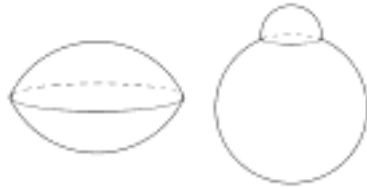


FIGURE 12. Examples of two-dimensional double discs. *Left*: convex double disc consisting of two identical caps. *Right*: non-convex double disc consisting of two non-identical caps.

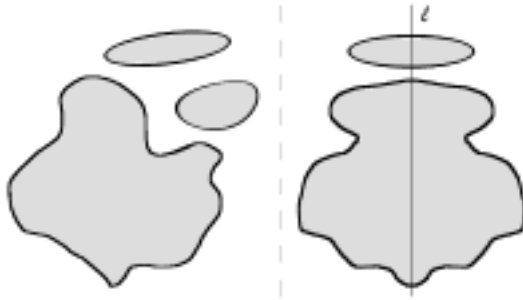


FIGURE 13. Schwarz symmetrization of a region about a line ℓ . Volume is maintained and boundary area does not increase.

Section 9.2]). The details of this argument, in the most general context of geometric measure theory in \mathbb{R}^n , can be found in [1]. Generally, we work modulo measure zero.

By Proposition 2.1, we know that a minimizer exists and has constant mean curvature away from the singularity, except possibly for a singular set of dimension at most $n - 8$. Let R denote a region of least area. Draw a geodesic ℓ connecting the poles of the two balls. For each $(n - 1)$ -sphere S centered at the pole of a given face, determine the area of the slice $R \cap S$. To obtain a more symmetric region R' , replace $R \cap S$ with a round $(n - 1)$ -ball on S enclosing the same area and centered on the geodesic ℓ . After doing this, $\text{vol}(R') = \text{vol}(R)$, and $\text{area}(\partial R') \leq \text{area}(\partial R)$, with strict inequality unless every slice of R is a round spherical cap centered on a geodesic (which we may assume is ℓ). We conclude that every slice of R is a round spherical cap; if the intersection with ∂R is empty, the slice is either the whole sphere or the empty set.

Consider the points of ∂R closest to the pole of one ball. Suppose there is more than one point. There must be a spherical cap of such points, and by analytic continuation, a whole sphere of such points. There can be no other components of ∂R on this ball, or we could slide this sphere until it touches another component, creating an illegal singularity.

On the other hand, suppose that there is a unique point p closest to the pole. Take a slice of the minimizer by an $(n - 1)$ -sphere centered at the pole, with radius slightly

larger than the distance from p to the pole. This sphere separates ∂R , and we consider the small component S inside this sphere. We choose spheres of mean curvature equal to that of the minimizer away from the singularity, position one above and one below S , and move them in toward the component until one touches ∂R on the interior at a point q . By a standard argument of geometric measure theory, q must be a regular point of ∂R . (If the tangent cone lies in a half-space, it must be a hyperplane and hence the point is a regular point.) By the maximum principle, this component of ∂R is locally and hence globally round on this ball; that is, a round sphere, a spherical cap, or the empty set. In any of the three cases, it must be the only component of ∂R on the ball.

We now know that the minimizer consists of two spherical caps or two round spheres, one on each side of the singularity, or a round sphere on one side. If it consists of two round spheres, we can slide the two spheres until they touch at the singularity, and then a simple variational argument shows that this cannot be a minimizer. We conclude that the minimizer is either a round spherical cap on each side of the singularity or a round sphere on one side. \square

In general, if the minimizer crosses the singularity we cannot determine the proportion of the area enclosed on each side. However, on manifolds consisting of two identical discs we can describe the minimizer precisely.

LEMMA 5.2. Construct an n -dimensional manifold M by taking two identical round discs of constant sectional curvature and gluing them together along their boundaries. If M is not a round sphere, then for all $0 < V < \text{vol}(M)$ there is a unique (up to rotations about the axis of symmetry) C^1 surface which encloses volume V and consists of a round spherical cap on each side of the singularity. It consists of two identical spherical caps meeting the singular curve orthogonally.

PROOF. We first consider the two-dimensional case. Without loss of generality, we can scale M so that its sectional curvature is 0, 1, or -1 . Consider two circular arcs of the same curvature, one on each face of M , meeting up so that they subtend the same edge length.

On each disc, draw a geodesic connecting the two points on the edge where the arcs meet. (Note that if M consists of two spherical caps greater than hemispheres, this geodesic will only exist on the discs completed to a sphere; this does not affect our argument.) Let α be the angle at which this geodesic meets the edge (negative if the discs are more than half spheres). Let the angles between these geodesics and the constant curvature arcs be β and γ on respective discs. Thus the angle at which the arc crosses is $\alpha + \beta$ on one side and $\alpha + \gamma$ on the other. (See Figure 14.)

By consideration of a sector of a circle of curvature κ , where the angle between the

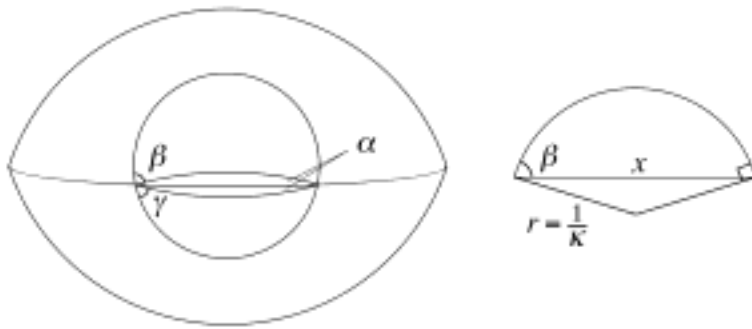


FIGURE 14. Calculation of a minimizer crossing the singularity of a double disc made up of two identical caps.

arc and the chord is β , we find that

$$\sin \beta = \cos \left(\frac{\pi}{2} - \beta \right) = \frac{f(x/2)}{f(r)},$$

where $f(t)$ is t if the sectional curvature of M is zero, $\tan t$ if it is 1, and $\tanh t$ if it is -1 , and where x is the length of the geodesic chord and r is the radius of the circle. (These formulae are standard in non-Euclidean geometry and can be found in any introductory text on the subject.) Since the lengths of the geodesic chords are the same on both sides, we consider the same wedge on the other side and find that $\sin \beta = \sin \gamma$.

Since both β and γ are in $[0, \pi]$, the only solutions to $\sin \beta = \sin \gamma$ are $\beta = \gamma$ and $\beta = \pi - \gamma$. The fact that these arcs must cross the edge in a C^1 fashion tells us that $2\alpha + \beta + \gamma = \pi$. Thus if $\beta = \pi - \gamma$, then $\alpha = 0$, and either M is a sphere, in which case the geodesic chord makes zero angle with the singularity, or the edge length is zero, in which case the curve lies entirely on one side (it is a circle tangent to the edge).

If $\beta = \gamma$, then we have $\alpha + \beta = \alpha + \gamma = \pi/2$, and thus the curve crosses the edge orthogonally and is the same on both sides. Since there are no other solutions for $\beta, \gamma \in [0, \pi]$, this curve is the unique one that has constant curvature and crosses the edge in a C^1 manner.

Uniqueness for a given area follows because if we have two curves of different curvatures meeting the edge orthogonally, it is clear that we can place them tangent to each other at the edge such that the curve of higher curvature is entirely contained in the one of lower curvature. Thus area enclosed is a strictly decreasing function of curvature, and thus the curve we have constructed is the unique one for a given area.

For $n > 2$, the fact that the surface must be a spherical cap on each side of the singularity means we need only consider a two-dimensional cross-section that

intersects the axis of symmetry. Since the surface is C^1 , the argument above applies, and we conclude that each cross-section consists of identical circular arcs meeting the edge orthogonally. Thus the surface consists of identical spherical caps meeting the edge orthogonally. \square

LEMMA 5.3. *Construct an n -dimensional manifold M by taking two identical round discs of constant sectional curvature and gluing them together along their boundaries. On this manifold, the symmetric surface crossing the singularity orthogonally encloses volume more efficiently than a sphere on one of the caps if and only if the mean curvature of the edge singularity as viewed from each face is greater than zero.*

PROOF. If curvature of the edge is positive, we consider a sphere entirely on one disc. By Figure 15, it is clear that if we cut this sphere in half and place each half against the singularity, we can enclose more volume with the same area. It is then easy to reduce area and restore the original volume.

If the curvature of the edge is negative, we consider the surface consisting of two identical spherical caps meeting the edge orthogonally. By Figure 15, it is clear that if we place the two caps together on the same disc, we can enclose more volume with the same area. It is then easy to reduce area and restore the original volume. If the two halves do not fit on the same cap, then they must each have area greater than half that of an equatorial sphere on one of the discs, since each disc is greater than a hemisphere. Since it is possible to enclose any volume with an area equal to or smaller than that of an equatorial sphere on one disc, such a surface crossing the singularity cannot be a minimizer.

If the curvature of the edge is zero, the surface is a round sphere, and thus the two candidates are identical and equally efficient. \square

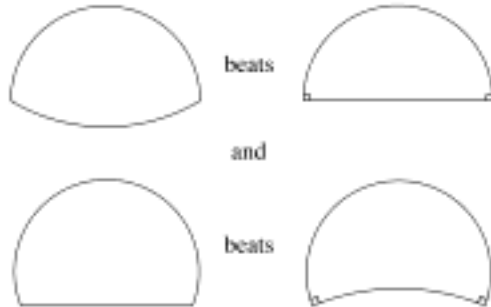


FIGURE 15. *Top*: If the curvature of the edge singularity is positive, it is more efficient to enclose area against the edge than with a circle on one cap. *Bottom*: If the curvature of the edge singularity is negative, it is less efficient to enclose area against the edge.

REMARK. Alternatively, one could prove Lemma 5.3 for the case $n = 2$ using the extension of the Gauss-Bonnet formula from Proposition 3.1. If the curvature of the edge singularity is positive, then enclosed Gaussian curvature is always greater for the family of curves that cross the edge than for circles which lie only on one face, so length is less for the curves crossing the edge. Similarly, if the curvature of the edge singularity is negative, then enclosed Gaussian curvature is less for the family of curves that cross the edge than for the curves which lie only one one face, and so length is greater for the curves crossing the edge.

We can now prove our double disc theorem:

THEOREM 5.4. *On an n -dimensional manifold M ($n \geq 2$) constructed by taking two identical round balls of constant sectional curvature and gluing them together along their boundaries, the least-area surface enclosing a given quantity of volume V is as follows. (See Figure 16.)*

- (1) *If the balls are hemispheres, M is a round n -sphere, and the surface is a round $(n - 1)$ -sphere anywhere on M .*
- (2) *If the balls are spherical caps greater than hemispheres, the surface is a round sphere on one cap (possibly the singular surface).*
- (3) *Otherwise, the surface is two congruent round spherical caps meeting the singular surface orthogonally.*

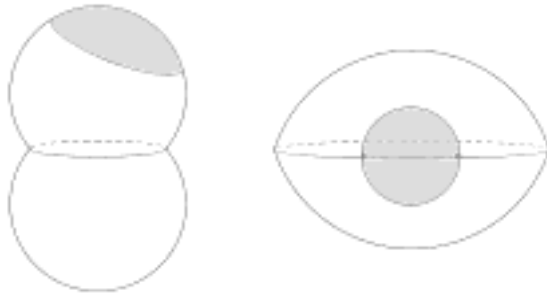


FIGURE 16. There are two types of least-area enclosures on a double disc composed of two identical caps: a circle on one cap (left), and a curve which is identical on both sides and meets the singularity orthogonally (right).

PROOF. By Proposition 5.1, the minimizer must be either a spherical cap on each side of the singularity or a round sphere on one side. In the former case, by Lemma 5.2 it consists of two identical spherical caps meeting the edge orthogonally. Lemma 5.3 allows us to determine when each case occurs. If the two balls are spherical caps

greater than hemispheres, then the curvature of the singularity is negative, and the surface crossing the edge is less efficient than a round sphere on one side. Since these are the only two candidates, the minimizer in this case must be a round $(n - 1)$ -sphere on one side. If the two balls are Euclidean, hyperbolic, or spherical caps less than hemispheres, then the curvature of the edge singularity is positive, and the minimizer is the unique surface crossing the edge orthogonally. Finally, if the two balls are hemispheres, then M is a round sphere, and by the standard isoperimetric theorem on spheres the minimizer is an $(n - 1)$ -sphere anywhere on M . \square

We conjecture that this result generalizes to manifolds made up of discs that are not identical. We believe that the statement analogous to Lemma 5.3 is that the choice between a round sphere on one side and a spherical cap on each side is determined by the *sum* of the curvatures of the singularity as viewed from each side. It is relatively straightforward to show that this sum is negative if and only if the manifold M consists of two spherical caps such that M is not convex when standardly embedded in \mathbb{R}^{n+1} . We thus have the following conjecture:

CONJECTURE 5.5. *On an n -dimensional manifold M ($n \geq 2$) constructed by taking two round balls of constant sectional curvature whose boundaries are congruent and gluing them together along their boundaries, the least area surface enclosing a given quantity of volume V is as follows:*

- (1) *Of course if M is a round n -sphere the surface is a round $(n - 1)$ -sphere anywhere on the n -sphere.*
- (2) *If M consists of two spherical caps such that M is not convex (standardly embedded in \mathbb{R}^{n+1}), the surface is a round $(n - 1)$ -sphere on one cap.*
- (3) *Otherwise the surface consists of round spherical caps in each ball meeting differentiably in a round $(n - 2)$ -sphere inside the singular set.*

In (2), when V is less than or equal to the volume of the ball of greater curvature, the surface is contained in the ball of greater curvature.

6. The cylindrical can

We continue to increase the complexity of the surfaces we are considering, this time by adding another edge singularity. The cylindrical can is a family of two-dimensional surfaces with two edge singularities, which are the circles where the lids meet the sides. We assume that the can is symmetric around the ‘vertical’ z -axis and the ‘horizontal’ xy -plane.

Like the double discs of Section 5, a cylindrical can is a $C^{1,1}$ Riemannian manifold everywhere. It follows by Proposition 2.1 that a perimeter-minimizing curve is $C^{1,1}$ everywhere. We will not need these facts to prove the results in this section.

Note: in this section, by *unique*, we mean unique up to rotation about the cylinder's axis of symmetry and reflection about the cylinder's horizontal plane of symmetry.

PROPOSITION 6.1. *The shortest curve enclosing a given quantity of area on a cylindrical can is a simple closed curve that crosses no edge more than twice.*

PROOF. We use Schwarz symmetrization as in the proof of Proposition 5.1. Let R be a region of least perimeter. Draw a straight line ℓ on the can connecting the centers of the two lids. For each circle S on the can, centered on the can's axis of symmetry, determine the length L of $R \cap S$. Replace $R \cap S$ with a circular arc of length L centered on ℓ , creating a new region R' . The areas of R and R' are the same, and the length of the boundary of R' is less than that of the boundary of R unless each slice of R by one of these circles was originally a circular arc. Each slice of ∂R by such circles thus consists of two points which subtend a circular arc (unless the minimizer coincides with the circle for positive length and thus has that circle as one of its components). We note that this construction ensures that the minimizer intersects each singularity at no more than two points, for it cannot coincide with the singularity for a positive length by Proposition 2.2.

By Proposition 2.1, we deduce that ∂R consists of one or more of the following:

- (1) Round circles on a lid or the side, centered on ℓ or the antipodal line ℓ' .
- (2) At most one constant-curvature curve crossing the top (or bottom) rim twice.
- (3) At most one constant-curvature curve crossing both rims twice.
- (4) Horizontal circles around the side.

Since any area may be enclosed with no more perimeter than that of (4), if there is one such curve it is all of ∂R . Also, if (3) occurs, it is all of ∂R , for otherwise curves of type (1) or (2) could be moved to touch it, creating an illegal singularity.

We now show that in the remaining cases the minimizer consists of one curve of either type (1) or type (2). We eliminate most possibilities by sliding curves until they touch, creating illegal singularities. There cannot be multiple curves of type (1) on either lid or the side, for then we could slide any two together until they touch. Similarly, if there is a curve of type (2), there can be no curve of type (1) on the side or on the lid it overlaps. Finally, if there are no type (2) curves, there cannot be curves of type (1) on the side and a lid, for otherwise we could move them to circles tangent at the singular curve, easily seen to be not minimizing.

At this point, there are three multi-component curves we have not ruled out: a type (1) curve on each lid, a type (2) curve on each edge, and one curve of each type on opposite lids. To eliminate these cases, we temporarily overlap these two regions and then separate them into two different curves, each with an illegal singularity. We first move the two components to the same lid and translate them until they intersect in

two places. Let M be the union of the two enclosed regions and N be the intersection. Replace the two original curves with the boundary of M on one lid and the boundary of N on the other (see Figure 17). These new curves do not touch on the side, because if they did then we could have moved the original curves until they touched. The new curves have the same length as the original ones and enclose the same total area, but both have illegal singularities, so we can reduce length and maintain area enclosed.

We conclude that the minimizer is a simple closed curve that crosses no edge more than twice. \square

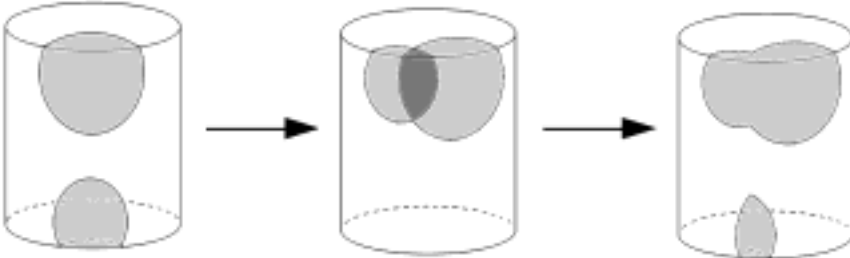


FIGURE 17. A minimizer on a cylinder cannot consist of two closed curves, one crossing each edge (left), because we can place the two curves on the same edge so that they overlap (center), and move the overlapping portion to the other edge (right). The resulting curves have the same length as the original ones and enclose the same total area, but both have illegal singularities. The same method can be used to eliminate the case of a circle on each lid and the case of a circle on one lid and a curve crossing the other edge.

We now show that a flat circle on a lid or the side cannot be a minimizer.

PROPOSITION 6.2. *On a cylindrical can, the least-perimeter enclosure of a given quantity of area cannot be a flat circle.*

PROOF. Consider a flat circle on the side. Translate this circle until it is tangent to an edge. Now truncate the circle a small distance from the edge, and translate the remainder upwards until it touches the singularity at two points p_1 and p_2 . Let A be the area lost by this truncation. On the lid, draw a circular arc connecting p_1 and p_2 that encloses an area A above the chord connecting p_1 and p_2 . Since we can make A arbitrarily small, we may assume that this arc is less than a semicircle. Since the points are closer together on the lid than on the side, the curvature of this arc is higher than that of the original circle. It is easy to show that for circular arcs less than a semicircle enclosing area against a given chord, L^2/A decreases as curvature increases. (See Figure 18.) Thus the length L_1 of the new arc between p_1 and p_2 is less than the length L_0 of the arc cut off from the circle on the side. Since the edge

of the lid curves outward, the total area enclosed is now greater than the original area, so we can perturb the portion of the circle that remains on the side so as to regain the original area and reduce length (for example, by replacing a small curved arc with a chord). Our new curve now encloses the same area as the original circle, but with a smaller length.



FIGURE 18. Replacing a small portion of a circle on the side of a cylinder with an arc on the lid increases area while decreasing length.

Now consider a flat circle on one lid. Again, we cut a small portion off the top of the circle and push the circle until it touches the edge at two points p_1 and p_2 . Let L be the length of the arc cut off. On the side, draw a circular arc of length L connecting p_1 and p_2 . (See Figure 19.) By the thread inequality [11, Theorem 2.3], this arc encloses more area than the area lost by the truncation. It is now easy to perturb the portion of the circle that remains on the side so as to regain the original area and reduce length. Again, our new curve encloses the same area as the original circle, but with a smaller length.



FIGURE 19. Replacing a small portion of a circle on the lid of a cylinder with an arc on the side increases area while maintaining length.

Alternatively, we can rule out a flat circle on one lid without using the thread inequality. If the height H of the cylinder is greater than or equal to twice the radius of this circle, we can draw this circle on the side instead and use the same argument as above. If not, find the length L of a chord a distance $H/2$ from the center of the circle. Draw chords of length L at the same place on the two lids (that is, so that the lines connecting their endpoints are vertical). Connect the four points where the chords intersect edges with curves of the same curvature as the circle. The length of this new curve is the same. However, the area has increased, since the two curves on the side are farther apart, and there is an extra lune enclosed on both lids. (See

Figure 20.) Thus the area enclosed is greater than A . To reduce the area to A , reduce the curvature of the connecting arcs identically on both lids. Since the curvature of our curve is greater than that of the edge (because the circle fit on the lid originally), the length decreases as we decrease the curvature of the arcs on top and bottom. The curve which results after we have restored the original area cannot coincide with both edges, for then then the whole closed curve would be a plane figure (on the side of the cylinder) enclosing area more efficiently than a circle. Our new curve now encloses the same area as the original circle, but with a smaller length. \square

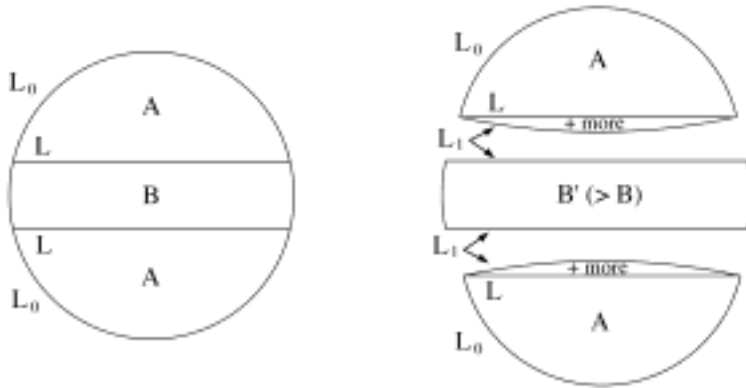


FIGURE 20. Replacing a circle on the lid of a cylinder with arcs of the same curvature on the two lids and side increases area while maintaining length.

The 1998 Geometry Group [7, Section 5] showed that on a cylindrical can of height H , for any given curvature $\kappa < 1/H$ there is a unique C^1 curve of curvature κ which crosses one edge twice, and for any given curvature $\kappa < 2/H$ there is a unique C^1 curve of curvature κ which crosses both edges twice and is symmetric about the cylinder's horizontal plane of symmetry. However, these uniqueness results are for given *curvature*, not area enclosed, so on a given cylinder, two curves of the same type, but different curvature, may enclose the same area. Such a pair of curves exists for the class of curves crossing one edge; the proof follows from Proposition 6.2. In addition, there will usually be asymmetric curves crossing both edges of the cylinder, and we have not yet found a way to classify these curves.

THEOREM 6.3. *The least-perimeter enclosure of a given area A on a cylindrical can is one of the following (see Figure 21).*

- (1) a horizontal circle around the side of the cylinder;
- (2) a C^1 constant curvature curve which crosses one edge twice; or

(3) a C^1 constant curvature curve which crosses both edges twice. (We conjecture that this curve must be symmetric about the cylinder's horizontal plane of symmetry.) Each case occurs as a minimizer for some area on some cylinder.

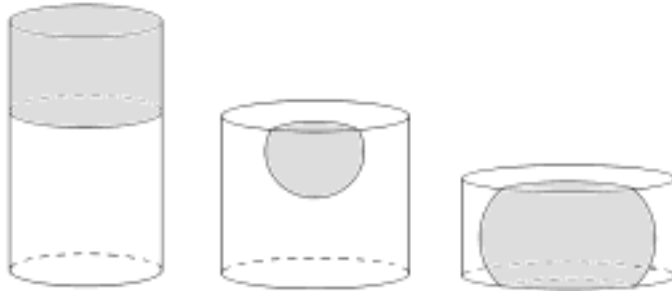


FIGURE 21. There are three types of least-perimeter enclosures on the cylinder.

PROOF. By Proposition 6.1, a minimizer must be a one-component constant-curvature curve that crosses no edge more than twice. There are two types of curves that do not cross any edges: flat circles and geodesics around the side of the cylinder. By Lemma 6.2, the flat circle cannot be a minimizer. Hence every minimizer is one of the three types stated above.

It remains to show that each type occurs. It is clear that curves of type (1) are best for enclosing half the area on a tall, thin cylinder. It is also clear that curves of type (2) are best for enclosing a very small area on any cylinder. We claim that curves of type (3) are best for enclosing half the area on a short, fat cylinder. To see this, observe that a type (1) curve would have length $2\pi R$, and a type (2) curve would have to enclose most of the top and go nearly all the way around the side, so it would have length nearly $2\pi R$ or greater. On the other hand, the type (3) curve that is a geodesic passing through the center of both lids would have length $4R + 2H$, which is less. \square

Note that a minimizer need not be unique. For example, at a transition point between type (2) and type (1) curves on a tall, thin cylinder, there would necessarily be a curve of each type that is a minimizer for the same area.

Numerical calculations with a computer program indicate that a curve about one edge with curvature $\kappa \approx 0.8207/R$ has length $L = 2\pi R$ and area $A \approx 3.7052R^2$. So for a sufficiently tall cylinder (that is, one for which minimizers never cross two edges), this will be the transition point between the two types of least-perimeter enclosures. Furthermore, if $3.7052R^2 \geq \pi R^2 + RH/2$ ($= \text{area}(S)/2$), that is, if $H/R \leq 1.1272$, then the least-perimeter enclosures will never be geodesics.

In addition, the smallest curves across two edges will have $r > H/2$, with perimeter $L > 2\pi r$, so that geodesics beat two-edge curves (both symmetric and asymmetric ones) if $2\pi R \leq \pi H$, that is, if $H/R \geq 2$. Hence a cylinder for which $H/R \geq 2$ will never have a minimizer that crosses both edges.

Beyond these results, the determination of which types of minimizers can occur on a given cylinder remains an open question. Calculation of the exact transition points on a given cylinder in terms of area enclosed also remains open.

We also note that in the limit where the height H is zero, the cylinder becomes a manifold consisting of two identical flat discs. By Theorem 5.4, all minimizers on this ‘double disc’ consist of two identical circular arcs meeting the edge orthogonally. This curve corresponds to the symmetric curve on the cylinder that crosses both edges twice.

Finally, Theorem 6.3 allows for the possibility that asymmetric curves crossing both edges can be minimizing. We believe that this is never the case, but the formulae are extremely complicated, and we have not found a nice geometric argument. We thus have the following conjecture:

CONJECTURE 6.4. *On a cylindrical can, there is at most one C^1 constant curvature curve which crosses both edges twice, encloses a given area, and is not symmetric about the horizontal plane of symmetry. Any such curve is unstable and thus cannot be a least perimeter enclosure.*

Acknowledgements

The authors were the members of the 1998 and 2000 Geometry Groups of the Williams College NSF SMALL undergraduate research program. Related work on geodesic nets began with the 1995 and 1996 Geometry Groups. Work was partially funded by the National Science Foundation and Williams College. The authors thank Professor Frank Morgan for his invaluable advising, the 2001 Geometry Group (Joseph Corneli, Paul Holt, Nicholas Leger, and Eric Schoenfeld) for their comments, and Sang Pakh for suggesting that the 1998 group investigate the cylinder.

References

- [1] F. Almgren, ‘Spherical symmetrization’, in: *Proceedings of the International Workshop on Integral Functionals in Calculus of Variations, (Trieste, 1985)* Suppl. Rend. Circ. Mat. Palermo (2) **15** (1987) pp. 11–25.
- [2] M. Barber, J. Tice and B. Wecht, ‘Geodesics and geodesic nets on regular polyhedra’, Williams College NSF ‘SMALL’ undergraduate research Geometry Group report (1995).

- [3] H. Bray and F. Morgan, 'An isoperimetric comparison theorem for Schwarzschild space and other manifolds', *Proc. Amer. Math. Soc.* **130** (2002), 1467–1472.
- [4] Yu. D. Burago and V. A. Zalgaller, *Geometric inequalities* (Springer, New York, 1980).
- [5] A. Erchak, T. Melnick and R. Nicholson, 'Geodesic nets on regular polyhedra', Williams College NSF 'SMALL' undergraduate research Geometry Group report (1996).
- [6] H. Federer, *Geometric measure theory* (Springer, New York, 1969).
- [7] A. Gnepp, T. Ng and C. Yoder, 'Isoperimetric domains on polyhedra and singular surfaces', Williams College NSF 'SMALL' undergraduate research Geometry Group report (1998).
- [8] A. Heppes, 'e-mail communication to M. Barber, J. Tice, B. Wecht and F. Morgan', 1995.
- [9] H. Howardsa, M. Hutchings and F. Morgan, 'The isoperimetric problem on surfaces', *Amer. Math. Monthly* **106** (1999), 430–439.
- [10] A. O. Ivanov and A. A. Tuzhlin, *Minimal networks. The Steiner problem, and its generalizations* (CRC Press, Boca Raton, 1994).
- [11] F. Morgan, 'An isoperimetric inequality for the thread problem', *Bull. Austral. Math. Soc.* **55** (1997), 489–495.
- [12] ———, *Geometric measure theory: a beginner's guide*, 3rd edition (Academic Press, San Diego, 2000).
- [13] ———, 'Area-minimizing surfaces in cones', *Comm. Anal. Geom.* **10** (2002), 971–983.
- [14] ———, 'Regularity of isoperimetric hypersurfaces in Riemannian manifolds', *Trans. Amer. Math. Soc.* **355** (2003), 5041–5052.
- [15] F. Morgan and M. Ritoré, 'Isoperimetric regions in cones', *Trans. Amer. Math. Soc.* **354** (2002), 2327–2339.

c/o Frank Morgan

Department of Mathematics

Williams College

Williamstown, MA 01267

USA

e-mail: Frank.Morgan@williams.edu
acotton@math.berkeley.edu
dfreeman@math.berkeley.edu
gnepp@post.harvard.edu
ngtf@seas.upenn.edu
John.Spivack@williams.edu
cyoder@wso.williams.edu