An isoperimetric theorem in plane geometry

Alan Siegel[1]

COURANT INSTITUTE OF MATHEMATICAL SCIENCES
NEW YORK UNIVERSITY
New York

Abstract

Let P be a simple polygon. Let the vertices of P be mapped, according to a counterclockwise traversal of the boundary, into a strictly increasing sequence of real numbers in $[0, 2\pi)$. Let a ray be drawn from each vertex so that the angle formed by the ray and a horizontal line pointing to the right equals, in measure, the number mapped to the vertex. Whenever the rays from two consecutive vertices intersect, let them induce the triangular region with extreme points comprising the vertices and the intersection point. It is shown that there is a fixed α such that if all of the assigned angles are increased by α, the triangular regions induced by the redirected rays cover the interior of P.

This covering implies the standard isoperimetric inequalities in two dimensions, as well several new inequalities, and resolves a question posed by Yaglom and Boltanskiĭ.

1 Introduction and summary

The isoperimetric theorem states that among all planar regions with a given boundary length p, the disk has the maximum area. This bound, which is also known as the isoperimetric inequality, dates back to antiquity; its literary debut occurs—albeit obliquely—in the Aeneid [9]. There are also more specialized forms of the inequality for simple polygons.

We take all polygons and boundary curves to be simple. If P is a polygon, let P^{rgn} denote the bounded polygonal region with $\partial(P^{rgn}) = P$. Let $Area(P) = Area(P^{rgn})$.

Definition 1. Let ℓ be a ray or directed line. The *angular direction* of ℓ is the measure of the angle formed by ℓ and a horizontal ray h that originates at some point on ℓ and runs to the right. The measure of such an angle is taken in a counterclockwise direction from h to ℓ.

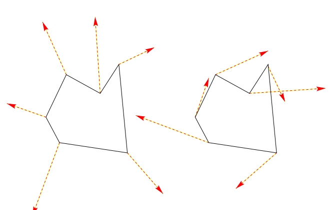

Figure 1

Definition 2. Let P be an n-gon with the consecutive vertices v_1, v_2, $\ldots$, v_n, which are sequenced in counterclockwise order around P. Let, for $i = 1, 2, \ldots, n$, r_i be an infinite ray that originates at vertex v_i and has an angular direction θ_i. The rays $r_1, r_2, \ldots, r_n$ are called a *monotone slicing for* P if $\theta_1 < \theta_2 < \cdots < \theta_n < 2\pi + \theta_1$. We say that the slicing rays are *associated* with P, and the angles $\theta_1, \ldots, \theta_n$ are *associated* with the rays.

Definition 3. Let P be an n-gon and $r_1, r_2, \ldots, r_n$ be a monotone slicing for P. Let t_j be the region whose boundary is the triangle defined by the base $\overline{v_j v_{j+1}}$ and the intersection of the consecutive[2] rays r_j and r_{j+1} as the opposing vertex, provided $r_j \cap r_{j+1}$ exists, and let $t_j = \emptyset$ otherwise. We call t_j a *slicing triangle* for P, and say that the set of slicing triangles $\{t_j\}_{j=1}^n$ is *associated* with the rays $\{r_i\}_{i=1}^n$.

[1]This research was supported in part by NSF grant CCR-9503793

[2]Of course, the index value $n + 1$ is interpreted as 1.

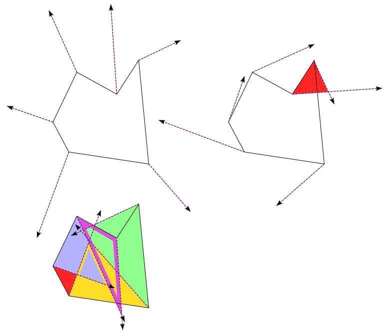

Figure 2

Theorem 1. Let P be an n-gon with consecutive vertices $v_1, v_2, \ldots, v_n$, and let $r_1, r_2, \ldots, r_n$ be a monotone slicing for P with associated angles $\theta_1, \theta_2, \ldots, \theta_n$. Let $R(\alpha) = \{r_j(\alpha)\}_{j=1}^n$ be the monotone slicing for P with associated angles $\theta_j + \alpha$, for $j = 1, 2, \ldots, n$. Let $T(\alpha)$ be the set of slicing triangles associated with $R(\alpha)$. Then for some fixed α,

$$P^{rgn} \subset \cup_j t_j(\alpha).$$

That Theorem 1 is an isoperimetric inequality can be seen from the following corollaries.

Corollary 1. Let F be a polygon. Then among all polygons with edges congruent to those of F, those polygons whose vertices lie on a circle have the greatest area.

Of course, Corollary 1 is standard. The proof is algorithmic. Let G be a polygon that is inscribed in a circle, and which has sides congruent to those of F. The construction splits G^{rgn} into a collection of n slicing triangles that are mapped onto a set of slicing triangles that cover F^{rgn}. The mapping is area reducing and elementary.

Let $\phi(s)$ be a curve parameterized by arclength. Let $\eta(s)$ be the angular direction of a normal to ϕ at $\phi(s)$. We say that ϕ is piecewise smooth if it can be partitioned into a finite number of pieces where $\eta(s)$ is continuously differentiable on the closure of each piece. Let $\eta(s+) = \lim_{t\downarrow s} \eta(t)$, and let $\eta(s-) = \lim_{t\uparrow s} \eta(t)$.

Corollary 2. Let C be a piecewise smooth closed simple curve with perimeter p and which encloses a region with area A. Let C be parameterized by $\phi(s)$, for the arclength parameter s, with $0 \le s \le p$. Let $\eta(s)$ be the angular direction of the outward normal to C at $\phi(s)$. Let $\theta(s)$ be monotone increasing with $\theta(0) = 0$ and $\theta(p) = 2\pi$. Let $\dot\theta(s) = \frac{d}{ds}\theta(s)$. Then

$$A \le \max_\alpha \left\{ \int_0^p \frac{\cos^2(\alpha + \theta(s) - \eta(s))}{2\dot\theta(s)} ds \right\}.$$

To see that this is an isoperimetric inequality, let $\theta(s) = \frac{2\pi s}{p}$. Substituting gives

$$A \le \max_\alpha \left\{ \int_0^p \frac{p\cos^2(\alpha + \theta(s) - \eta(s))}{4\pi} ds \right\} \le \frac{p}{4\pi}\int_0^p ds = \frac{p^2}{4\pi},$$

and equality can hold only if $\frac{d\eta(s)}{ds}$ is constant, which is to say that C is a circle. This special case is the most common formulation of the isoperimetric inequality in two dimensions.

A strong consequence of the proof given for Theorem 1 is the following.

Corollary 3. Let C, s, ϕ, η and $\theta(s)$ be as in Corollary 2. Let R be the finite region bounded by C. Let $\ell(s, \alpha)$ be a directed line segment that originates at $\phi(s)$, has an angular direction of $\theta(s) + \alpha$, and has length $\max(0, \frac{-\cos(\alpha+\theta(s)-\eta(s-))}{\dot\theta(s)}, \frac{-cos(\alpha+\theta(s)-\eta(s+))}{\dot\theta(s)})$. Then for some constant α,

$$R \subset \cup_s \ell(s, \alpha).$$

Theorem 1 also has an immediate interpretation as an area minimization formulation.

Corollary 4. Let P be an n-gon with side lengths of $\ell_1, \ell_2, \ldots, \ell_n$. Let $\theta_1, \theta_2, \ldots, \theta_n$ be nonnegative values that sum to 2π. Then

$$\sum_{j=1}^n \frac{1}{4}\ell_j^2 \cot(\frac{\theta_j}{2}) \ge Area(P),$$

and equality holds if and only if P is inscribed in a circle and each θ_j equals the radial angle of the arc subtended by a chord of length ℓ_j.

2 Proofs of Corollaries 1–4

The main part of Corollary 1 is established first.

Corollary 1a. Let P be an n-gon with consecutive vertices $v_1, v_2, \ldots, v_n$, and let S be a polygon that is inscribed in a circle, and which has edges with the same lengths as those of F. Suppose that S^{rgn} contains the center of its circumscribing circle.

Then

$$Area(F) \leq Area(S),$$

and equality holds if and only of F can be inscribed in a circle.

Proof: The first step is to show that S exists. Create a path of $n-1$ segments that are congruent to the sides of P, apart from some side of maximum length s. Let the path have its vertices placed along a circle of huge radius, and consider how the path is forced to curl up as the radius is decreased. Evidently, the endpoints of the path would begin at a distance that equals the sum of the segment lengths, when the circle radius is infinite, and decrease continuously as the radius diminishes. The existence of P ensures that s is no more than than this initial sum, so the radius can be decreased until the distance between the endpoints equals s. Then the path can be closed with the omitted edge. For completeness, it should be noted that the construction would fail if the diameter of the circle were decreased to a value that is smaller than some edge in the chain of $n-1$ segments, since such a segment cannot be a chord of such a circle. But since all such lengths are bounded by s, and since the diameter must be at least s when the termination condition is satisfied, this failure cannot occur.

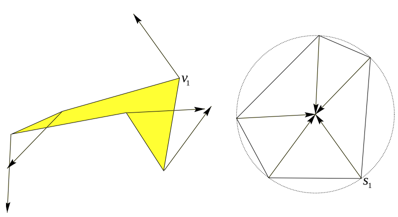

Figure 3

Let S have the consecutive vertices $s_1, s_2, \ldots, s_n$. Suppose that S has a natural correspondence with P where $|\overline{v_j v_{j+1}}| = |\overline{s_j s_{j+1}}|$ for all j. In addition, suppose that $s_1, s_2, \ldots, s_n$ circulate about S with the same rotational sense as $v_1, \ldots, v_n$ rotate about P. Let c be the circumcenter of S. Define $\{\hat{r}_j\}_{j=1}^{n}$ to be the monotone slicing of S where $\hat{r}_j = \overrightarrow{s_j c}$, and let $\{r_j\}_{j=1}^{n}$ be the monotone slicing of P where r_j begins at v_j and is parallel to $\hat{r}_j$. Let the slicing triangles for $\{\hat{r}_j\}_{j=1}^{n}$ be $\hat{t}_j = (\triangle s_{j+1} s_j c)^{rgn}$, for $j = 1, 2, \ldots, n$.

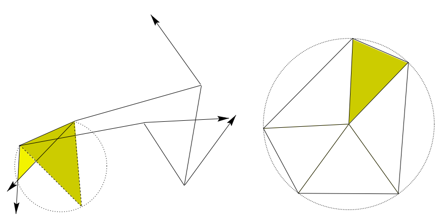

Figure 4

According to Theorem 1, the rotated slicing rays $\{r_j(\alpha)\}_{j=1}^{n}$ must, for some fixed α, induce a set of slicing triangles that cover P^{rgn}. But the slicing triangle $t_j(\alpha)$, if it exists, will have a base that is congruent to the corresponding base of $\hat{t}_j$. Moreover, the angle subtending $t_j(\alpha)$'s base must equal $\angle s_j c s_{j+1}$ since rotating r_j and r_{j+1}, by α will keep their angle of intersection the same as that of $\hat{r}_j$ and $\hat{r}_{j+1}$, provided $r_j(\alpha)$ and $r_{j+1}(\alpha)$, intersect. The Peripheral Angle Theorem in basic Euclidean Geometry says that $r_j \cap r_{j+1}$ will move along the boundary of a fixed circle and will have, therefore, a maximum altitude precisely when $t_j(\alpha)$ is isosceles, in which case it will be congruent to $\hat{t}_j$. Of course $Area(S) = \sum_{j=1}^{n} Area(\hat{t}_j)$, and hence

$$Area(P) \leq \sum_j Area(t_j(\alpha)) \leq \sum_j Area(\hat{t}_j) = Area(S). \tag{1}$$

Evidently, equality can only hold only if $t_j(\alpha) \cong \hat{t}_j$ for all j, and $Area(t_i \cap t_j) = 0$, for $i \neq j$, in which case P will be congruent to S.

Finally, it should be noted that the edges of S need not have a natural correspondence with those of P. Since each $\hat{t}_j$ has radial sides of length r, where r is the radius of the circumcircle, these triangles can be permuted to ensure that the edges of S are sequenced to correspond with those of P. ∎

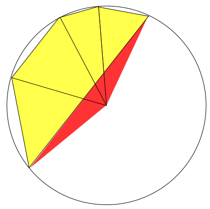

Figure 5

The less interesting part of Corollary 1, which concerns the case when S^{rgn} does not contain its circumcenter, is postponed. This case is different because the largest side in S will be the base of a triangle with a reflex angle, and this triangle will contribute a negative term in the formulation of $Area(S)$ as given in equation 1. The specifics comprise Corollary 1b in Section 4.

Corollary 2 is just a matter of applying Theorem 1 and passing to the limit.

Proof of Corollary 2. Let C be approximated by an n-gon P. Let the vertices of P be $v_j = \phi(\frac{pj}{n})$, for $j = 1, 2, \ldots, n$. Let $r_j(\alpha)$ and $r_{j+1}(\alpha)$ be slicing rays with angular directions $\theta(p\frac{j}{n}) + \alpha$ and $\theta(p\frac{j+1}{n}) + \alpha$ and respective origins v_j and v_{j+1}. Suppose that the two rays intersect, and let their intersection be the point w_j. Let $|\overline{v_j v_{j+1}}| = \Delta_j$, and let η_j be the angular direction of the outward normal to $\overline{v_j v_{j+1}}$. Then by the law of sines, $|\overline{v_j w_j}| = \frac{\Delta_j}{\sin(\theta_{j+1} - \theta_j)} \sin(\eta_j - \frac{\pi}{2} - \theta_{j+1} - \alpha)$, and the area of the slicing triangle is

$$Area(\triangle v_j v_{j+1} w_j) = \frac{(\Delta_j)^2}{2\sin(\theta_{j+1} - \theta_j)} \cos(\eta_j - \theta_{j+1} - \alpha) \cos(\eta_j - \theta_j - \alpha).$$

Theorem 1 says that for some α,

$$\sum_j \frac{(\Delta_j)^2}{2\sin(\theta_{j+1} - \theta_j)} \cos(\eta_j - \theta_{j+1} - \alpha) \cos(\eta_j - \theta_j - \alpha) \geq Area(P).$$

Passing to the limit gives the formulation: For some α,

$$\int \frac{\cos^2(\eta(s) - \theta(s) - \alpha)}{2\frac{d\theta}{ds}} ds \geq A. \quad \blacksquare$$

Proof of Corollary 3 (Sketch). Corollary 3 is little more than an application of Corollary 2 in the limit. The restriction to positive lengths is a consequence of the actual covering argument given for Theorem 2, which establishes that every point $x \in P$ will be covered by some slicing triangle $(\triangle v_j v_{j+1} w_j)^{rgn}$ whose base $\overrightarrow{v_j v}_{j+1}$ has a counterclockwise orientation with respect to x. $\blacksquare$

Proof of Corollary 4. Similarly, Corollary 4 is just an analytic interpretation of Theorem 1, since each term $\frac{1}{4}\ell_j^2 \cot(\frac{\theta_j}{2})$ is the largest possible area that a triangle with a base length of l_j and an opposing vertex angle of θ_j can have.

However, if some angle θ_j exceeds π, then the corresponding area term is negative, and additional justification is necessary. For specificity, let the reflex angle be θ_n. Let P be an n-gon that is inscribed in a circle and has side lengths ℓ_j, for $j = 1, 2, \ldots, n$, and let its vertices be $v_1, v_2, \ldots, v_n$. Let Q be the $(n+1)$-gon resulting from adjoining P with the exterior isosceles triangle $\triangle v_0 v_1 v_n$, where vertex v_0 satisfies $\angle v_1 v_0 v_n = 2\pi - \theta_n$.

Let $r_1 = \overrightarrow{v_1 v_0}$, and define $r_2, r_3, \ldots, r_n$ so that r_j and r_{j+1} intersect at an angle of θ_j. By construction, $r_n = \overrightarrow{v_n v_0}$. Theorem 2 as stated below and its mild extension in Corollary 5 guarantee that

$$\sum_{j=1}^{n-1} \frac{1}{4}\ell_j^2 \cot(\frac{\theta_j}{2}) \geq Area(Q).$$

It follows that

$$\sum_{j=1}^{n-1} \frac{1}{4}\ell_j^2 \cot(\frac{\theta_j}{2}) - Area(\triangle v_0 v_1 v_n) \geq Area(P).$$

Theorem 1 combined with Corollary 1b establish the circumstances where equality can hold. $\blacksquare$

3 Proof of main theorem

Theorem 1 follows from the special case described next.

Theorem 2
Let P be an n-gon with consecutive vertices $v_1, v_2, \ldots, v_n$ that sequence counterclockwise about P, and let $r_1, r_2, \ldots, r_n$ be a monotone slicing for P with associated angles $\theta_1, \theta_2, \ldots, \theta_n$, and associated slicing triangles $\{t_j\}_{j=1}^n$. Suppose that $r_1 = \overrightarrow{v_1 v_n}$, $r_n = \overrightarrow{v_n v_1}$, and let v_1 and v_n be on the x axis with v_1 to the right of v_n, so that $r_2, r_3, \ldots r_{n-1}$ point downward. Let D be the upper halfplane.

Then
$$P^{rgn} \cap D \subset \cup_{j=1}^{n-1} t_j.$$

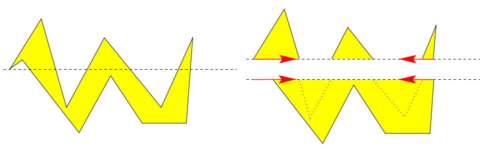

Figure 6

Notice that t_n is not used to cover $P^{rgn} \cap D$. This suggests that Theorem 1 might be established by locating a line that partitions P^{rgn}—convex or otherwise—in a way where Theorem 2 can be applied to each piece.

Proof of Theorem 2: Suppose that $P^{rgn} \cap D$ is not contained in $\cup_{j=1}^{n-1} t_j$. Then there is a point $Q \in P^{rgn} \cap D$ that is in general position and does not belong to $\cup_{j=1}^{n-1} t_j$. In particular, Q can be selected so that none of the lines from Q through the vertices of P is parallel to any of the rays $r_1, r_2, \ldots, r_n$.

Let the boundary of P be parameterized in polar coordinates with respect to the origin Q, and let $(\Theta(t), r(t))$ define the boundary curve, for $t \in [0, 1]$. Let $(\Theta(0), r(0)) = v_1$, and let this parameterization traverse the boundary with a counterclockwise rotation.

Draw a horizontal line ℓ through Q. By hypothesis, ℓ lies above $\overline{v_1 v_n}$. We can assume that the line does not intersect any of the vertices. A polar representation is used to select a subset of edges that all rotate counterclockwise about Q, and that define a figure with the following properties. The polygonal figure should intersect Q, lie in the closed upper halfplane defined by ℓ, and be star-convex with respect to Q. See the shaded region in Figure 7. Formally, we find a sequence $t_0, t_1, \ldots, t_k$ where

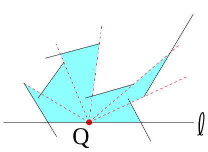

Figure 7

a) $0 \leq t_0 < t_1 < \cdots < t_k \leq 1$,

b) $\Theta(t_0) < 0 < \Theta(t_1) < \cdots < \Theta(t_k) < \pi < \Theta(t_{k+1})$,

c) $\Theta(t_{i+1}) - \Theta(t_i) < \pi$,

d) P contains an edge that begins at $(\Theta(t_j), r(t_j))$, has a counterclockwise orientation with respect to Q, and terminates at some location $(\Theta(w_j), r(w_j))$ where $\Theta(w_j) \geq \Theta(t_{j+1})$, for $j = 0, 1, 2, \ldots, k$.

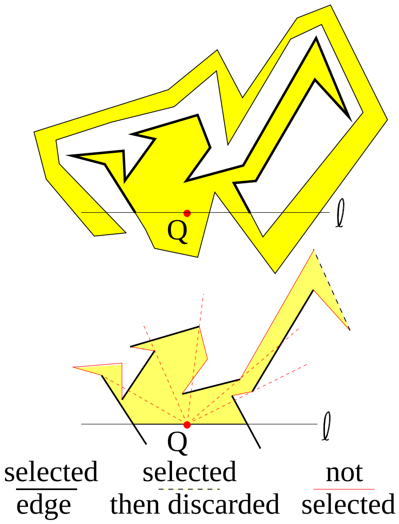

Figure 8

The existence of these edges follows from the fact that the boundary has a winding number of 1 about any interior point. An argument that appears to lie within Euclidean geometry is illustrated in Figure 8, and is sketched as follows. The horizontal line ℓ partitions the boundary into a finite number of pieces. Make each piece a closed curve by drawing a line segment between its two endpoints. Of the pieces that are in the upper halfplane, there will be one whose new edge contains Q and is of minimal length. Let this portion of the boundary be called ϕ. Let us walk along the edges of ϕ in a counterclockwise direction with respect to the interior of ϕ, and monitor the polar angle for the endpoints of each edge. Every time a new maximum is reached in the polar angle, the current edge is selected regardless of whether it can be seen from Q. This selected subset is then postprocessed by discarding any edge that cannot be seen from Q, because of other edges in the selected subset.

A more precise formulation is the following. Let $e_1, e_2, \ldots, e_z$ be the edges of ϕ in consecutive order, beginning with the edge that intersects ℓ to the right of Q. A suitable set of edges can be selected for subsequent postprocessing as follows.

Create the empty set S;
$\Theta \leftarrow 0$;
for $j \leftarrow 1$ **to** z **do**
    Assign the maximum of the polar angle values for the endpoints of e_j to α;
    if $\alpha > \Theta$ **then**   { *α is a new maximum angle* }
        $\Theta \leftarrow \alpha$;   { *Remember the new maximum* }
        Include e_j in S
    endif
endfor.

The postprocessing is equally direct; the interested reader may wish to consider how to design an elegant solution.

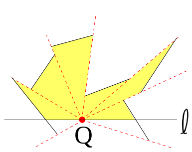

Figure 9

Let each of the resulting edges (after postprocessing) be trimmed so that consecutive segments have an endpoint on a common radial line as shown. It is convenient to use a trimming procedure that uses radial lines through vertices of P. Similarly, let the first and last segment in the sequence penetrate below ℓ as illustrated in Figure 9. Let $s_1, s_2, \ldots, s_k$ be the sequence of segments. The desired set of angles $\{\Theta(t_i)\}$ is defined by the angles of the radial lines through the endpoints of these segments.

Each s_i, for $i = 1, 2, \ldots, k$ will be assigned two slicing rays. Let s_i be a trimmed version of the edge e_j, Then each endpoint of s_i will have an emanating slicing ray that is parallel to the ray from the corresponding endpoint of e_j.

Since s_i has rays that are parallel to the rays of its corresponding untrimmed edge, the slicing triangle formed from any s_i and its rays must be contained within the corresponding slicing triangle for P. Thus, it suffices to show that Q is contained in the slicing triangle formed from some s_i and its associated rays.

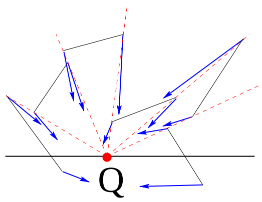

Figure 10

Since none of the original slicing rays was parallel to any of the lines from Q through the vertices of P, none of the slicing rays associated with the s_i will pass through Q. Let a slicing ray be called *right-oriented* if it begins above ℓ and intersects ℓ to the right of Q. Similarly, a ray is *left-oriented* if it begins above ℓ and intersects ℓ to the left of Q.

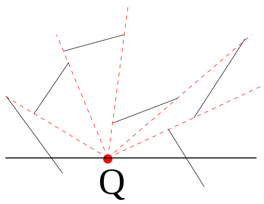

Figure 11

For any edge s_i, let its two endpoints be denoted by the forward endpoint and the rear endpoint, where the polar angle for the forward endpoint has a larger value than that of the rear. Similarly, each radial line will have two incident edges, which are identified as forward and rear. As shown in Figure 11, the edges $s_1, s_2, \ldots, s_k$ have a natural order that begins with an edge that intersects ℓ on the right, and proceeds in counterclockwise order through to the last edge, which intersects ℓ on the left.

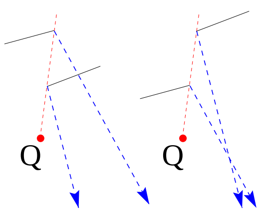

Figure 12

Remark: as Figure 12 shows, if the ray emanating from the forward endpoint of a radial line's rear incident edge is right-oriented, then so is the ray that emanates from the rear vertex of the radial line's forward edge. This property is just a straightforward consequence of the monotone rotational increment in the direction of consecutive slicing rays. Of course, the analogous fact holds when the words left and right are exchanged, along with the notions forward and rear.

To locate a slicing triangle that covers Q, let these edges be searched in sequential polar order for the first associated ray that is left-oriented. In view of the preceding remark, this ray, if it exists, must emanate from the forward endpoint of its associated segment. There are three cases.

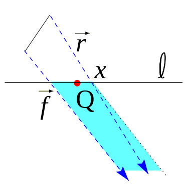

Figure 13

Case 1: The sought ray is associated with a segment that lies entirely above ℓ. Let the segment's rear ray be $\vec{r}$, and let the forward ray be $\vec{f}$. In this situation, Q is trapped between $\vec{f}$ and $\vec{r}$. Let $\vec{r}$ intersect ℓ at x. Draw a ray at x that is parallel to $\vec{f}$. These two rays define a half infinite strip that $\vec{r}$ must enter at x (due to monotonicity). Consequently, $\vec{f}$ and $\vec{r}$ must intersect below ℓ to form a slicing triangle that contains Q.

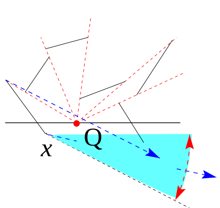

Figure 14

Case 2: All segments that lie entirely above ℓ have emanating rays that are right-oriented. In this circumstance, the reasoning of Case 1 applies to segment s_k, which has a forward endpoint x that lies below ℓ. Notice that the associated ray emanating from x must lie in the cone that has its vertex at x, one side parallel to the ray associated with the rear of s_k (due to monotonicity) and the other parallel to ℓ and pointing the right (since no ray can be rising).

Case 3: All segments that lie entirely above ℓ have emanating rays that are left-oriented. This circumstance is, mutatis mutandis, the same as case 2 with all notions of left and right exchanged.

To ensure correctness, we are obliged to discuss a case that cannot occur. The underlying question is why the above argument fails to hold for the portion of the polygon that lies below the x axis. The answer is that such points Q will not be trapped by either r_1 or r_n, and therefore might not be caught between any pair of intermediate rays. Moreover, the monotonicity argument as illustrated in Figure 12 would not hold if $\overline{v_n v_1}$ were an intermediate (and omitted) edge that belonged to the portion of boundary named ϕ. But this event cannot occur since ϕ lies above ℓ.

The proof is now complete. ∎

Two mild extensions of Theorem 2 are needed to complete the proof of Theorem 1.

Corollary 5. Let P be an n-gon with consecutive vertices $v_1, v_2, \ldots, v_n$, and let $r_1, r_2, \ldots, r_n$ be a monotone slicing for P with associated angles $\theta_1, \theta_2, \ldots, \theta_n$, and associated slicing triangles $\{t_j\}_{j=1}^{n}$. Let v_1 and v_n be on the x axis, with v_1 to the right of v_n. Let D be the upper halfplane.

5a) Suppose that r_1 and r_n intersect, and the intersection point lies below the x-axis. Then

$$P^{rgn} \cap D \subset \cup_{j=1}^{n-1} t_j.$$

5b) Suppose that r_1 points at v_n, v_{n-1} lies in the upper halfplane, and r_{n-1} intersects the x-axis at or to the left of v_n. Then

$$P^{rgn} \cap D \subset \cup_{j=1}^{n-1} t_j.$$

Proof: For both 5a and 5b, the rays $r_2, \ldots, r_{n-1}$ are constrained to point downward. Moreover, if r_1 and r_n are replaced by rays that point at at each other, then the slicing rays would still be monotone and satisfy the conditions of Theorem 2.

For 5a, this change only affects three slicing triangles, which each might lose coverage below the x-axis but are otherwise unaffected.

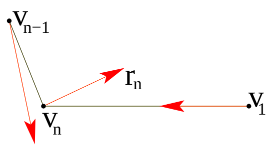

Figure 15

For 5b, only triangles t_{n-1} and t_n could possibly be affected by the change in r_n. However, t_n will remain fixed (as a segment with zero area), and t_{n-1} will not exist for either direction of r_n, due to the errant pointing of r_{n-1}, unless it points at v_n, in which case t_n will also be a degenerate triangle. So the coverage will be the same. $\blacksquare$

The intended application for part 5b is instances where r_n points upward as shown.

Proof of Theorem 1: Let, for $j = 1, \ldots, n$, angle β_j satisfy $\beta_j = \theta_{j+1} - \theta_j$. Intuitively, β_j is associated with edge $\overline{v_j v_{j+1}}$. Suppose, for the moment, that $\beta_j < \pi$, for all j.

Given an edge and its associated angle β, a suitable interpolation scheme is needed to compute the appropriate portion of β to associate with a given portion of the edge. The interpolation, it turns out, will only be needed for edges that are intersected by a transversal.

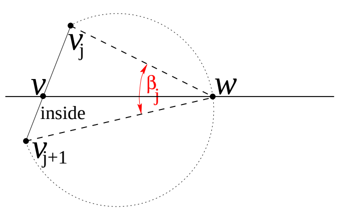

Figure 16

Let the transversal ℓ intersect the edge $\overline{v_j v_{j+1}}$ at the point v, and suppose that one side of $\overline{v_j v_{j+1}}$ is designated as the inside. Let w be a point on ℓ that is on the inside half of ℓ as defined by the designation for $\overline{v_j v_{j+1}}$, and suppose that $\angle v_j w v_{j+1} = \beta_j$. Then interpolation rule associates $\angle v w v_{j+1}$ with the segment $\overline{v v_{j+1}}$, and $\angle v_j w v$ with the segment $\overline{v_j v}$.

The interpolation procedure introduces v as an artificial vertex on $\overline{v_j v_{j+1}}$. The vertex will also need an associated slicing ray r_v. Let r_v emanate from v with an angular direction of $\theta_v \equiv \theta_j + \angle v_j w v$.

Now let ℓ be a horizontal line that lies above P, and let ℓ be translated downward in a continuous motion, so that ℓ will eventually intersect P and will eventually wind up below the polygon. We can suppose that ℓ is not parallel to any of the lines defined by pairs of vertices of P. Let this family of sweep lines be denoted by $\ell(\rho)$, where ρ ranges from 0 to 1.

Let $p_1(\rho)$ be the rightmost intersection of $\ell(\rho)$ with the boundary of P, and let $p_2(\rho)$ be the leftmost such point. For transversals of interest, $p_1(\rho)$ and $p_2(\rho)$ will split the boundary of P into two pieces, which we name the upper and lower boundaries.

Let $\sigma_1(\rho)$ and $\sigma_2(\rho)$ be the sums of the angles associated with the segments belonging to, respectively, the upper and lower boundaries of P. If $p_1(\rho)$ or $p_2(\rho)$ are not vertices, the interpolation rule is used to apportion a fraction of the relevant angle(s) to each sum.

Ideally, there ought to be a ρ and transversal $\ell(\rho)$ where $\sigma_1(\rho) = \sigma_2(\rho) = \pi$. While this would would be evident if $\sigma(\rho)$ were continuous, the function can have jumps when ℓ intersects a vertex. We defer, momentarily, an analysis of the discontinuous case, which includes the possibility that the interpolation procedure might fail.

So suppose that there is a ρ_0 where $\sigma_1(\rho_0) = \sigma_2(\rho_0) = \pi$. Then the conditions of Theorem 2 can be satisfied as follows. First, interpolation is used to create artificial vertices and associated rays as needed at $p_1(\rho_0)$ and $p_2(\rho_0)$. Let α be the angle between the ray emanating from $p_1(\rho_0)$ and a ray pointing from $p_1(\rho_0)$ to $p_2(\rho_0)$. Then a rotation of the slicing rays by α will force the ray emanating from $p_1(\rho_0)$ to point at $p_2(\rho_0)$. Similarly, the ray emanating from p_2 will point at $p_1(\rho_0)$.

In this case, Theorem 2 shows that the coverage is complete for both halves of P^{rgn} as determined by $\ell(\rho_0)$. Moreover, the interpolation rule ensures that the artificial rays are unnecessary, since any such ray and its two neighboring rays intersect at a common point as illustrated in Figure 16.

However, the interpolation process might fail for two reasons.

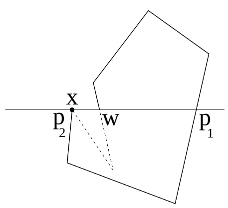

Figure 17

First, suppose that $\ell(\rho)$ intersects a vertex x that has both of its connecting edges on the same side of the transversal, and suppose that x equals $p_1(\rho)$ or $p_2(\rho)$. For specificity, let $x = p_2(\rho)$, and let the edges connecting $p_2(\rho)$ lie in the lower halfplane bounded by $\ell(\rho)$. Let w be, in this case, the next-to-the-leftmost intersection of $\ell(\rho)$ and the boundary of P (i.e., the leftmost intersection that is to the right of $p_2(\rho)$). By the choice of ℓ, w cannot be a vertex of P and must therefore belong to some edge that actually crosses ℓ.

In this circumstance, it is possible that $\sigma_2(\rho)$ might jump from a value greater than $\pi + \epsilon$ to a value that is less than $\pi - \epsilon$ as the transversal descends to x. This jump is harmless, since the covering problem for the upper and lower halves of P can be treated separately. An artificial vertex at $p_1(\rho)$ is added to both boundary portions, and the upper boundary is terminated at the artificial vertex w. The interpolation rule determines the angular portions to assign each subsegment of a partitioned edge. Lastly, a slicing ray is introduced at $p_1(\rho)$, along with an angular adjustment α that forces the ray at $p_1(\rho)$ to point toward $p_2(\rho)$. Vertex w also gets a slicing ray.

Since neither collection of edges will have a total rotational sum that exceeds π, Corollary 5a suffices to establish the coverage of each half.

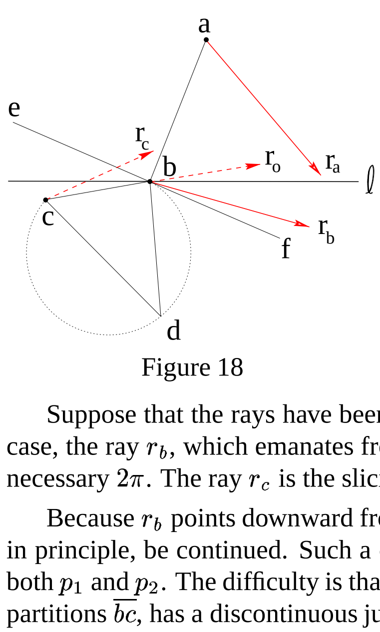

Figure 18

Second, there is a discontinuity that can occur from a failure in the interpolation rule. As shown in Figure 18, the transversal is horizontal. The point b is the current p_2, so that ℓ does not intersect P to the left of b. The rotational value β associated with edge $\overline{bc}$ is $\angle bdc$. The possible locations for vertex d as used in the interpolation procedure is shown as the circular arc $b\widehat{dc}$. Line $\overline{ef}$ is tangent to $b\widehat{dc}$ at b.

Suppose that the rays have been rotated so that the ray at p_1, which is not shown, points at p_2. In this case, the ray r_b, which emanates from p_2, has an associated angle $\theta_b = \sigma_2 + \pi$ that is a little less than the necessary 2π. The ray r_c is the slicing ray for vertex c, and has an associated angle equal to $\theta_b + \angle bdc$.

Because r_b points downward from p_2 (and the ideal direction is to point at p_1), the descent of ℓ should, in principle, be continued. Such a change would cause additional rotation to occur from the relocation of both p_1 and p_2. The difficulty is that in the illustration, the interpolated contribution from $\overline{bp_2(\rho)}$, once $\ell(\rho)$ partitions $\overline{bc}$, has a discontinuous jump.

It is easy to see that the interpolation rule will be jump-free at a vertex b if and only if the line that is parallel to ℓ and contains b also intersects the interior of arc $b\widehat{dc}$ or is tangent at b.

However, when this criterion fails, as illustrated in Figure 18, descending ℓ from b introduces a jump

θ_J in the interpolated angle, which equals the angle between the tangent $\overrightarrow{bf}$ and the rightward pointing transversal from b.[3] To resolve this case, it suffices to use the transversal through b. The upper boundary will have monotone slicing rays that begin with one pointing from p_1 toward p_2, and end with r_b, which has some kind of downward direction. Corollary 5a addresses this case.

As for the lower half, Corollary 5b will apply if r_c, which is the next-to-the-last slicing ray for the lower half of P (as viewed from a mirrored perspective for Corollary 5), intersects ℓ at or to the left of p_2 as illustrated. To prove that r_c satisfies this condition, let r_0 be collinear with $\overline{bc}$ and point upward from b. Then the angle between r_0 and the tangent $\overrightarrow{bf}$ is precisely $\angle bdc$. Failure will occur only if r_b lies within the cone bounded by $\overrightarrow{bf}$ and r_0, since r_b must point downward, and the smallest interpolated angle that can be constructed from $\angle bdc$ must be too large. Consequently, a rotation of r_b by θ_J will cause it to become rising, and a rotation by $\angle bdc$ will cause it to pass r_0. But the rotation by $\angle bdc$ gives a ray emanating from b that is parallel to r_c. Consequently, r_c has a direction that comprises a positive rotation of $\overrightarrow{cb}$ with respect to c, and a rotation of the falling r_b by an angle equal to $\angle bdc < \pi$. It follows that r_c rises to intersect ℓ at or to the left of b, which ensures that Corollary 5b is applicable.

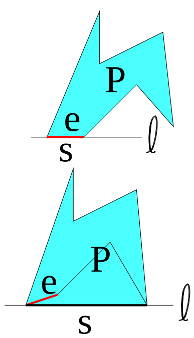

Figure 19

Lastly, suppose that some edge e has an associated angle of size π or more. We show that such an occurrence is harmless; in fact, e's slicing triangle is not needed to cover P^{rgn}.

Let s be a longest segment that belongs to the convex hull of P and either contains e or seals off e inside of the region enclosed by $P \cup s$. Let P be redefined to include the enclosed pocket, and let the angle associated with s equal the sum of the angles associated with the edges that no longer belong to P, including e. Thus, the associated angle is at least π. Let the extension of s be the transversal ℓ, so that the endpoints of s are p_1 and p_2.

Let the rays be rotated so that the ray from p_1 points at p_2. Then each slicing ray will intersect ℓ, and Corollary 5a shows that the coverage is complete. ∎

4 Minor extensions

For completeness, we give a reduction to show that Corollary 1b follows from Corollary 1a.

Corollary 1b. Let F be a simple polygon in the plane, and let S be a polygon that is inscribed in a circle, and which has edges with the same lengths as those of F. Suppose that S^{rgn} does not contain the center of its circumscribing circle.

Then the area of S is at least as large as that of F.

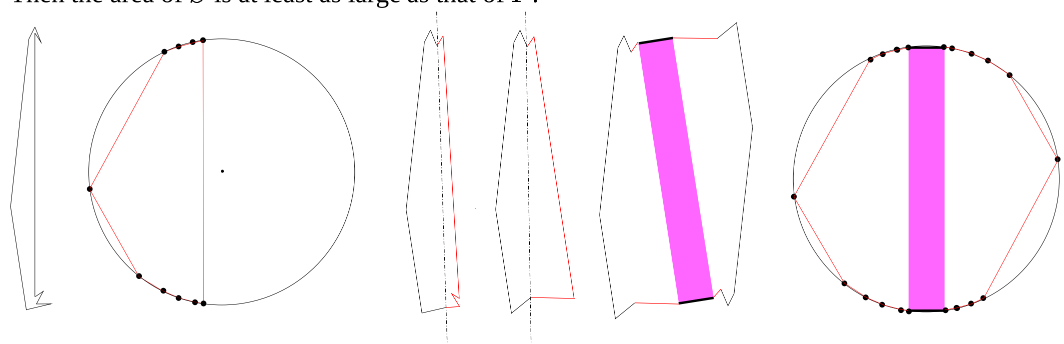
Figure 20

[3] Notice that c must lie below ℓ as otherwise the discontinuity must be of the first type associated with Figure 17. Because c lies below ℓ, segment $\overline{bc}$ induces no comparable jump at c. Similarly, the segment would not induce a jump at b if $\overrightarrow{bc}$ were to point down and to the right, rather than down and to the left, because the upper side of $\overline{bc}$ would then be designated as the inside.

Proof (Sketch):
One way to see that this is so to use reflection and other simple local improvements to expose the longest edge so that its extension as an infinite line will not have any additional intersections with P. Then a rotated copy of the improved figure can be combined with a (virtual) rectangular filler to build a new figure that will have the same size circumscribing circle as shown, and that meets the conditions of Corollary 1a. If P has n edges, then the composite figures will have $2n$ edges. In particular, each edge in P, apart from the largest, will have two comparable edges in the composite. The last two edges in the composite will represent the top and bottom edges of the filler rectangle. By construction, the filler rectangle in the modified polygon and the rectangle inscribed in the circle have side lengths equal to the length of the longest edge in P, and twice the distance from the longest edge of the inscribed polygon to the center of its circumcircle. ∎

Lastly, we conclude this section by outlining extensions of Theorem 2 to cases where the rays r_1 and r_n do not point at each other. While Corollary 5 addressed this case, it was somewhat weak.

Theorem 3. Let P be an n-gon with consecutive vertices $v_1, v_2, \ldots, v_n$ that sequence counterclockwise about P, and let $r_1, r_2, \ldots, r_n$ be a monotone slicing for P with associated angles $\theta_1, \theta_2, \ldots, \theta_n$, and associated slicing triangles $\{t_j\}_{j=1}^{n}$. Let v_1 and v_n be on the x axis with v_1 to the right of v_n, and suppose that r_1 and r_n intersect at a unique point p_1.

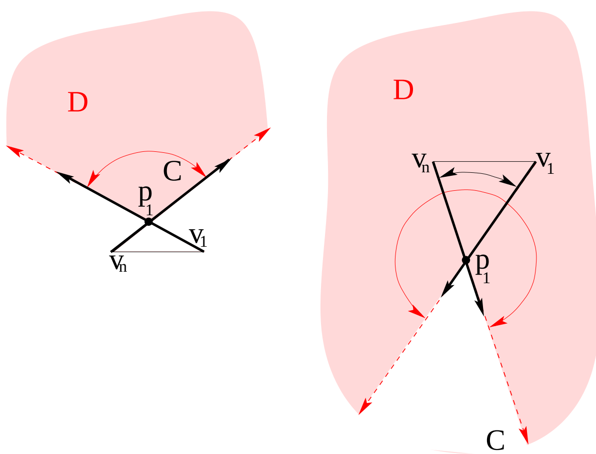

Figure 21

Define $\hat{r}_1$ to be the ray that emanates from p_1 and is parallel to (and coincides with) r_1. Let $\hat{r}_n$ be the ray that emanates from p_1 and is parallel to (and coincides with) r_n. Let C be the cone defined by the boundary $\hat{r}_1 \cup \hat{r}_n$.

a) If p_1 lies above the line $\overleftrightarrow{v_1 v_n}$, let $D = C$.

b) If p_1 lies below $\overleftrightarrow{v_1 v_n}$, let D comprise $(R^2 \setminus C) \cup \hat{r}_1 \cup \hat{r}_n$.

Then in cases 1) and 2):

$$P^{rgn} \cap D \subset \cup_{j=1}^{n-1} t_j.$$

Proof (Sketch):
The basic reasoning is the same as that of Theorem 2.

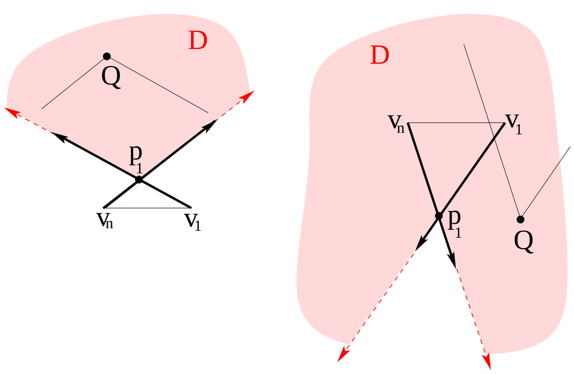

Figure 22

The only difference is that the horizontal line ℓ through Q is replaced by a pair of rays that emanate from Q and are anti-parallel to $\hat{r}_1$ and $\hat{r}_n$ as shown. ∎

For completeness, we observe that Theorem 3b facilitates a direct proof of Corollary 1b without any need to transform the problem into an instance of Corollary 1a. The proof begins with a convexification step as outined next.

Let P be oriented. If the figure is not convex, let $\ell = \overleftrightarrow{v_i v_j}$, be a global support line that intersects P at v_i and v_j, for $i \neq j$.

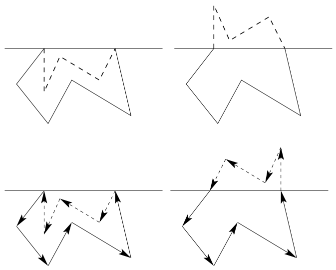

Figure 23

Suppose that $P \cap \overline{v_i v_j} = \{v_i, v_j\}$, so that $\overline{v_i v_i}$ seals off some pocket of P as shown. Steiner symmetrization flips (reflects) one of the subpaths terminating at v_i and v_j about ℓ to increase the area of the resulting figure while preserving the edge lengths. As a practical matter, it is simpler to reverse the sequencing of these edges, which effectively rotates the boundary portion about the midpoint of $\overline{v_i v_j}$ (but does not reverse their direction). While iterations of either operation lead, eventually, to a convex region with increased area, the virtual rotation operation is easier to analyze. It simply rearranges the segment ordering without changing their directions. Since this procedure can produce no more than $(n-1)!$ such arrangements, one of these polygons must have a maximal area. This polygon must be convex, since otherwise this procedure could further increase the area.

At this point, the mapping between S^{rgn} and the convexified P is straightforward. The radial rays of S should be transferred to the convexified P with a rotational, offset that forms t_n as an isosceles triangle exterior to the figure. It follows that

$$|P| \leq \sum_{j-1}^{n-1} |t_j| - |t_n|.$$

This construction resolves a question impicitly posed by Yaglom and Boltanskiĭabout the difficulty of finding a geometric proof of Corollary 1 that does not rely on an intermediate isoperimetric inequality [10, p. 56].

5 Conclusions

Theorem 1 was derived as a parallelized interpretation of a much simpler but less general divide-and-conquer proof of the most basic isoperimetric inequality as presented in [8]. Its topological perspective enables Corollary 1a to be proven directly without any convexifying steps or reductions to the isoperimetric inequality for the disk. Moreover, the proof can be established as an algorithmic construction in traditional Euclidean geometric algebra.

Construction. Let P be an n-gon, and let Q be a polygon that is inscribed in a circle and has sides that are congruent to those of P. Let the vertices of P be $v_1, v_2, \ldots, v_n$, and be in correspondence with $q_1, q_2, \ldots, q_n$ in Q. Let both vertex sets have the same orientation about their respective interiors. Let q_0 be the center of the circle circumscribing Q. Suppose that v_1 is a corner of the convex hull of P.

Form the slicing triangles $\hat{t}_j = (\triangle q_j q_0 q_{j+1})^{rgn}$, for $j = 1, 2, \ldots, n$. Draw a line through the points q_1 and q_0. Let the line exit Q^{rgn} at $q \in \overline{q_k q_{k+1}}$ for some vertex index k. The proof of Theorem 1 shows that a suitable transversal ℓ is either $\overrightarrow{v_1 v_k}$, $\overrightarrow{v_1 v_{k+1}}$, or $\overrightarrow{v_1 v}$, where v, if it exists, is located so that $\angle v_1 v v_k = \angle q_1 q q_k$, $\angle v_k v v_{k+1} = \angle q_k q q_{k+1}$ and $v \in (\triangle v_1 v_k v_{k+1})^{rgn}$.

In any case, the image of $\hat{t}_j$ is defined by $\overline{v_j v_{j+1}}$ and the intersection (if it exists) of the slicing rays r_j and r_{j+1}. Ray r_1 emanates from v_1 with an angular direction θ_1 that points it along ℓ. Ray r_j emanates from v_j and has the direction $\theta_j = \theta_{j-1} + \angle q_{j-1} q_0 q_j$, for $j = 2, 3, \ldots, n$.

Evidently, the construction is trivial, but its proof of correctness is somewhat subtle. On the other hand, each step of this topological approach is elementary. Yet the perspective gives a variety of related results that are sometimes not so obvious.

References

[1] Benson, R.V. Euclidean Geometry and Convexity, McGraw-Hill, Inc., 1966.

[2] Courant, Richard, and Herbert Robbins. What is Mathematics? Oxford University Press, 1960.

[3] Kazarinoff, N.D., Geometric Inequalities. Random House, 1961.

[4] Lax, P.D. A Short Path to the Shortest Path, *American Mathematical Monthly*, 158–159, Feb., 1995.

[5] Osserman, R. The isoperimetric inequality, Bulletin of the American Mathematical Society, 84 (1978), 1182–1238.

[6] Pappus, of Alexandria. La collection mathematique; oeuvre traduite pour la premiere fois du grec en francais par Paul Ver Eecke. Paris, Bruges, Desclee, de Brouwer et cie, 1933.

[7] Rademacher, H. and O. Toeplitz. The Enjoyment of Mathematics, Princeton University Press, 1994.

[8] Siegel, A. A naive proof of the 2-D Isoperimetric Theorem in Elementary Geometry. Submitted.

[9] Virgil (A.K.A. Publius Vergilius Maro). Aeneid I, 365–368.

[10] Yaglom, I.M., and V.G. Boltanskiĭ. Convex Figures, translated by P.J. Kelly and L.F. Walton, Holt, Rinehart and Winton, 1961.