Parabola Volume 37, Issue 2 (2001)

THE ISOPERIMETRIC PROBLEM

George Szekeres[1]

Amongst all closed curves with given length (or perimeter) L, which one encloses the largest area? This is called the isoperimetric problem, and has apparently occupied geometers from the earliest times of recorded civilization. Legend has it that Queen Dido, founder of the ancient city of Carthage, had acquired from the inhabitants "as much land as can be surrounded by the hide of an ox". With the deal accomplished, the Queen ordered the hide to be cut up into narrow strips, and by arranging them in a large semi-circular arc with the straight shore line as diameter she secured the largest possible area of land for her city. Dido's problem was thus to enclose the maximum area between a long straight line and a curved arc of given length, and this is closely related to the isoperimetric problem. (You will see from what follows that Dido gave the correct answer: the area is maximum when the arc is semi-circular).

We have already anticipated the solution of the isoperimetric problem: among all curves of given perimeter, the one which encloses the largest area is the circle. This is equivalent to the statement that among all figures of given area, the circle has the smallest perimeter. Since a circle with perimeter $L = 2\pi r$ has area

$$A_0 = \pi r^2 = \pi (L/2\pi)^2 = L^2/4\pi$$

both the above statements can be expressed by the inequality

$$A \leq L^2/4\pi$$

between the length L of a curve and the area A enclosed by it. This is called the isoperimetric inequality; the only curve for which equality occurs (i.e., for which $A = L^2/4\pi$) is the circle.

The following table gives the perimeters of various geometrical figures of unit area:

Circle	$2\sqrt{\pi} \approx 3.55$
Square	4.00
Semicircle	$(2+\pi)\sqrt{2/\pi} \approx 4.10$
Rectangle $(1:2)$	$3\sqrt{2} \approx 4.24$
Equilateral triangle	$6/\sqrt[4]{3} \approx 4.56$
Isosceles right triangle ...	$2(1+\sqrt{2}) \approx 4.84.$

[1]This is a reprint of an article which appeared in Parabola Vol.1 No.1, written by Professor Szekeres who turned 90 this year (see Vol.37 No.1, page 2).

Thus among simple geometrical figures of unit area the circle has clearly the smallest perimeter. However, the examination of special cases will never yield a general proof, even if the number of cases examined is vastly increased. Neither can we accept as conclusive proof the considerable amount of physical evidence, derived from observation of electrostatic phenomena, raindrops, soap bubbles, etc., which seems to confirm the isoperimetric property of the circle (or of the sphere in three dimensions).

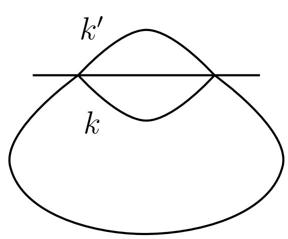
An elementary solution of the isoperimetric problem was first given by the famous Swiss geometer, Jacob Steiner (1796-1863) over 130 years ago. The idea of Steiner's proof was to increase the area of the given figure while keeping the perimeter fixed. He showed that such increase is always possible unless the figure is a circle, and concluded that the circle is the figure of given perimeter which encloses the largest area.

In the following we shall describe the essential steps in Steiner's proof; the interested reader will find a great deal of further material in Polya [1] (see references).

We first note that a convex closed curve is one with the property that if P, Q are two points inside the curve then the whole of the line segment PQ joining them also lies inside the curve. (Roughly, this means that the curve has no "bumps" inwards.)

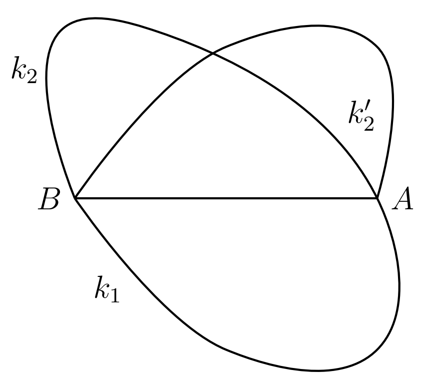
We start with a closed curve k of given length, and the steps of the argument are as follows:

1st Step. — Suppose that the curve k is not convex. Then it is intuitively clear from the diagram (and can be proved) that we can obtain a new curve k' which has the same length as k but encloses a larger area than k.



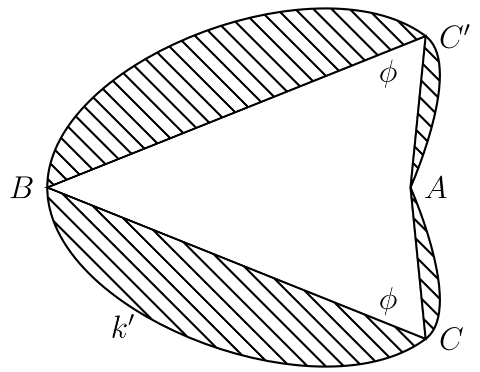
2nd Step. — Suppose now that the curve k is convex. Let A, B be points on k with the property that they divide k in two parts, k_1 and k_2, of equal length.

The existence of such a pair of points is intuitively evident and will be accepted here without proof. A correct proof requires careful mathematical analysis which is beyond the scope of the present exposition.



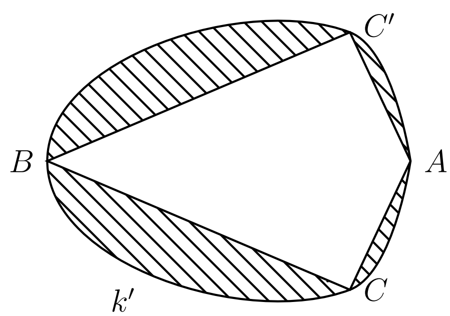
Now let A_1, A_2 be the areas between the diameter AB and k_1, k_2 respectively (so that $A_1 + A_2$ is the area enclosed by k) and suppose that $A_1 \geq A_2$. Replace k_2 by the arc k_2', obtained through reflection of k_1 with respect to the diameter AB. The new curve k', consisting of k_1 and k_2', has the same length as k but the included area is $2A_1 \geq A_1 + A_2$, so that the area has certainly not diminished.

3rd Step. — Let C be a point on k_1 and C' the corresponding (reflected) point on k_2'. Because of the symmetry of the figure, the interior angles of the quadrilateral $ACBC'$ at the vertices C and C' are equal; let this common angle be ϕ. Suppose that ϕ is not a right angle; then we can increase the area enclosed by k' (while the circumference is kept fixed) as follows:



imagine that $ACBC'$ is made up of four rigid rods freely jointed at the vertices A, C, B, C'. Let us pull the vertices A, B (if $\phi < 90°$) or C, C' (if $\phi > 90°$) apart, until the angle at C (and C') becomes a right angle. Imagine that the four shaded segments enclosed between k' and the rods AC, CB, BC', $C'A$ are carried along during this process as shown.

Then the new curve k'' has the same circumference as k' (since it consists of the same four pieces of arc), but its area is larger because the shaded positions are the same but the area of the triangle ACB in the new position is larger than in the original position (having the same sides AC, CB, but a right angle at C).



Thus we were able to increase the area provided that there existed a point C on k_1 with $\angle ABC \neq 90°$. But if $\phi = 90°$ for every C on k_1 then k_1 is a semicircle above the diameter AB, and k_1 is a circle. Hence the proof is complete.

Unfortunately Steiner's ingenious argument contains a well-concealed logical error, so well concealed indeed that it remained undetected for nearly half a century. The error was eventually discovered and rectified by the German mathematician Weierstrass, himself one of the founders of modern analysis. What Steiner actually did was to define a procedure by which the given curve k can be changed into one with the same length and larger included area, **provided that k was not a circle**. It follows from Steiner's argument that **if** the isoperimetric problem has a solution at all then the solution must be a circle; what he omitted to show was that a solution **does** in fact exist.

To illustrate the fallacy in Steiner's argument, consider a similar but quite nonsensical problem: among all positive integers determine the one which has the largest possible value. Here we know of course that a solution does not exist; nevertheless we can invent a procedure which will give a definite answer, by much the same type of argument as the one used by Steiner. Take the square of each number. By this procedure every positive integer is increased, except the number 1; for $1^2 = 1$ and $n^2 > n$ if $n \neq 1$. We conclude therefore that if there exists a largest integer then it must be equal to 1. This is undoubtedly a true statement, but our argument does not prove of course that 1 is actually the largest number.

References

[1] G. Polya Intuition and Analogy in Mathematics (Oxford University Press, 1954) chapter 8.

[2] R. Courant and H. Robbins, What is Mathematics (Oxford University Press, 1951) chapter VII, 8