ACTION, MINKOWSKI ADDITION, AND TWO PROOFS OF THE ISOPERIMETRIC INEQUALITY

JEREMY BOOHER

Of all figures in the plane with a given perimeter, which has the greatest area? The Greeks knew the answer to this question, although they could not prove it. Part of the problem was that the notion of perimeter and area are tricky notions, requiring analysis to deal with properly. A rigorous proof had to wait until the 19th century. There are now many proofs of the isoperimetric inequality. We will focus on the case where the boundary of the figure is a piecewise C^1 curve in order to sweep the analytic issues about area and perimeter under the rug. Remember that being piecewise C^1 means that the closed curve is continuously differentiable except at a finite number of points.

Theorem 1 (Isoperimetric Inequality)**.** *Let γ be a closed curve in the plane that is piecewise C^1. Let $A = \mathcal{A}(\mathrm{Int}(\gamma))$ be the area enclosed by the curve and $L = \mathcal{L}(\gamma)$ the perimeter. We have*

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

with equality holding if and only if the curve is a circle.

Remark 2. The boundary curve need not be piecewise C^1: with more analysis, less needs to be assumed about γ. However, this is more than sufficient to encompass common figures likes polygons and circles.

Here we will present two proofs of the isoperimetric inequality, one using the concept of the action of a curve and one using Minkowski addition of sets.

1. Hurwitz's Proof via Actions

Hurwitz's proof uses calculus. For this to make sense, the curve must be piecewise C^1 as we will assume throughout. Several useful facts are:

- If $\gamma(t) = (x(t), y(t))$ is a curve from $[a, b] \to \mathbb{R}^2$, its length can be calculated by the formula
$$\mathcal{L}(\gamma) = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt.$$
- If $\gamma(t)$ is a curve from $[0, \pi] \to \mathbb{R}^2$ with the polar coordinates of $\gamma(t)$ given by $r(t)$ and $\theta(t)$, the length of the curve is
$$\mathcal{L}(\gamma) = \int_0^\pi \sqrt{\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta}{dt}\right)^2} dt.$$
This follows by making the change of variable $x = r\cos(\theta)$ and $y = r\sin(\theta)$ in the first formula.

Date: July 28, 2010.

- With the notation above, the area inside the curve is given by
$$\mathcal{A}(\mathrm{Int}(\gamma)) = \frac{1}{2}\int_0^\pi r(t)^2 \frac{d\theta}{dt} dt.$$
- The Cauchy-Schwarz inequality for integrals says that for square integrable functions f and g from $[0,\pi] \to \mathbb{R}$ we have
$$\left(\int_0^\pi f(t)g(t)dt\right)^2 \leq \int_0^\pi f(t)^2 dt \int_0^\pi g(t)^2 dt.$$
Equality is achieved if and only if g is a multiple of f.

These are fairly standard facts from calculus: you should be able to find all of them in a good calculus book.

The concept of action appears in physics and allows a reformulation of classical mechanics in terms of action and Lagrangians. For our purposes, all we need is the following definition of the action of a curve, which should be thought of as a physical property of a particle following that trajectory.

Definition 3. Let $\gamma : [a,b] \to \mathbb{R}^2$ be a plane curve. The action of the path γ is defined to be

$$\mathcal{E}(\gamma) := \frac{1}{2}\int_a^b \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 dt = \frac{1}{2}\int_a^b \left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta}{dt}\right)^2 dt \tag{1}$$

Unlike the length of the curve, the action of the curve depends on the parametrization of the curve. Therefore, we need to standardize how to parametrize curves.

Definition 4. The standard parametrization of a curve γ is the one for which the parameter's interval is $[0,\pi]$.

Since the definitions of action and length are similar, it is no surprise that they are related.

Proposition 5. *Let γ be a curve with standard parametrization. Then*
$$\mathcal{L}(\gamma)^2 \leq 2\pi\mathcal{E}(\gamma)$$
with equality achieved if and only if the parameter is a multiple of arc length along γ.

Proof. By Cauchy-Schwarz,
$$\left(\int_0^\pi f(t)g(t)dt\right)^2 \leq \int_0^\pi f(t)^2 dt \int_0^\pi g(t)^2 dt.$$
Taking $f(t) = 1$ and $g(t) = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}$ gives the desired inequality. Equality is achieved if $g(t)$ is a constant, which says that the point $(x(t), y(t))$ moves along the curve with constant speed. Thus the parameter is a multiple of arc length. □

The relevance to the isoperimetric inequality is that the action is also related to the area the curve encloses.

Proposition 6. *Let γ be a simple closed curve with standard parametrization. Then*
$$2\mathcal{A}(\mathrm{Int}(\gamma)) \leq \mathcal{E}(\gamma)$$
with equality achieved if and only if γ is a circle.

The proof uses the following analytic lemma:

Lemma 7 (Wirtinger Inequality)**.** *Let $f : [0,\pi] \to \mathbb{R}$ be a piecewise C^1 function with $f(0) = f(\pi) = 0$. Then we have*

$$\int_0^\pi f'(t)^2 dt \geq \int_0^\pi f(t)^2 dt.$$

Equality is achieved if and only if $f(t)$ is a multiple of $\sin(t)$*.*

Proof. Consider the function $g(t) = \frac{f(t)}{\sin(t)}$. It is differentiable at the endpoints by l'Hopital's rule and the fact that $f(0) = f(\pi) = 0$. Then

$$f'(t) = g(t)\cos(t) + g'(t)\sin(t)$$

so

$$\int_0^\pi f'(t)^2 dt = \int_0^\pi \left(g(t)^2\cos(t)^2 + 2g(t)g'(t)\cos(t)\sin(t) + g'(t)^2\sin(t)^2\right)$$

Integrating the middle term by parts and using the boundary conditions gives

$$2\int_0^\pi g(t)g'(t)\cos(t)\sin(t) = -\int_0^\pi g(t)^2\left(\cos(t)^2 - \sin(t)^2\right)dt.$$

Thus

$$\int_0^\pi f'(t)^2 dt = \int_0^\pi \left(g(t)^2 + g'(t)^2\right)\sin(t)^2 dt = \int_0^\pi f(t)^2 dt + \int_0^\pi g'(t)^2\sin(t)^2 dt \geq \int_0^\pi f(t)^2 dt.$$

Equality is achieved if and only if $g'(t) = 0$ on the interval, which implies $g(t)$ is constant and $f(t)$ is a multiple of $\sin(t)$. □

We can now prove Proposition 6.

Proof. Remember that $\mathcal{A}(\mathrm{Int}(\gamma)) = \frac{1}{2}\int_0^\pi r(t)^2 \frac{d\theta}{dt} dt$. Then using (1) we have

$$\mathcal{E}(\gamma) - 2\mathcal{A}(\mathrm{Int}(\gamma)) = \frac{1}{2}\int_0^\pi \left(\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta}{dt}\right)^2 - 2r^2\frac{d\theta}{dt}\right)dt.$$

Completing the square turns the right side into

$$\frac{1}{2}\int_0^\pi r^2\left(\frac{d\theta}{dt} - 1\right)^2 dt + \frac{1}{2}\int_0^\pi \left(\left(\frac{dr}{dt}\right)^2 - r^2\right)dt.$$

The first term is non-negative since it is a square, and the second is non-negative by Wirtinger's inequality (the curve is closed, so we may pick the origin of the polar coordinate system to be at $\gamma(0) = \gamma(\pi)$). To have equality, $\frac{d\theta}{dt} = 1$ and $r(t)$ is a multiple of $\sin(t)$. Putting these together, $r = a\sin(\theta + \varphi)$. This is the polar equation of a circle. □

The isoperimetric equality now follows easily.

Proof. Pick a standard parametrization for γ. By the two propositions, we know that

$$\mathcal{A}(\mathrm{Int}(\gamma)) \leq \frac{1}{2}\mathcal{E}(\gamma) = \frac{1}{4\pi}\mathcal{L}(\gamma)^2.$$

Equality is obtained if and only if the curve is a circle. □

2. The Isoperimetric Inequality Through Minkowski Sums

The second proof uses the Minkowski sum of two sets. It avoids directly invoking calculus, although it is hidden in the approximation of general regions by polygons.

Definition 8. Let A and B be subsets of $\mathbb{R}^2$. The Minkowski sum is defined to be

$$A \boxplus B = \{x + y : x \in A, y \in B\}. \tag{2}$$

This is easiest to deal with for rectangles. By an open rectangle, I mean the interior of a rectangle in $\mathbb{R}^2$. All that we will deal with will have sides parallel to the coordinate axes, so are of the form $(a, b) \times (c, d)$.

2.1. **Steiner's Inequality.** Let Ω be a closed and bounded set with piecewise smooth C^1 boundary. Let B_r denote the circle of radius r centered at the origin. Define

$$\Omega_r = \Omega \boxplus B_r. \tag{3}$$

By definition, it is the set of points at most distance r from a point of Ω.

Theorem 9 (Steiner)**.** *With the notation above, the following inequality holds:*

$$\mathcal{A}(\Omega_r) \leq \mathcal{A}(\Omega) + Lr + \pi r^2$$

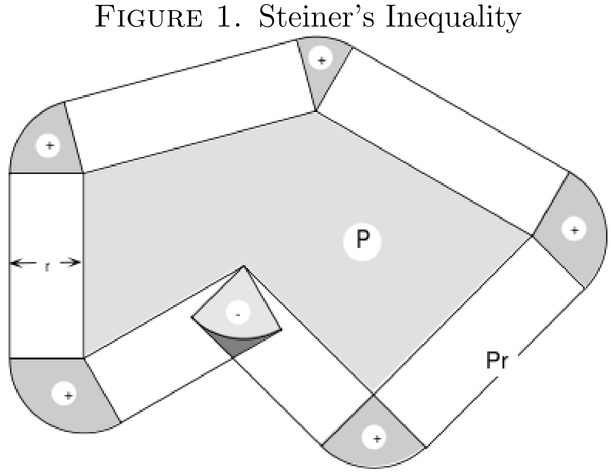
Equality happens if and only if Ω is convex.

Remark 10. The same proof also shows that

$$\mathcal{L}(\partial\Omega_r) \leq \mathcal{L}(\partial\Omega) + 2\pi r.$$

Proof. To Steiner's inequality, exhaust the inside of the Ω by polygons. If Ω is convex, these polygons may be taken to be convex. The limit of the area of the polygons approach the area of Ω.[1] If the inequalities in Steiner's theorem hold when extending these polygons by r, then taking the limit as the polygons better approximate Ω gives the desired inequalities.

Figure 1. Steiner's Inequality



Let P be a polygon with vertices v_i, and deconstruct P_r as shown in Figure 1. Let θ_i be the angle (in radians) of the circular sector based at v_i, and σ_i be $+1$ if the polygon is convex

[1] This is actually tricky. Look up Hausdorff Convergence in the appendix of "Inequalities that Imply the Isoperimetric Inequality".

at v_i, -1 if it is concave. Let L be the perimeter of P. The rectangles of height r added to each side have total area Lr, but may contribute less to P_r if they overlap. Two rectangles overlap if they are next to a common vertex v_i where P is concave. In this case the overlap is larger than the area of the circular wedge with angle θ_i, so the total contribution of the rectangles is at most $Lr - \sum_{i:\sigma(i)=-1} \frac{1}{2} r^2 \theta_i$.

Note that $2\pi = \sum \sigma_i v_i$. At any vertex where the polygon is convex, the area of the circular wedge is $\frac{1}{2} r^2 \theta_i$. Therefore the total area of the wedges is

$$\frac{1}{2} r^2 \sum_i \theta_i = \pi r^2 + \frac{1}{2} r^2 \sum_{i:\sigma(i)=-1} \theta_i.$$

The total area of P_r is at most $\mathcal{A}(P) + Lr + \pi r^2$. Note in the case that the polygon is convex this is an equality. □

2.2. The Brunn-Minkowski Theorem.

We now move on to the Brunn-Minkowski Theorem, which gives a lower bound on the area of $A \boxplus B$ in terms of the areas of A and B.

Theorem 11 (Brunn-Minkowski). *Let A and B be bounded sets with piecewise C^1 boundary. Then we have*

$$\sqrt{\mathcal{A}(A \boxplus B)} \geq \sqrt{\mathcal{A}(A)} + \sqrt{\mathcal{A}(B)}.$$

Remark 12. Equality happens if A and B are homeolithic: $A = rB \boxplus x$, for some scaling factor r and point x.

Similar to the proof of Steiner's Inequality, we will prove the statement when A and B are finite unions of open rectangles and then approximate general A and B by such rectangles.

Lemma 13. *Let A be the union of n pairwise disjoint open rectangles, and B be the union of m pairwise disjoint open rectangles. Then we have*

$$\sqrt{\mathcal{A}(A \boxplus B)} \geq \sqrt{\mathcal{A}(A)} + \sqrt{\mathcal{A}(B)}.$$

Proof. The proof proceeds by induction on $n + m$. For the base case, suppose $n + m = 2$ so that $A = (a, b) \times (c, d)$ and $B = (e, f) \times (g, h)$. Then

$$\begin{aligned}
\mathcal{A}(A \boxplus B) &= \mathcal{A}((a + e, b + f) \times (c + g, d + h)) \\
&= (b + f - a - e)(d + h - c - g) \\
&= (b - a)(d - c) + (f - e)(h - g) + (b - a)(h - g) + (f - e)(d - c) \\
&\geq (b - a)(d - c) + (f - e)(h - g) + 2\sqrt{(b - a)(h - g)(f - e)(d - c)} \\
&= \left(\sqrt{\mathcal{A}((a, b) \times (c, d))} + \sqrt{\mathcal{A}((e, f) \times (g, h))}\right)^2
\end{aligned}$$

where the second to last step uses that $x + y \geq 2\sqrt{xy}$, a form of the arithmetic-geometric mean inequality.

Now suppose $n + m = l > 2$, and the assertion holds for all pairs of sets with $n + m < l$. We may assume that $n \geq 2$. Pick two rectangles R_1 and R_2 that are part of A. Because R_1 and R_2 are disjoint and open, there is a horizontal or vertical line that does not intersect them that passes between them (draw a picture). Without loss of generality, assume it is the vertical line $x = x_1$. It divides the rectangles composing A up into two sets

$$A' = \{(x, y) \in A : x < x_1\} \quad \text{and} \quad A'' = \{(x, y) \in A : x > x_1\}.$$

The line cuts some rectangles of A into two parts, but does not split R_1 and R_2 since they are completely on one side of the line. Thus both A' and A'' are composed of less than n rectangles. Let

$$\theta = \frac{\mathcal{A}(A')}{\mathcal{A}(A)} \quad \text{and} \quad \Psi = \frac{\mathcal{A}(A')}{\mathcal{A}(A'')}$$

be the fraction of the area of A appearing in A' and the ratio of the areas of A' and A''.

By moving a vertical line $x = t$ through B, we see that because the area of B to the left of the line is a continuous function of t there will be a vertical line $x = x_2$ that cuts B into two sets, the ratio of whose areas is Ψ. Let

$$B' = \{(x, y) \in B : x < x_2\} \quad \text{and} \quad B'' = \{(x, y) \in B : x > x_2\}.$$

Note that B' and B'' are each the union of at most m pairwise disjoint open rectangles, that $\mathcal{A}(B) = \mathcal{A}(B') + \mathcal{A}(B'')$, and that $\theta = \dfrac{\mathcal{A}(B')}{\mathcal{A}(B)}$.

Now because A' and A'' are in A and B' and B'' are in B we know that

$$A' \boxplus B' \cup A'' \boxplus B'' \subset A \boxplus B.$$

Since $A' \boxplus B'$ is to the left of the line $x = x_1 + x_2$ and $A'' \boxplus B''$ is the right of the line $x = x_1 + x_2$, the two sets are disjoint. Since A' and A'' are composed of fewer than n rectangles and B' and B'' are composed of at most m rectangles, the inductive hypothesis applies to show that

$$\begin{aligned}
\mathcal{A}(A \boxplus B) &\geq \mathcal{A}(A' \boxplus B') + \mathcal{A}(A'' \boxplus B'') \\
&\geq \left(\sqrt{\mathcal{A}(A')} + \sqrt{\mathcal{A}(B')}\right)^2 + \left(\sqrt{\mathcal{A}(A'')} + \sqrt{\mathcal{A}(B'')}\right)^2 \\
&= \theta \left(\sqrt{\mathcal{A}(A)} + \sqrt{\mathcal{A}(B)}\right)^2 + (1-\theta) \left(\sqrt{\mathcal{A}(A)} + \sqrt{\mathcal{A}(B)}\right)^2 \\
&= \left(\sqrt{\mathcal{A}(A)} + \sqrt{\mathcal{A}(B)}\right)^2 .
\end{aligned}$$

□

We can now prove the general version of the Brunn-Minkowski theorem.

Proof. The idea is to construct a sequence of sets A_n and B_n that contain A and B, $\lim_{n\to\infty} \mathcal{A}(A_n) = \mathcal{A}(A)$ and $\lim_{n\to\infty} \mathcal{A}(B_n) = B$, and with A_n and B_n a finite union of open rectangles. Tessellate the plane with squares of side length $\frac{1}{n}$, and let A_n be the union of (the interiors) of all squares that intersect A. Since A is bounded, A_n is a finite union of open rectangles. Do the same to construct B_n. As $A = \cap A_n$ and $B = \cap B_n$, the areas of A_n and B_n approach the areas of A and B as n goes to infinity.[2] Using this and the

[2]Again, this is actually tricky. Look at properties of the Lebesgue measure and how it interacts with limits and intersections.

Brunn-Minkowski theorem for finite unions of rectangles,

$$\begin{aligned}\sqrt{\mathcal{A}(A \boxplus B)} &= \lim_{n\to\infty} \sqrt{\mathcal{A}(A_n \boxplus B_n)} \\ &\geq \lim_{n\to\infty} \sqrt{\mathcal{A}(A_n)} + \sqrt{\mathcal{A}(B_n)} \\ &= \sqrt{\mathcal{A}(A)} + \sqrt{\mathcal{A}(B)}\end{aligned}$$

which establishes the Brunn-Minkowski inequality. □

2.3. **The Isoperimetric Inequality.** The Brunn-Minkowski inequality along with Steiner's inequality immediately imply the isoperimetric inequality.

Proof. Let A be the area of a set Ω with piecewise C^1 boundary of length L. Let B_r be the disc of radius r. Then for any $r \geq 0$ using Steiner's inequality and the Brunn-Minkowski inequality

$$A + Lr + \pi r^2 \geq \mathcal{A}(\Omega \boxplus rB) \geq \left(\sqrt{A} + \sqrt{\pi r^2}\right)^2 = A + \pi r^2 + 2\sqrt{A\pi}r$$

Combining these and simplifying gives

$$L^2 \geq 4\pi A.$$ □

3. An Application of the Isoperimetric Inequality to Quadrilaterals

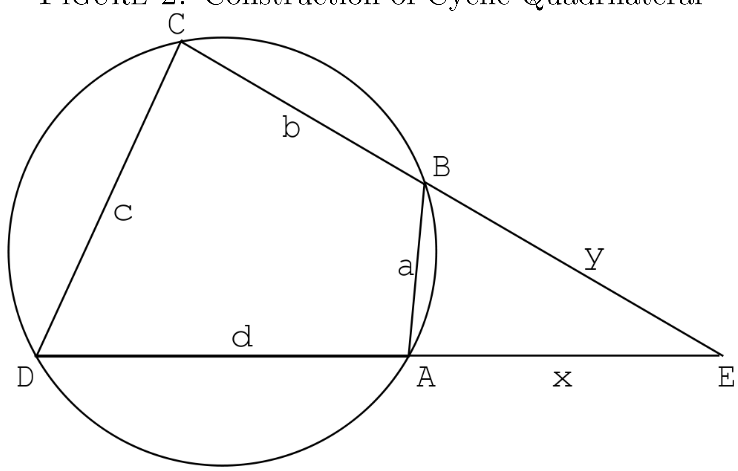
The isoperimetric inequality is quite useful. One easy application is to the problem of finding quadrilateral has the maximum area when the four sides are specified.

Given the four side lengths a, b, c, and d, a quadrilateral exists with these side lengths (in order) provided the obvious inequalities $a + b + c > d$, $b + c + d > a$, etc. are satisfied.

Corollary 14. *Given a, b, c, and d as above, the quadrilateral with these side lengths and maximum area is cyclic (all four vertices lie on a circle).*

Proof. It is a lesser-known geometric fact that there exists a cyclic quadrilateral with sides of length a, b, c, and d. If $a = c$ then a trapezoid works. Otherwise, suppose that such a quadrilateral existed. Extend the sides of length b and d to meet in a point E as shown in Figure 2. Let x and y be the lengths of the extensions. Because $\angle BCD$ and $\angle BAD$ are

Figure 2. Construction of Cyclic Quadrilateral



supplementary (they subtend opposite arcs), triangles $\triangle EAB$ and $\triangle ECD$ are similar. This gives algebraic formula for x and y in terms of a, b, c, and d.

To construct the cyclic quadrilateral, construct the triangle with side lengths a, x, and y, and then extend EA and EB to get the points C and D. Triangles $\triangle EAB$ and $\triangle ECD$ are then similar, which implies $\angle BCD$ and $\angle BAD$ are supplementary so the quadrilateral is cyclic.

The vertices of this cyclic quadrilateral divide the circle into four arcs. For any other quadrilateral Q with side lengths a, b, c, and d place these arcs on the sides of Q. This results in a piecewise C^1 curve γ with the same perimeter as the circle. The area of the cyclic quadrilateral is the area of the circle minus the area of the four circular sectors, while the area of Q is the area enclosed by γ minus the area of the same four circular sectors. The area enclosed by the curve is maximized when the curve is a circle, so the area of the quadrilateral is maximized when it is cyclic. $\square$

4. Acknowledgments and Sources

The proof using Wirtinger's Inequality is due to Hurwitz and comes from *Elementary Geometry* by John Roe. The second proof using Minkowski Addition comes from "Inequalities that Imply the Isoperimetric Inequality" by Andrejs Treibergs, available online.[3] Figure 1 and 2 are taken from it. It also includes extensive bibliographic references and some information about the analytic details about area and length that are lurking.

[3] `www.math.utah.ede/~treiberg/isoperim/isop.pdf`