Inequalities that Imply the Isoperimetric Inequality

Andrejs Treibergs
University of Utah

Abstract. The isoperimetric inequality says that the area of any region in the plane bounded by a curve of a fixed length can never exceed the area of a circle whose boundary has that length. Moreover, if some region has the same length and area as some circle, then it must be the circle. There are dozens of proofs. We give several arguments which depend on more primitive geometric and analytic inequalities.

"The cicrle is the most simple, and the most perfect figure. "
Poculus. *Commentary on the first book of Euclid's Elements.*
"Lo cerchio è perfettissima figura." Dante.

In these notes, I will present a few of my favorite proofs of the isoperimetric inequality. It is amusing and very instructive to see that many different ideas can be used to establish the same statement. I will concentrate on proofs based on more primitive inequalities. Several important proofs are omitted (using the calculus of variations, integral geometry or Steiner symmetrization), due to the fact that they have already been discussed by others at previous occasions [B], [BL], [S].

Among all closed curve of length L in the plane, how large can the enclosed area be? Since the regions are enclosed by the curves with the same perimeter, we are asking for the largest area amongst isoperiometric regions. Which curve (or curves) encloses the largest possible area? We could ask the dual question: Among all regions in the plane with prescribed area A, at least how long should the perimiter be? Since the regions all have the same area, we are asking for the smallest length amongst isopiphenic regions. Which figures realize the least perimeter? The answer to both questions is the circle. A way to formulate a statement equivalent to both is in the form of the Isoperimetric Inequality.

Theorem. Isoperimetric Inequality. *Among all regions in the plane, enclosed by a piecewise C^1 boundary curve, with area A and perimeter L,*

$$4\pi A \le L^2.$$

If equality holds, then the region is a circle.

One doesn't need to assume this much smoothness on the boundary for the isoperimetric inequality to hold, rectifiability suffices [B],[G]. This is for simplicity. The student may be more familiar with rigorous arguments under such hypotheses.

1. Some applications of the isoperimetric inequality.

This is a useful inequality. Following Polya[P] let's illustrate how it may be used to prove things that may be much trickier to prove using calculus alone.

The largest quadrilateral problem. Suppose four side lengths a, b, c, d are given. Which quadrilateral in the plane with these side lengths maximizes the area? The answer is the cyclic quadrilateral, the one whose vertices can be placed in order around a circle, *i.e.* the one inscribable in some circle.

First of all, can a quadrilateral whose side lengths are four numbers be consrtructed at all? The answer is yes provided that the side lengths satisfy the inequalities that require that the sum of the lengths of any three of the sides exceeds the length the remaining one. (See Lemma 1. of the Appendix.) One can also prove that there is a cyclic quadrilateral with side lengths a, b, c, d provided that the inequalities are strict [B]. (See Lemma 2. of the Appendix.)

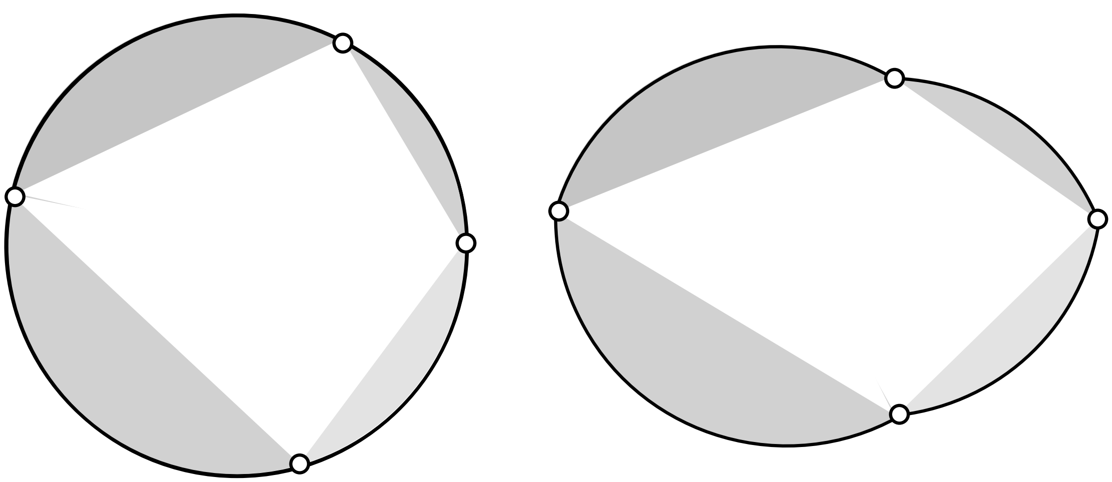


Fig. 1.

Suppose we wished to show that among quadrilaterals in the plane with given four side lengths a, b, c, d, that the cyclic quadrilateral has the largest area. This problem can be handled by calculus, but it may become messy depending on your formulation. We will show this fact independently of the isoperimetric inequality later in two ways, using trigonometry (Brahmagupta's Inequality) and as a consequence of Ptolemy's inequality. After all we will wish to deduce the isoperimetric inequality from these! But assuming the isoperimetric inequality gives an easiest argument that cyclic quadrilaterals have the largest area.

By Lemma 2. we can construct a cyclic quadrilateral with the given side lengths. Let $ABCD$ be its vertices in order on Z, the circle containing the vertices. Now imagine the regions between the sides of the quadrilateral and the circle are made of a rigid material such as a stiff plastic. Furthermore, imagine that the quadrilateral with its plastic flanges attached has hinges at the vertices, so that the entire figure can flex in the plane. The problem of finding the largest area enclosed amongst these flexings is the same as finding the largest area enclosed by the quadrilateral is the same as finding the largest area enclosed by the the circular arcs, because the additional plastic includes a fixed area. By the isoperimetric inequality, the circular figure has the largest area, thus the cyclic quadrilateral contains the largest area.

This argument applies to polygons with any number of sides. Hence, the polygon with given side length has largest area if and only if the polygon is inscribed in a circle.

2. Princess Dido's Problem. Here is another application we shall motivate by a tale from antiquity [P]. Princess Dido, daughter of a Tyrian king and future founder of Carthage purchased from the North African natives an amount of land along the coastline "not larger than what an oxhide can surround." She cut the oxhide into strips and made a very long string of length L. And then she faced the geometrical problem of finding the region of maximal area enclosed by a curve, given that she is allowed to use the shoreline as part of the region boundary. In the interior of the continent the answer would be the circle, but on the seashore the problem is different. Assuming that the seashore is a straight line (the x-axis), then the maximum area enclosed is a semicircle with its diameter on the shoreline and area $A = L^2/\pi$. This easily follows from the isoperimetric inequality. We regard the shoreline as a mirror and reflect the curve Γ whose points are $(x(s), y(s))$ to Γ' whose corresponding points are $(x(s), -y(s))$ on the other side of the axis. Then the composite curve $\Gamma \cup \Gamma'$ has total length $2L$. The maximal area enclosed is the circle. Centering the circle on the shoreline means that the upper and lower semicircles coincide under reflection, thus the lower semicircle is the maximal area region.

Exercise. Suppose Dido bargained to buy land along a cape (in $\{(x, y) : y \leq -c|x|\}$, for $c > 0$ some constant.) What now would be the shape of the largest land she could surround with a string of length L? (see [P].)

3. Ptolemy's Inequality, Complex Numbers and the Quadrilateral Inequality. We envision a theorem about the four vertices of a quadrilateral, but the statement holds for any four points.

Theorem. Ptolemy's Inequality. *Let $ABCD$ be four points in the plane. Let a, b, c, d be the lengths of the sides AB, BC, CD, DA and p, q be the lengths of the diagonals AC, BD, resp. Then*

$$pq \leq ac + bd. \tag{1}$$

If equality holds, then $ABCD$ is cyclic (its vertices lie in order on a circle) or it's contained in a line segment (the length of one side equals the sum of the other three.)

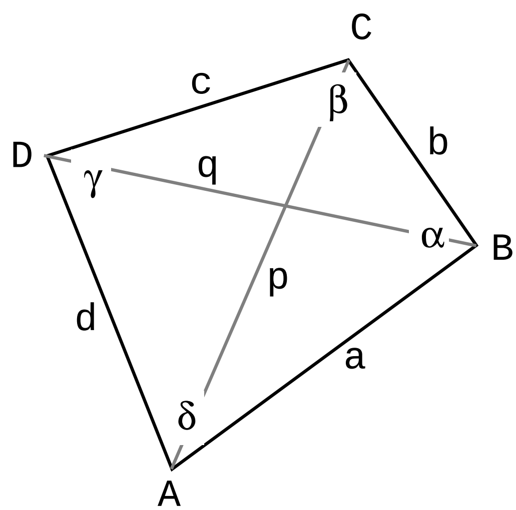


Fig. 2. Quadrilateral for Ptolemy's Theorem.

It may happen that the quadrilateral degenerates to a triangle or bigon or a point, all of which we regard as cyclic figures. If equality holds in any of these cases, then cyclicity implies that the figures are convex. Under equality, the quadrilateral may also degenerate to a line segment in which the length of one side equals the sum of the other three. Then the vertices are not contained in a circle, but since the proof uses projective transformations, the line may be regarded as a circle through infinity.

Proof. We consider two cases: that one of the points is distinct from the other three or not.

We use a little bit of complex arithmetic for the first case where A is distinct from all BCD. If the vertices $ABCD$ are assigned to complex numbers 0, z_1, z_2 and z_3 with $z_i \neq 0$ then $a = |z_1|$, $b = |z_1 - z_2|$, $c = |z_2 - z_3|$, $d = |z_3|$, $p = |z_2|$ and $q = |z_1 - z_3|$. Now we consider triangle inequality for the inversions

$$\left|\frac{1}{z_1} - \frac{1}{z_3}\right| \le \left|\frac{1}{z_1} - \frac{1}{z_2}\right| + \left|\frac{1}{z_2} - \frac{1}{z_3}\right|.$$

or

$$\frac{|z_1 - z_3|}{|z_1||z_3|} \le \frac{|z_1 - z_2|}{|z_1||z_2|} + \frac{|z_2 - z_3|}{|z_2||z_3|}.$$

Multiplying by $|z_1||z_2||z_3|$ gives

$$|z_2||z_1 - z_3| \le |z_3||z_1 - z_2| + |z_1||z_2 - z_3|$$

which is the same as (1). Equality in the triangle inequality holds if the points z_1^{-1}, z_2^{-1} , z_3^{-1}, ∞ lie (in order) on a line.

If that line contains the origin, then $ABCD$ lie on a line through the origin. Since z_2^{-1} lies between z_1^{-1} and z_3^{-1} then, depending on where 0 lies relative to the other points, the figure $ABCD$ has the property that one of the sides has a length equal to the sum of the other three. Unless the three points coincide $B = C = D$, the figure is not cyclic.

However, if the line does not pass through the origin, inversion maps it to a circle through the origin. Thus z_1, z_2 , z_3, 0 lie in order on a circle through the origin. It may happen that two points $B = C$ or $C = D$ coincide yielding a triangle. But, since the points occur in order around the circle, $B = D$ implies $B = C = D$ or the figure is a bigon.

The last case is if none of the points is distinct from the other three. If all points coincide $A = B = C = D$ and (1) is $0 = 0$. There three remaining degenerate subcases. If $A = B \neq C = D$ then $a = c$ and $b = d = p = q > 0$ and equality holds in (1). If $A = C \neq B = D$ then $p = q = 0$ and $a = b = c = d > 0$ and strict inequality holds trivially in (1). If $A = D \neq B = C$ then $b = d = 0$ and $a = c = p = q > 0$ and equality holds in (1). □

An isoperimetric inequality for quadrilaterals is based on a sharp upper bound for area of the quadrilateral. By the triangle inequality, the length of each side is less than the sum of the other three sides. Conversely, this condition on lengths is sufficient to construct a quadrilateral in the plane with those given side lengths. If one side length equals the sum of the other three sides then the quadrilateral degenerates to a line segment.

In order to avoid fussing about signed areas or worrying about cases in which opposite sides cross, we shall insist that by quadrilateral we mean a figure which is bounded by a simple (non-selfintersecting) polygonal curve consisting of four line segments connected end to end. Area then is the measure of the bounded region enclosed by the polygon. We allow some of the side lengths to be zero, thus triangles, bigons and points are degenerate examples. In any case, these degenrerate figures may be approximated by nondegenerate ones, so that both the length and area of a quadrilateral approach those of the limiting degenerate figure and so that the inequality for nondegenerate figures may be passed to the limit. The maximum property for cyclic quadrilaterals was first observed by Steiner.

Maximum Property for Cyclic Quadrilaterals. *Let $ABCD$ be the vertices of a quadrilateral in the plane. Let a, b, c, d be the lengths of the sides AB, BC, CD, DA, resp., and F the area. Then*

$$(2) \qquad 16F^2 \leq (a+b+c-d)(a+b-c+d)(a-b+c+d)(-a+b+c+d).$$

If equality holds, then $ABCD$ is cyclic (its vertices lie on a circle) or it's contained in a line segment (with one side as long as the sum of the other three lengths.)

Proof. A proof of this statement may be based on Calculus or on the observation that the maximum area of a triangle for which two sides are given occurs for the right triangle. (see Polya [P].)

We prefer to deduce it from Ptolemy's theorem. If the quadrilateral degenerates to a line segment, then F is zero and (2) holds. If the right side of (2) is also zero, then one of the terms is zero and one side is the sum of the other three.

If $ABCD$ is a convex quadrilateral, then let E be the intersection of AC and BD and p_1, q_1, p_2, q_2 the lengths of EC,ED,EA, EB resp. Note that $\theta = \angle CED = \angle AEB$ and $\pi - \theta = \angle DEA = \angle BEC$. Then the area of the quadrilateral is the sum of the areas of the triangles

$$\begin{aligned}2F =& p_1 q_1 \sin\theta + p_2 q_1 \sin(\pi-\theta) + p_2 q_2 \sin\theta + p_1 q_2 \sin(\pi - \theta)\\ =&(p_1q_1 + p_2q_1 + p_2q_2 + p_1q_2)\sin\theta\\ =&(p_1+p_2)(q_1+q_2)\sin\theta\\ =&pq\sin\theta.\end{aligned}$$

On the other hand, the lengths of the four sides are expressed by the cosine law

$$\begin{aligned}a^2 =& p_2^2 + q_2^2 - 2p_2q_2\cos\theta\\ b^2 =& p_1^2 + q_2^2 - 2p_1q_2\cos(\pi-\theta)\\ c^2 =& p_1^2 + q_1^2 - 2p_1q_1\cos\theta\\ d^2 =& p_2^2 + q_1^2 - 2p_2q_1\cos(\pi-\theta)\end{aligned}$$

Thus

$$\begin{aligned}a^2 - b^2 + c^2 - d^2 =& -2(p_2q_2 + p_1q_2 + p_1q_1 + p_2q_1)\cos\theta\\ =& -2(p_1+p_2)(q_1+q_2)\cos\theta \qquad (3)\\ =& -2pq\cos\theta\end{aligned}$$

Thus, rewriting the area, we get using Ptolemy's inequality, and (3)

$$\begin{aligned}16F^2 =& 4p^2q^2\sin^2\theta\\ =& 4p^2q^2 - 4p^2q^2\cos^2\theta\\ \le& 4(ac+bd)^2 - (a^2-b^2+c^2-d^2)^2\\ =& [2ac+2bd+a^2-b^2+c^2-d^2][2ac+2bd-a^2+b^2-c^2+d^2]\\ =& [(a+c)^2-(b-d)^2][(b+d)^2-(a-c)^2]\\ =& (a+b+c-d)(a-b+c+d)(a+b-c+d)(-a+b+c+d).\end{aligned}$$

thus (2) holds. If equality in (2) holds, then there is equality in Ptolemy's inequality and the points $ABCD$ lie in order on a circle. □

Note that if one of the sides of the quadrilateral degenerates (e.g. $d = 0$), then the resulting figure is a triangle, which is cyclic. It follows that (2) is equality, which is nothing more than *Heron's formula* for the area of a triangle:

$$4\,\mathrm{Area}(ABC) = \sqrt{(a+b+c)(a+b-c)(a-b+c)(-a+b+c)}.$$

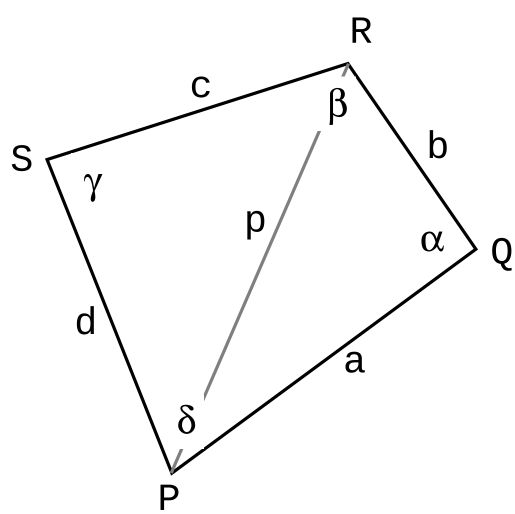
4. Brahmagupta Inequality.

Brahmagupta was an mathematician from Ujjain, India, living in the early seventh century. Brahmagupta was primarily concerned with number theory and integral solutions of equations. He found many integral triangles and quadrilarerals which motivated his generalization of Heron's formula for the area of a triangle in terms of its side lengths to the case of the quadrilateral. The basic formula he gave applied to cyclic quadrilaterals, those whose vertices occur in order around a circle. By giving a trigonometric proof, it is easy to adjust his formula for all quadrilaterals in the plane. An immediate corollary is the maximum property for quadrilaterals obtained by Steiner. In fact, Brahmagupta's formula gives an expression for the error term in Steiner's inequality. There are several equivalent formulations of this formula.

Brahmaguptas's Quadrilateral Formula. *Suppose a quadrilateral $PQRS$ is given in the plane whose side lengths are $a = PQ$, $b = QR$, $c = RS$, and $d = SP$ order around the quadrilateral and whose interior angles are $\alpha = \angle PQR$, $\beta = \angle QRSZ$, $\gamma = \angle RSP$, $\delta = \angle SPQ$ then the area of the quadrilateral F is given by*

$$16F^2 = (a+b+c-d)(a+b-c+d)(a-b+c+d)(-a+b+c+d) - 16abcd\cos^2\left(\frac{\alpha+\gamma}{2}\right)$$

Equality implies that that the quadrilateral is cyclic (the vertices occur in order around some circle) or that the vertices occur on a line segment such that one side length is the sum of the other three side lengths.



Quadrilateral used in the proof of Brahmagupta's Formula

Corollary. Maximum Property for Cyclic Quadrilaterals. *Among quadrilateral given in the plane with side lengths a, b, c, d (satisfying the triangle relations, each side length is less than the sum of the other three, e.g. $a \leq b+c+d$) the largest area is attained for a cyclic quadrilateral or a degenerate one whose verticies lie on a line such that one side length is the sum of the other three.*

Proof. Take the angles as in Brahmagupta's formula. Since the angles may be taken so $0 \le \alpha + \gamma = 2\pi - \beta - \delta \le 2\pi$ in Brahmagupta' formula, the max occurs when $\alpha + \gamma = \pi$ so $\beta + \delta = \pi$ or one of the sides, say $d = 0$. If a side has length zero, the figure degenerates to a triangle, biangle or point, which are all cyclic quadrilaterals. If all lengths are nonzero and one of the angles, say $\alpha = 0$ is zero, then the other $\gamma = \pi$ and S is an interior point of the line segment RP. It follows that the four points are collinear and $a = b+c+d$ or $b = c+d+a$. Similarly if $\alpha = \pi$ then $\gamma = 0$ and so either $a+b+c = d$ or $d+a+b = c$. Let $p = PR$. If $0 < \alpha < \pi$ then the chord PR of the circle through PQR subtends an angle 2α. Similarly, the chord PR of the circle through RSP subtends an angle 2γ. But since $2\alpha + 2\gamma = 2\pi$ the circles agree. and the two triangles PQR and RSP are on opposite sides of the chord PR since the subtended angles are on opposite sides of the circle. If $\alpha = \gamma = \pi/2$ then In other words, the quadrilateral is cyclic. ◻

Proof of Brahmagupata's formula. The idea is that the area of a quadrilateral is the sum of the areas of the triangles on opposite sides of a diagonal PR Thus

$$2F = 2\,\mathrm{Area}(PQR) + 2\,\mathrm{Area}(RSP) = ab\sin\alpha + cd\sin\gamma.$$

Squaring,

$$\begin{aligned}16F^2 =& 4(ab\sin\alpha + cd\sin\gamma)^2\\ =& 4(a^2b^2\sin^2\alpha + 2abcd\sin\alpha\sin\gamma + c^2d^2\sin^2\gamma)\\ =& 4(a^2b^2 - a^2b^2\cos^2\alpha + 2abcd\sin\alpha\sin\gamma + c^2d^2 - c^2d^2\cos^2\gamma\\ &- 2abcd\cos\alpha\cos\gamma + 2abcd\cos\alpha\cos\gamma)\\ =& 4a^2b^2 + 4c^2d^2 - (2ab\cos\alpha - 2cd\cos\gamma)^2 - 8abcd\cos(\alpha+\gamma)\,.\end{aligned} \tag{4}$$

Using the fact that the diagonal $p = PR$ can be expressed using the cosine law from either triangle,

$$p^2 = a^2 + b^2 - 2ab\cos\alpha = c^2 + d^2 - 2cd\cos\gamma,$$

we get

$$2ab\cos\alpha - 2cd\cos\gamma = a^2 + b^2 - c^2 - d^2. \tag{5}$$

The double angle formula implies the identity

$$\cos(\alpha+\gamma) = 2\cos^2\left(\tfrac{\alpha+\gamma}{2}\right) - 1 \tag{6}$$

Substituting (5) and (6) into (4) yields

$$
\begin{aligned}
16F^2 &= 4a^2b^2 + 4c^2d^2 - \left(a^2+b^2-c^2-d^2\right)^2 - 16abcd\cos^2\left(\tfrac{\alpha+\gamma}{2}\right) + 8abcd \\
&= (2ab+2cd)^2 - \left(a^2+b^2-c^2-d^2\right)^2 - 16abcd\cos^2\left(\tfrac{\alpha+\gamma}{2}\right) \\
&= (2ab+2cd+a^2+b^2-c^2-d^2)(2ab+2cd-a^2-b^2+c^2+d^2) \\
&\quad - 16abcd\cos^2\left(\tfrac{\alpha+\gamma}{2}\right) \\
&= \left((a+b)^2-(c-d)^2\right)\left((c+d)^2-(a-b)^2\right) - 16abcd\cos^2\left(\tfrac{\alpha+\gamma}{2}\right) \\
&= (a+b+c-d)(a+b-c+d)(a-b+c+d)(-a+b+c+d) \\
&\quad - 16abcd\cos^2\left(\tfrac{\alpha+\gamma}{2}\right)
\end{aligned}
$$

as to be proved. □

Exercise. Give another proof of Brahmagupta's inequality based on Ptolemy's inequality.

Mechanical Argument. Consider the following mechanical "proof" of the quadrilateral inequality. Imagine that the quadrilateral is allowed to flex in three space. Attach the ends of four sufficiently long but equally long new edges to the existing vertices and bind the four opposite ends together to make a new vertex. Assume that the whole structure is allowed to flex at the vertices in three space. There is still a degree of freedom but the quadrilateral is no longer necessarily planar. However, if the whole apparatus is dropped onto a planar surface (table top), with the new vertex held above the quadrilateral base, it settles to a planar quadrilateral, with the four new edges forming a pyramid above the base. Orthogonally projecting the four new edges to the table top give four equally long segments in the plane of the quadrilateral connecting at an interior vertex, showing that the mechanically stable pyramid configuration has a cyclic quadrilateral base. One could build a model by tying drinking straws together with string.

Exercise (Solution Unknown to the author!) Show that there is only one physically stable configuration (one with the lowest center of gravity) and this one has the largest area of the base. Perhaps one should look among pyramids that have the property that the new vertex must have a projection inside the projection of the base.

5. Steiner's Four Hinge "Proof".

Jakob Steiner(1796–1863), a self made Swiss farmer's son and contemporary of Gauss was the foremost "synthetic geometer." He hated the use of algebra and analysis and distrusted figures and once wrote "Calculating replaces thinking while geometry stimulates it"[St]. He proposed several arguments to prove that the circle is the largest figure with given boundary length. One very important method is Steiner symmetrization [B]. The other is his four-hinge method that has great intuitive appeal, but is limited to two dimensions.

If the curve bounding the region is not a circle, then there are four points on the boundary which are not cyclic. Consider the quadrilateral with these vertices, and regard as the rest of the domain rigidly attached to the sides of the quadrilateral as before. Suppose that the boundary can be articulated at these points. Then flexing the quadrilateral so to move the points to a circle results in a larger area by the quadrilateral inequality, with the same boundary length. Call this the "Four Hinge Maneuver." Steiner argued that for noncircular figures, the area could be increased, and therefore the circle has the largest area.

The fact that this is not a proof because there is no demonstration of the existence of a maximizing figure was pointed out to Steiner by Dirichlet. Weierstraß eventually gave the first rigorous demonstration. If we denote the application of a Four Hinge Maneuver to a noncircular domain Ω bounded by a piecewise C^1 curve by $F(\Omega)$, then we can envision a sequence $\Omega_0 = \Omega$, $\Omega_{n+1} = F(\Omega_n)$, for $n = 0, 1, 2, \ldots$ such that $\mathrm{Area}(\Omega_{n+1}) > \mathrm{Area}(\Omega_n)$ and $\mathrm{L}(\partial\Omega_{n+1}) = \mathrm{L}(\Omega_n)$. If it could be proved that for a (subsequence taken from a) sequence of Four Hinge Maneuvers, $\Omega_n \to B$ an $n \to \infty$ where B is a closed unit disk, using the semicontinuity (see Appendix), then the proof of the isoperimetric inequality would be given by

$$\begin{aligned}\mathrm{Area}(\Omega) < \lim_{n\to\infty} \mathrm{Area}\,\Omega_n = \mathrm{Area}\left(\lim_{n\to\infty}\Omega_n\right) = \mathrm{Area}(B),\\ \mathrm{L}(\partial\Omega) = \liminf_{n\to\infty} \mathrm{L}(\partial\Omega_n) \geq \mathrm{L}\left(\partial\left(\lim_{n\to\infty}\Omega_n\right)\right) = \mathrm{L}(\partial B).\end{aligned}$$

Here area is in the sense of Lebesgue measure, length is in the sense of one dimensional Hausdorff measure (mass of a rectifiable 1-current) and convergence of domains is in the sense of Hausdorff distance (see the definitions and discussion in the Appendix about Hausdorff convergence and the continuity of areas and lengths.) This program has been carried out for Steiner Symmetrization.

One of the issues is whether a Four Hinge Maneuver sequence can be found that actually forces the domains to converge to the circle. Consider the following suggestive reasoning. Starting with the domain Ω with sufficiently nice boundary, choose a point $X_0 \in \partial\Omega$, and some $\varepsilon > 0$. Starting from X_0, go both directions around $\partial\Omega$ and mark off points $X'_1, X'_2, \ldots, X'_m$ going one direction $X''_1, X''_2, \ldots, X''_m$ in the other whose $\mathbf{R}^2$ distances apart are ε, *i.e.*

$$\begin{aligned}\varepsilon &= \mathrm{dist}(X'_m, X'_{m-1}) = \mathrm{dist}(X'_{m-1}, X'_{m-2}) = \cdots\\ &= \mathrm{dist}(X'_1, X_0) = \mathrm{dist}(X_0, X''_1) = \cdots = \mathrm{dist}(X''_{m-1}, X''_m).\end{aligned}$$

Then construct $m-1$ quadrilaterals $Q_k = X'_k X'_{k+1} X''_{k+1} X''_k$ and apply the Four Hinge Maneuver $m-1$ times based on the Q_k's. Since the two sides

$$\varepsilon = \mathrm{dist}(X'_k, X'_{k+1}) = \mathrm{dist}(X''_k, X''_{k+1})$$

the result of the maneuver is a trapezoid whose sides $F(X'_{k+1}X''_{k+1})$ and $F(X'_k X''_k)$ are parallel an have a common perpendicular bisector. Lining up $F^{m-1}(\Omega)$ on this bisector results in a domain that is nearly symmetric about its axis, has the same boundary length as Ω but has larger area. If ε is small, then the polygon length approximates the length of $\partial\Omega$ and so is Hausdorff close to the polygon which is the convex hull of the new vertices, which is symmetric along a line. Now ε can be made to decrease and the choice of starting point, hence direction of symmetry, can be made arbitrarily, thus the sequence of Four Hinge Maneuvers converges to a domain such that every line is a line of symmetry, *i.e.* a circle. One expects that a proof of convergence of such a sequence will follow the lines of the proofs of convergence of Steiner Symmetrization sequences, although the author doesn't know if this has been done. (The circle in the figure is a radius ε-circle.)

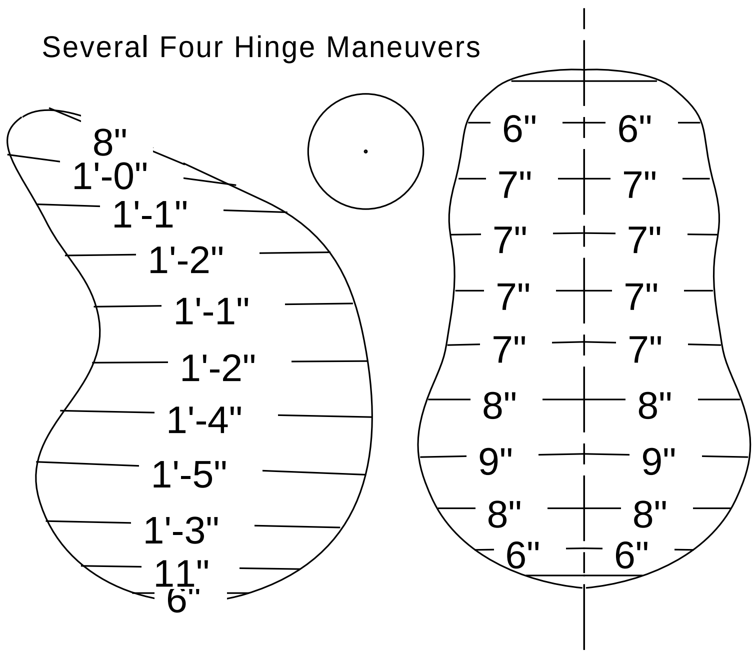


Fig. 3. Left region is converted by Four Hinge Maneuvers to the nearly symmetric one on right.

6. Hurwitz's proof using the Wirtinger inequality.

The Wirtinger inequality bounds the L^2 norm of a function by the L^2 norm of its derivative. In more general settings, the inequality is also known as the *Poincaré Inequality.* [GT,PS]. The best constant in the Poincaré Inequality is known as the first eigenvalue of the Laplace operator, and has been the inspiration of much geometric study (see, e.g. [LT].) We state stronger hypotheses than necessary.

Theorem. Wirtinger's inequality. *Let $f(\theta)$ be a piecewise $C^1(\mathbf{R})$ function with period 2π (for all θ, $f(\theta+2\pi)=f(\theta)$). Let $\bar{f}$ denote the mean falue of f*

$$\bar{f} = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)\,d\theta.$$

Then

$$\int_0^{2\pi} \left(f(\theta)-\bar{f}\right)^2\,d\theta \le \int_0^{2\pi} \left(f'(\theta)\right)^2\,d\theta.$$

Equality holds if and only if

$$f(\theta) = \bar{f} + a\cos\theta + b\sin\theta \tag{7}$$

for some constants a, b.

Proof. The idea is to express f and f' in Fourier series. Thus, since the derivative is bounded and f is continuous, the Fourier series converges at all θ [W]

$$f(\theta) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \{a_k \cos k\theta + b_k \sin k\theta\}$$

where the Fourier coefficients are determined by formally multiplying by $\sin m\theta$ or $\cos m\theta$ and integrating to get

$$a_m = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\cos m\theta\, d\theta, \qquad b_m = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\sin m\theta\, d\theta,$$

hence $2\bar{f} = a_0$. Since the sines and cosines are complete, the Parseval equation holds

$$\int_0^{2\pi} \left(f - \bar{f}\right)^2 = \pi \sum_{k=1}^{\infty} \left(a_k^2 + b_k^2\right). \tag{8}$$

Formally, this is the integral of the square of the series, where after multiplying out and integrating, terms like $\int \cos m\theta \sin k\theta = 0$ or $\int \cos m\theta \cos k\theta = 0$ if $m \neq k$ drop out and terms like $\int \sin^2 k\theta = \pi$ contribute π to the sum. The Fourier Series for the derivative is given by

$$f'(\theta) \sim \sum_{k=1}^{\infty} \{-k a_k \sin k\theta + k b_k \cos k\theta\}$$

Since f' is square integrable, Bessel's inequality gives

$$\pi \sum_{k=1}^{\infty} k^2 \left(a_k^2 + b_k^2\right) \le \int_0^{2\pi} (f')^2. \tag{9}$$

Wirtingers inequality is deduced form (8) and (9) since

$$\int_0^{2\pi} (f')^2 - \int_0^{2\pi} \left(f - \bar{f}\right)^2 \ge \pi \sum_{k=2}^{\infty} (k^2 - 1)\left(a_k^2 + b_k^2\right) \ge 0.$$

Equality implies that for $k \ge 2$, $(k^2 - 1)\left(a_k^2 + b_k^2\right) = 0$ so $a_k = b_k = 0$, thus f takes the form (7). □

Recall Green's theorem. If p and q are differentiable functions on the plane and Γ is a piecwise C^1 curve bounding the region Ω then

$$\oint_\Gamma p\,dx + q\,dy = \iint_\Omega (q_x - p_y)\;dx\,dy.$$

If we take $q = x$ and $p = 0$ then Green's theorem says

$$\oint_\Gamma x\,dy = \text{Area}(\Omega). \tag{10}$$

The same formula can be used to make sense of area even for curves that are merely rectifiable, namely, those whose length is the limit of lengths of approximating polygonal curves. (see e.g. [B].)

Hurwitz's proof of the Isoperimetric Inequality. We suppose that the boundary curve has length L is parameterized by arclength, thus given by two piecewise C^1 and L periodic functions $x(s)$, $y(s)$ that satisfy

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 = 1.$$

We convert to 2π periodic functions

$$f(\theta) = x\left(\frac{L\theta}{2\pi}\right), \qquad g(\theta) = y\left(\frac{L\theta}{2\pi}\right)$$

so that writing "$\,'\,$" $= d/d\theta$ gives

$$(f')^2 + (g')^2 = \frac{L^2}{4\pi^2}. \tag{11}$$

We now simply to estimate the area integral (10). Using $\int g'\,d\theta = 0$, the Wirtinger inequality and (11),

$$\begin{aligned} 2A =& 2\int_0^{2\pi} fg'\,d\theta = 2\int_0^{2\pi} (f - \bar f)g'\,d\theta \\ =& \int_0^{2\pi} (f - \bar f)^2 + (g')^2 - (f - \bar f - g')^2\,d\theta \\ \le& \int_0^{2\pi} (f')^2 + (g')^2\,d\theta = \int_0^{2\pi} \frac{L^2}{4\pi^2}\,d\theta = \frac{L^2}{2\pi}, \end{aligned}$$

which is the isoperimetric inequality. Equality forces equality in the Wirtinger inequality so

$$f(\theta) = \bar f + a\cos\theta + b\sin\theta$$

for some constants a, b. Equality also forces the dropped term to vanish

$$\int_0^{2\pi} (f - \bar{f} - g')^2 \, d\theta = 0$$

so that

$$g' = f - \bar{f}.$$

Hence

$$g(\theta) = \bar{g} + a \sin\theta - b \cos\theta.$$

Hence (11) implies

$$a^2 + b^2 = \frac{L^2}{4\pi^2}$$

so $(x(s), y(s))$ is a circle of radius $L/2\pi$. □

7. Steiner's Inequality for Parallel Sets.

Let $\Omega \subset \mathbf{R}^2$ be a closed bounded region with piecewise C^1 boundary. For $\rho \geq 0$ the outer ρ-parallel set is

$$\Omega_\rho = \Omega \boxplus \rho B = \{x \in \mathbf{E}^2 : \operatorname{dist}(x, \Omega) \leq \rho\}$$

the set of points in the plane whose distance to A is at most ρ. (B is the closed unit ball.)

Theorem. Steiner's Inequality. *Let $\Omega \in \mathbf{E}^2$ be a closed and bounded set with piecewise C^1 boundary whose area is A and whose boundary has length L. Let $\rho \geq 0$. Then the ρ-parallel set satisfies the inequalities*

$$\begin{gathered} \operatorname{Area}(\Omega_\rho) \leq A + L\rho + \pi\rho^2, \\ \mathrm{L}(\partial\Omega_\rho) \leq L + 2\pi\rho. \end{gathered}$$

If Ω is convex, then the inequalities are equalities.

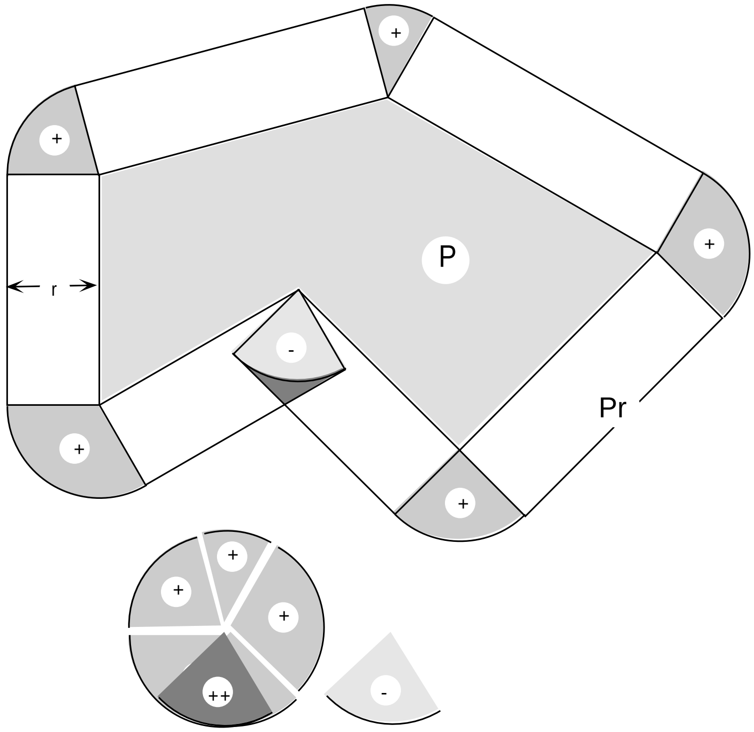


Fig. 4. Parallel Set for a polygon.

Proof. Let $P_n \subset \Omega$ be a sequence of polygons (convex if Ω is convex) converging to Ω whose boundary curves approach the boundary of Ω so that $\mathrm{Area}(P_n) \to A$ and $\mathrm{L}(\partial P_n) \to L$ as $n \to \infty$. Since Minkowski addition is continuous with respect to Hausdorff convergence (see [H]) then $(P_n)_\rho \to \Omega_\rho$ and furthermore $\mathrm{Area}((P_n)_\rho) \to A$ and $\mathrm{L}(\partial P_n) \to L$ as $n \to \infty$. Thus if we can prove the Steiner inequality for polygons then the general one follows. But for polygons, the Steiner inequality is almost obvious (see Figure 4.) It is proved in the appendix. □

8. Brunn's Inequality and Minkowski's Proof.

The *Minkowski Addition* of two arbitrary sets $S, T \subset \mathbf{R}^2$ in the plane is defined to be

$$aS \boxplus bT := \{x + y : x \in aS \text{ and } y \in bT\}$$

where a, b are nonnegative numbers and aS is the dilation by factor a

$$aS = \{ax : x \in S\}.$$

For example, the Minkowski sum of the rectangles is a rectangle

$$r([0,a] \times [0,b]) \boxplus s([0,c] \times [0,d]) = [0, ra + sc] \times [0, rb + sd]. \tag{12}$$

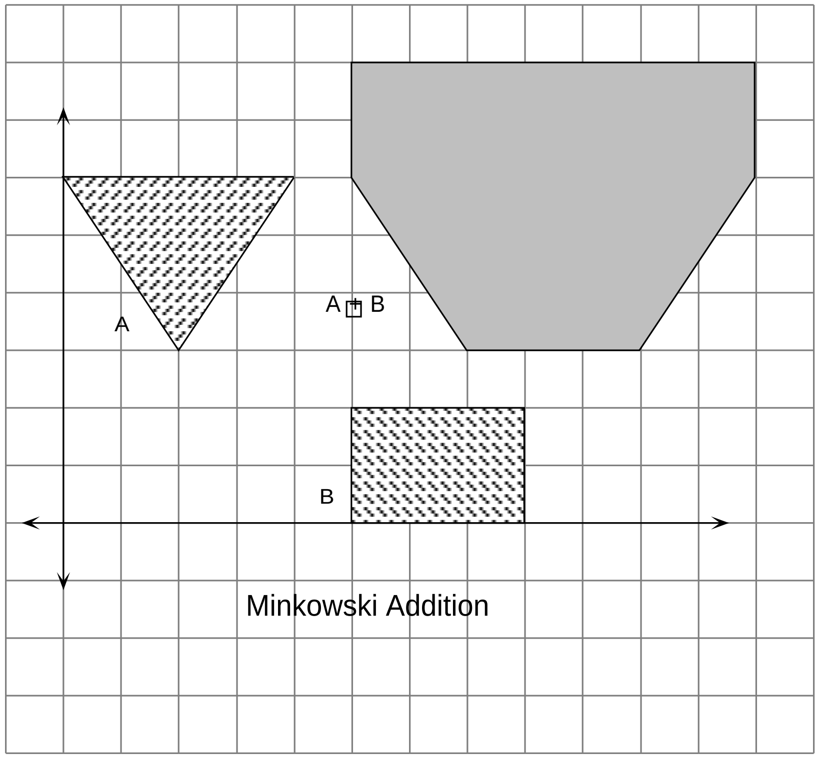


Fig. 5. The Minkowski Sum of a triangle and a rectangle.

The theorem of Brunn-Minkowski says that since the Minkowski addition tends to "round out" the figures being added, the area of the added figure exceeds the area of the summands.

Theorem. Brunn's Inequality. *Let $A, B \in \mathbf{R}^2$ be arbitrary bounded measurable sets in the plane. Then,*

$$\sqrt{\text{Area}(A \boxplus B)} \geq \sqrt{\text{Area}(A)} + \sqrt{\text{Area}(B)}. \tag{13}$$

Minkowski proved that equality holds if and only if A and B are homothetic. Two figures A and B are *homothetic* if and only if they are similar and are similarly situated, which means there is a point x and $r \geq 0$ so that $A = rB \boxplus \{x\}$.

Proof. We choose to present the proof of Minkowski, given in [F], using induction. The inequality is proved for finite unions of rectangles first and then a limiting process gives the general statement. Suppose that $A = \cup_{i=1}^{n} R_i$ and $B = \cup_{j=1}^{m} S_j$ where R_i and S_j are pairwise disjoint open rectangles, that is $R_i \cap R_j = \emptyset$ and $S_i \cap S_j = \emptyset$ if $i \neq j$. The proof is based on induction on $\ell = m + n$. For $\ell = 2$ there are two rectangles as in (12). The area

$$
\begin{aligned}
\operatorname{Area}((a,b)\times(c,d)\boxplus(e,f)\times(g,h)) &= \operatorname{Area}((a+e,b+f)\times(c+g,d+h))\\
&= (b-a+f-e)(d-c+h-g)\\
&= (b-a)(d-c)+(f-e)(h-g)+(b-a)(h-g)+(f-e)(d-c)\\
&\geq (b-a)(d-c)+(f-e)(h-g)+2\sqrt{(b-a)(h-g)(f-e)(d-c)} \qquad (14)\\
&= \left(\sqrt{(b-a)(d-c)}+\sqrt{(f-e)(h-g)}\right)^2\\
&= \left(\sqrt{\operatorname{Area}((a,b)\times(c,d))}+\sqrt{\operatorname{Area}((e,f)\times(g,h))}\right)^2,
\end{aligned}
$$

where we have used the Arithmetic-Geometric Mean Inequality

$$\frac{|X|+|Y|}{2}-\sqrt{|X||Y|}=\frac{1}{2}\left(\sqrt{|X|}-\sqrt{|Y|}\right)^2\geq 0.$$

Equality in the Arithmetic-Geometric Mean Inequality implies $|X|=|Y|$, so that equality in (14) implies $(b-a)(h-g)=(f-e)(d-c)$ or both rectangles have the same ratio of height to width. In other words, R and S are homothetic.

Now assume the induction hypothesis: suppose that (13) holds for $A=\cup_{i=1}^{n}R_i$ and $B=\cup_{j=1}^{m}S_j$ with $m+n\leq \ell-1$. For A and B so that $m+n=\ell$, we may arrange that $n\geq 2$. Then some vertical or horizontal plane, say $x=x_1$, can be placed between two rectangles. Let $R_i'=R_i\cap\{(x,y):x<x_1\}$ and $R_i''=R_i\cap\{(x,y):x>x_1\}$ and put $A'=\cup_i R_i'$ and $A''=\cup_i R_i''$. By choice of the plane, the number of nonempty rectangles in $\#A'<n$ and $\#A''<n$, but both A' and A'' are nonempty. Select a second plane $x=x_2$ and set $S_i'=S_i\cap\{(x,y):x<x_2\}$ and $S_i''=S_i\cap\{(x,y):x>x_2\}$ and put $B'=\cup_i S_i'$ and $B''=\cup_i S_i''$. Note that $\#B'\leq m$ and $\#B''\leq m$. x_2 can be chosen so that the area fraction is preserved

$$\theta=\frac{\operatorname{Area}(A')}{\operatorname{Area}(A')+\operatorname{Area}(A'')}=\frac{\operatorname{Area}(B')}{\operatorname{Area}(B')+\operatorname{Area}(B'')}.$$

By definiton of Minkowski sum, $A\boxplus B\supset A'\boxplus B'\cup A''\boxplus B''$. Furthermore, observe that $A'\boxplus B'$ is to the left and $A''\boxplus B''$ is to the right of the plane $x=x_1+x_2$, so they are disjoint sets. Now we may use the additivity of area and the induction hypothesis on $A'\boxplus B'$ and $A''\boxplus B''$.

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

Thus the induction step is complete.

Finally every compact region can be realized as the intersection of a decreasing sequence of open sets $A_n \supset A_{n+1}$ so that $A = \cap_n A_n$. A_n can be taken as the interiors of a union of finitely many closed squares. For each $\varepsilon = 2^{-n} > 0$ consider the closed squares in the grid of side ε which meet the set. Then the interior of the union of these squares is A_n. Removing the edges of the squares along gridlines A'_n results in a set with the same area. The result follows since Lebesgue measure of the limit is limit of the Lebesgue measure for decreasing sequences. Since the Minkowski sum of a decreasing set of opens is itself a decreasing set of opens, it follows that

$$\begin{aligned}\sqrt{\operatorname{Area}(A \boxplus B)} &= \lim_{n\to\infty} \sqrt{\operatorname{Area}(A_n \boxplus B_n)} \geq \lim_{n\to\infty} \sqrt{\operatorname{Area}(A'_n \boxplus B'_n)} \\ &\geq \lim_{n\to\infty} \left(\sqrt{\operatorname{Area}(A'_n)} + \sqrt{\operatorname{Area}(B'_n)}\right) = \sqrt{\operatorname{Area}(A)} + \sqrt{\operatorname{Area}(B)}\end{aligned}$$

and we are done. □

Brunn's inequality can be proved by other means.

Exercise. Give a proof of Brunn's inequality using the Wirtinger inequality. Hint: Suppose $K \subset \mathbf{R}^2$ is strictly convex with C^2 boundary and contains a neighborhood of the origin. For each $\theta \in [0, 2\pi)$ let $N(\theta) = (\cos\theta, \sin\theta)$ be the outward normal vector and $h(\theta)$ be the support function in the θ direction. That is, the line given by $N(\theta) \cdot X = h(\theta)$ is tangent to K so $h(\theta)$ is the distance from the origin to the supporting line. Show that if the curve is parameterized by θ then the arclength is $ds = (h'' + h)\, d\theta$ so the curvature is $\kappa = 1/(h + h'')$, the area is $2A = \int_0^{2\pi} h(\theta)\, ds = \int_0^{2\pi} h^2 + hh''\, d\theta = \int_0^{2\pi} h^2 - (h')^2\, d\theta$ and that the length is $L = \int_{\partial K} ds = \int_0^{2\pi} h + h''\, d\theta = \int_0^{2\pi} h\, d\theta$.

Minkowski's proof of the isoperimetric inequality using the Brunn Inequality. The result follows immediately from the Brunn inequality and Steiner's inequality. Thus, taking A to be the area of a set Ω with piecewise C^1 boundary and B the closed unit ball, we obtain for any $r \geq 0$

$$\begin{aligned}A + Lr + \pi r^2 &\geq \operatorname{Area}(\Omega_r) = \operatorname{Area}(\Omega \boxplus rB) \geq \left(\sqrt{\operatorname{Area}(\Omega)} + \sqrt{\operatorname{Area}(rB)}\right)^2 \\ &\geq \left(\sqrt{A} + r\sqrt{\pi}\right)^2 = A + r\sqrt{4\pi A} + \pi r^2,\end{aligned}$$

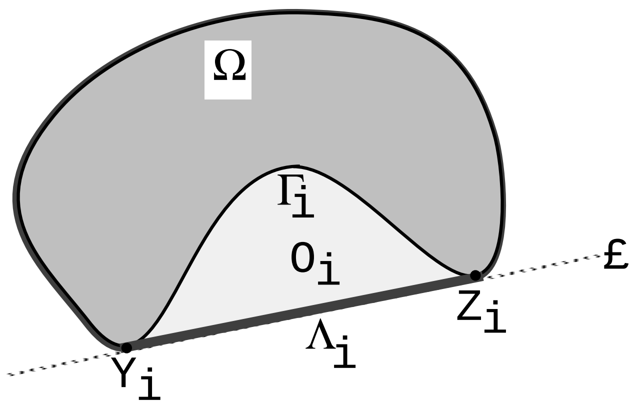
as desired. □

Exercise. Give another proof of the isoperimetric inequality using Wirtinger's inequality and the hint from the previous exercise. (This was Hurwitz's first proof.)

9. Convex Hulls and why the Isoperimetric Inequality need only be proved for convex sets. For if not, then we can construct a new curve, the convex hull of the figure, whose boundary has smaller length but encloses strictly more area, and thus has stricly greater isoperimetric ratio. The convex hull is defined to be the smallest convex set enclosing the figure. Thus if Ω is the closed region bounded by the curve, then

$$\hat{\Omega} = \bigcap \{K' : K' \text{ is convex and } \Omega \subset K'.\} .$$

If K is convex then $K = \hat{K}$. The key fact is that taking the convex hull increases the area and decreases the boundary length. Thus it suffices to prove the isoperimetric inequality for convex domains.



Convex Hull.

Convexity Lemma. *Let $\Omega \subset \mathbf{R}^2$ be a closed and bounded set surrounded piecewise C^1 Jordan boundary curve. Let $\hat{\Omega}$ be the convex hull. Then*

$$\text{Area}(\hat{\Omega}) \geq \text{Area}(\Omega) \qquad \textit{and} \qquad \text{L}(\partial(\hat{\Omega})) \leq \text{L}(\partial\Omega).$$

If Ω is nonconvex then both of the inequalities are strict.

Proof. Let $K = \hat{\Omega}$ be the convex hull. K must have interior points or else it is contained in a line as is $\partial\Omega$ which therefore can't be a $1-1$ image of a circle as a Jordan curve must be. The interior points of K which are not in Ω are denoted by the (at most countable since O_i are open) disjoint union

$$\amalg_i O_i = K^\circ - \Omega,$$

where O_i are the open connected components. If Ω is convex, this is the empty set. Every point X in the boundary ∂K is either a point of $\partial\Omega$ or not. If not, by a theorem of Caratheodory, it is a point of a line $\pounds \subset \mathbf{R}^2$ such that K is on one side of the line $\pounds$ and there are points $Y_i, Z_i \in \partial\Omega \cap \pounds$ such that X is strictly between Y_i and Z_i. Let O_i be the "lagoon" region between $\pounds$ and Ω. Since $X \notin \partial\Omega$ then O_i contains part of a ball neighborhood of X and thus it has positive area. It follows that $\text{Area}(K) = \text{Area}(\Omega) + \sum_i \text{Area}(O_i) > \text{Area}(\Omega)$. On the other hand, we cut the boundary curve of Ω into two pieces at the points Y_i and Y_i. Call Γ_i the part of $\partial\Omega$ from Y_i to Y_i along ∂O_i. Call $\Lambda_i = \partial O_i - \Gamma_i$ the line segment from Y_i to Z_i. Since Λ_i is a straigt line in the plane from Y_i to Z_i, its length $L(\Lambda_i) < L(\Gamma_i)$. On the other hand the boundary of $K = \Omega \cup \left(\cup_i \overline{O_i}\right)$ so that, by semicontinuity (see Appendix,) $L(\partial K) \le L(\partial\Omega) + \sum_i (L(\Lambda_i) - L(\Gamma_i)) \le L(\partial\Omega)$. □

It should be remarked that if a set has various components with various boundary pieces, then its $\mathcal{IR} = 4\pi A/L^2$ can be increased. Suppose that the $K = K_1 \cup K_2$ so that $\partial K = \partial K_1 \cup \partial K_2$. Then $L(\partial K) = L_1 + L_2$ where $L_i = L(K_i)$. Similarly $A = A_1 + A_2$. However, if both $A_1 > 0$,

$$\begin{aligned}\frac{\mathcal{IR}}{4\pi} = \frac{A_1 + A_2}{(L_1 + L_2)^2} &= \frac{L_1^2}{(L_1 + L_2)^2}\frac{A_1}{L_1^2} + \frac{L_2^2}{(L_1 + L_2)^2}\frac{A_2}{L_2^2} \\ &< \max\left\{\frac{A_1}{L_1^2}, \frac{A_2}{L_2^2}\right\} = \max\left\{\frac{\mathcal{IR}_1}{4\pi}, \frac{\mathcal{IR}_2}{4\pi}\right\}.\end{aligned}$$

Thus we do better by taking one or other part of a nontrivial partition. Finally, if K is not simply connected, then the isoperimetric ratio is improved by filling in the area and throwing out the boundary of any holes. Thus to find the extreme figure for the isoperimetric ratio, it suffices to consider convex domains K with interior.

10. Hadwiger's Proof using only Steiner's Inequality.

The proof of Hadwiger depends only on Steiner's inequality. For this purpose we need to define the *incircle*, *incenter* and *inradius* of a compact set $\Omega \subset \mathbf{E}^2$. The inradius is

$$r_I = \sup\{r \ge 0 : \text{there is } x \in \mathbf{E}^2 \text{ such that } x \boxplus rB \subset \Omega\}$$

An incenter is any such x_I so that the incircle $x_I \boxplus r_I B \subset \Omega$.

Isoperimetric Inequality of Hadwiger. *Suppose Ω is a compact set with piecewise C^1 boundary of area $\mathcal{A}$ and boundary length $\mathcal{L}$. Let M be a line through the incenter X_I of the convex hull K of Ω and $a = \mathrm{L}(K \cap \Omega)$ the length of a chord passing through the center. Then*

$$\mathcal{L}^2 - 4\pi\mathcal{A} \ge \frac{\pi^2}{4}\left(a - 2r_I\right)^2.$$

Observe that equality in the isoperimetric inequality implies that all chords through the incenter of K agree with the diameter of the incircle, hence K is a circle.

Proof. [G] Let A be the area of K and L the length of its boundary. Then by the convexity lemma, $\mathcal{L}^2 - 4\pi\mathcal{A} \ge L^2 - 4\pi A$ so we may estimte the deficit of the convex hull from below. Choose $R > 0$ so that the interior of $B_R(X_I) \supset K$. Let Υ be the annular region between the ball and K, $\Upsilon = B_R(X_I) - K^\circ$ and let F and G be the closures of the two halves of the annulus split by the line $\Upsilon - M$. Let P and Q denote the disjoint line segments $P \cup Q = M \cap \Upsilon$. Let ℓ_1 and ℓ_2 be the boundary components of ∂K in F and G respectively, and L_1 and L_2 be the two components of boundary $\partial F \cap \partial B_R(X_I)$ and $\partial G \cap \partial B_R(X_I)$ resp. We have

$$L = \ell_1 + \ell_2,$$
$$\pi R^2 = \text{Area}(F) + \text{Area}(G) + A.$$

Choose a number $2r_I \le 2r \le a$. Consider the parallel neighborhood of the ring domain Υ_r. Since $r \ge r_I$ the entire interior region is covered so $\text{Area}(\Upsilon_r) = \pi(R+r)^2$. If p and q denote the lengths of P and Q then $\text{Area}(P_r) = 2pr + \pi r^2$ and $\text{Area}(Q_r) = 2qr + \pi r^2$.

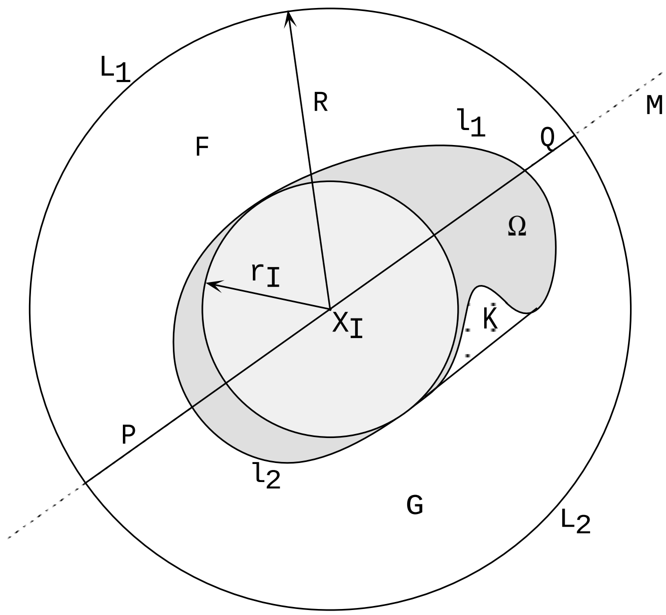


Fig. 6. Diagram for Hadwiger's proof.

Finally we apply Steiner's Inequality to F_r and G_r

$$\text{Area}(F_r) \le \text{Area}(F) + (p + L_1 + q + \pi R)r + \pi r^2,$$
$$\text{Area}(G_r) \le \text{Area}(G) + (p + L_2 + q + \pi R)r + \pi r^2.$$

Observe that any point of Υ_r is either a point of F_r or G_r. Also any point of P_r or Q_r are in both F_r and G_r. Hence

$$\text{Area}(\Upsilon_r) + \text{Area}(P_r) + \text{Area}(Q_r) \le \text{Area}(F_r) + \text{Area}(G_r)$$

which implies

$$\pi(R+r)^2 + 2(p+q)r + 2\pi r^2 \le \text{Area}(F) + \text{Area}(G) + (2p + L_1 + L_2 + 2q + 2\pi R)r + 2\pi r^2$$

or

$$A + \pi r^2 \leq Lr$$

which is equivalent to

$$L^2 - 4\pi A \geq (L - 2\pi r)^2.$$

Sunstituting $r = r_I$ and $r = a/2$ implies

$$L^2 - 4\pi A \geq \frac{1}{2}(L - 2\pi r_I)^2 + \frac{1}{2}(L - \pi a)^2 \geq \frac{\pi^2}{4}(a - 2r_I)^2,$$

using $2A^2 + 2B^2 \geq (A - B)^2$. □

11. Knothe's Proof and Crofton's formula for curves.

This is one of the directest proofs, it only requires one of the most basic formulas from integral geometry. (See [S] for a thorough exposition.) From the Convexity Lemma, it suffices to prove the isoperimetric inequality for convex domains K with interior points in the plane. Let K have area $A > 0$. By translating, we may suppose $0 \in K$. For simplicity, we assume that the boundary curve Γ of K is piecewise C^1, is parameterized by arclength and has length L. Thus there are L-periodic functions $(x(s), y(s)) : [0, L] \to \mathbf{R}^2$ that circumscribe the boundary.

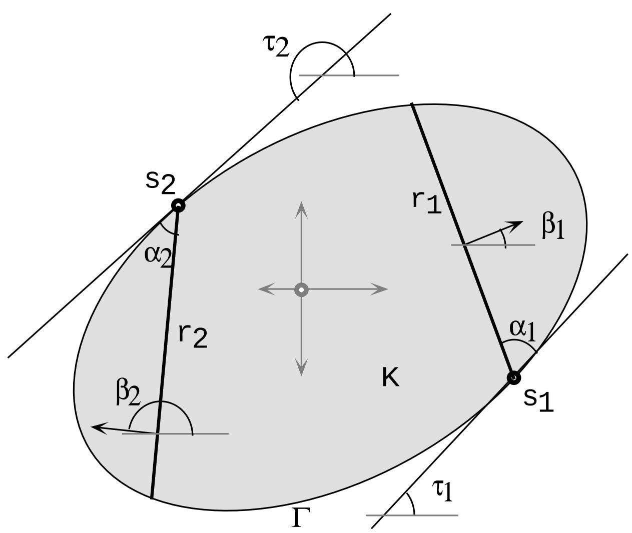


Fig. 7. Diagram for Knothe's Proof.

An unoriented line $L(\beta, p)$ in the plane may be parameterized by the angle $\beta \in [0, 2\pi)$ its normal makes with the x-axis and the distance $p \geq 0$ to the origin. Thus the line $L(\beta, p)$ has a normal vector $(\cos\beta, \sin\beta)$ and its equation is given by

$$x\cos\beta + y\sin\beta = p.$$

The kinematic measure on the space of lines, invariant under rigid motions of the plane, is $dL = dp \wedge d\beta$. Knothe's proof will amount to an integration over all pairs of lines that meet K. So let $\mathcal{L} = \{L : K \cap L \neq \emptyset\}$ denote the space of lines that meet K. The essence of intergral geometry is to describe the space of lines in two ways and derive different but equal expressions of the integral for these ways, utilizing the change of variables formula for the integration. The first way is $\mathcal{L} = \{L(p,\beta) : \beta \in [0, 2\pi),\ p \in [0, P(\beta)]\}$ where $P(\beta) \geq 0$ is the largest value of p so that the lines $L(p, \beta)$ meet K. A simple formula, due to Crofton, for the integral of chord lengths of lines meeting K is

$$\begin{aligned}\int_{\mathcal{L}} r(L)\, dL &= \int_0^{2\pi} \int_0^{P(\beta)} r(p,\beta)\, dp \wedge d\beta \\ &= \int_0^{\pi} \int_{-P(\beta+\pi)}^{P(\beta)} r(p,\beta)\, dp \wedge d\beta = \int_0^{\pi} A\, d\beta = \pi A\end{aligned}$$

where the length of the intersection is $r = \mathrm{L}(K \cap L)$.

The other way to describe a line is by a point $(x(s), y(s)) \in \Gamma$ and by the angle between the tangent vector to Γ and the line $\alpha = \angle((x'(s), y'(s)), L)$. Thus $\mathcal{L} = \{L(s,\alpha) : s \in [0, L), \alpha \in [0,\pi)\}$. To express dL in the (r, α) variables, it is convenient to let the tangent vector have angle τ so that

$$x'(s) = \cos\tau, \qquad y'(s) = \sin\tau, \qquad x(s)\cos\beta + y(s)\sin\beta = p, \qquad \tau + \alpha = \beta + \frac{\pi}{2}.$$

To compute the Jacobean for the change of variables $(s, \alpha) \to (p, \beta)$ we take exterior derivatives and exterior multiply the forms. Thus

$$\begin{aligned}(x\sin\beta - y\cos\beta)\, d\beta + dp &= (x' \ \cos\beta + y'\sin\beta)\, ds, \\ d\beta &= \frac{d\tau}{ds} ds + d\alpha.\end{aligned}$$

Wedging yields

$$\begin{aligned}dp \wedge d\beta &= (x' \ \cos\beta + y'\sin\beta)\, ds \wedge d\alpha \\ &= (\cos\tau\cos\beta + \sin\tau\sin\beta)\, ds \wedge d\alpha \\ &= \cos(\tau - \beta)\, ds \wedge d\alpha \\ &= \cos\left(\frac{\pi}{2} - \alpha\right)\, ds \wedge d\alpha \\ &= \sin\alpha\, ds \wedge d\alpha.\end{aligned}$$

Thus

$$2\pi A = \int_{\mathcal{L}} n r(L)\, dL = \int_0^{\pi} \int_0^{L} r(s,\alpha) \sin\alpha\, ds \wedge d\alpha \tag{15}$$

where n is the number of intersection points in $L \cap \Gamma$. $n = 2$ for almost every line, *i.e.* except for a set of lines of measure zero, because there are generically two points (s, α) corresponding to almost every $L \in \mathcal{L}$.

The second formula is for area in polar coordinates of a convex domain.

$$A = \int_0^\pi \int_0^{r(s,\alpha)} r\,dr\,d\alpha = \frac{1}{2}\int_0^\pi r(s,\alpha)^2\,d\alpha. \tag{16}$$

Knothe's Proof of the Isoperimetric Inequality. We compute using (15) and (16),

$$\begin{aligned}
0 \le & \int_\Gamma \int_\Gamma \int_0^\pi \int_0^\pi (r_1 \sin\alpha_2 - r_2 \sin\alpha_1)^2 \; d\alpha_1 d\alpha_2 ds_1 ds_2 \\
=& 2\int_\Gamma \int_0^\pi \int_0^\pi \int_\Gamma r_1^2 \sin^2\alpha_2 \, ds_1 d\alpha_1 d\alpha_2 ds_2 \\
& - 2\int_\Gamma \int_\Gamma \int_0^\pi \int_0^\pi r_1 \sin\alpha_1 r_2 \sin\alpha_2 \, d\alpha_1 d\alpha_2 ds_1 ds_2 \\
=& 4A \int_\Gamma \int_0^\pi \int_\Gamma \sin^2\alpha_2 \, ds_1 d\alpha_2 ds_2 - 2\left(\int_\Gamma \int_0^\pi r_1 \sin\alpha_1 \, d\alpha_1 ds_1\right)^2 \\
=& 2\pi A\left(L^2 - 4\pi A\right).
\end{aligned}$$

Equality implies that for all choices $\alpha_1, \alpha_2, s_1, s_2$, there holds

$$r_1 \sin\alpha_2 - r_2 \sin\alpha_1 = 0.$$

Thus for s_1, s_2 constant and $\alpha_2 = \pi/2$, $r_1(s_1, \alpha_1) = r_2(s_2, \pi/2)\csc\alpha_1$ so Γ is a circle with diameter $r_2(s_2, \pi/2)$ which is independent of s_2 (by setting $\alpha_1 = \alpha_2 = \pi/2$).□

Appendix. Definitions, Geometric Background, Various Facts.

The first two lemmata deal with the realizability of quadrilaterals and cyclic quadrilaterals.

Lemma 1. *Suppose four nonnegative numbers a, b, c, d are given that satisfy*

$$a \le b + c + d, \qquad b \le c + d + a, \qquad c \le d + a + b, \qquad d \le a + b + c. \tag{17}$$

Then there is a quadrilateral in the plane whose side lengths are, in order, a, b, c, d.

Proof. We cyclicly permute the lengths if necessary so that

$$m = a + b = \min\{a+b,\ b+c,\ c+d,\ d+a\}$$

then in fact the three numbers m, c, d satisfy

$$m \le c + d, \qquad c \le d + m, \qquad d \le m + c.$$

Thus there is a triangle with sides m, c, d which is also a quadrilateral whose sides $a+b$ line up with m. If one of the inequalities in (17) is equality, then the quadrilateral degenerates to a bigon. □

Lemma 2. *Suppose there are four positive numbers a, b, c, d that satisfy*

$$a < b + c + d, \qquad b < c + d + a, \qquad c < d + a + b, \qquad d < a + b + c. \tag{18}$$

Then there is a cyclic quadrilateral in the plane whose side lengths are given in order by the numbers a, b, c, d.

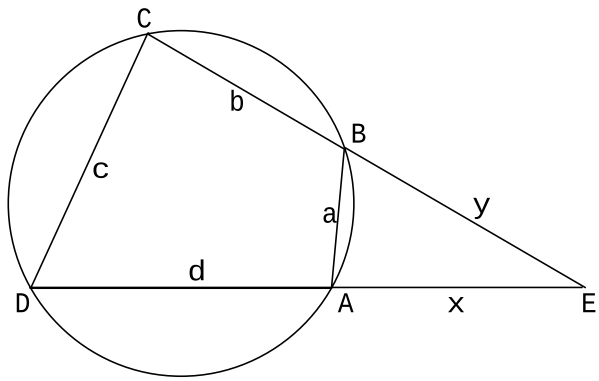


Fig. 8. Construction of Cyclic Quadrilateral

Proof. If $a = c$ then the figure is a trapezoid with b and d sides parallel. Otherwise we may assume that $a < c$ and $b \neq d$ (by cyclically permuting the sides if needed.) Then we may simply construct the cyclic quadrilateral whose sides are given as follows: Imagine that there is a cyclic quadrilateral with vertices $ABCD$ around a circle with $a = AB$, $b = BC$ $c = CD$ and $d = DA$. Then extend DA and CB to their intersection E and suppose that the lengths of the segment $x = AE$ and $y = BE$. Observe that the triangles CED and AEB are similar. To see this, observe that the angles $\angle CBA$ and the angle $\angle ADC$ are supplementary because they subtend opposite sides of the circle. It follows that $\angle CDE = \angle ABE$ hence the triangles are similar since all angles are the same. The similarity implies that

$$\frac{x}{y+b} = \frac{a}{c} \qquad \text{and} \qquad \frac{y}{d+x} = \frac{a}{c}.$$

Since $a < c$ we may solve

$$x = a\frac{bc+ad}{c^2-a^2}, \qquad y = a\frac{ab+cd}{c^2-a^2}.$$

Finally, if we are able to construct the triangle with sides axy then by putting $A = a \cap x$, $B = a \cap y$, $E = x \cap y$, and extending EB by a distance b to find C and extending EA a distance d to find D constructs the cyclic quadrilateral. But $c - a < b + d$, and $(c-a)(a+b+c-d) > 0$ implies $c^2 - a^2 + bc + ad > ab + cd$, and $(c-a)(a-b+c+d) > 0$ implies $c^2 - a^2 + ab + cd > bc + ad$, so, resp.,

$$\begin{aligned}
x + y =& \frac{a(b+d)}{c-a} > a, \\
a + x =& \frac{a(c^2-a^2+bc+ad)}{c^2-a^2} > \frac{a(ab+cd)}{c^2-a^2} = y, \\
a + y =& \frac{a(c^2-a^2+ab+cd)}{c^2-a^2} > \frac{a(bc+ad)}{c^2-a^2} = x
\end{aligned}$$

so the triangle axy is constructible. $\square$

Convergence of Sets in the Hausdorff sense. Given two compact sets in the plane $S, T \subset \mathbf{R}^2$ (compact is equivalent to closed and bounded by the Heine-Borel Theorem, [Wa]) their *Hausdorff Distance*, a measure of their separation is defined by

$$\eta(S,T) = \max\Big\{\sup\{\operatorname{dist}(x,T) : x \in S\}, \sup\{\operatorname{dist}(S,y) : y \in T\}\Big\}.$$

In terms of Minkowski sums this can be defined equivalently by

$$\eta(S,T) = \inf\{r > 0 : S \subset T \boxplus rB \text{ and } T \subset S \boxplus rB\}$$

where, as usual B is the closed unit ball in the plane. A sequence of compact sets $T_n \subset \mathbf{R}^2$ is said to *converge in the Hausdorff sense* to a compact set $T \subset \mathbf{R}^2$ if

$$\eta(T_n, T) \to 0 \qquad \text{as} \qquad n \to \infty.$$

Similarly one can define a Cauchy sequence of compact sets. $\{T_n\}$ is Cauchy if for every $\varepsilon > 0$ there is an $N \in \mathbf{N}$ so that $m, n \geq N$ implies $\eta(T_n, T_m) < \varepsilon$. The compact sets and the compact convex sets are complete: if T_n is a Cauchy sequence of compact (convex) sets, then there is a compact (convex) set T such that $T_n \to T$ as $n \to \infty$.

Blaschke-Hadwiger Selection Theorem. *Let $T_n \in \mathbf{R}^2$ be a uniformly bounded sequence of compact sets: $T_n \subset B$ for some all n where B is some fixed closed disk. Then there is a compact set $T \subset B$ and a subsequence $\{n'\} \subset \mathbf{Z}$ so that $T_{n'} \to T$ as $n' \to \infty$* [B], [F], [H]. *If the T_n are convex then so is T.*

Definitions of Area and Length. For nice sets with piecewise C^1 boundary or which are convex, the various notions of area and length coincide. However, the classes of sets for which these notions are defined differ. Moreover they have different convergence properties. See [C], [H] or [R]. For us we shall take area to mean Lebesgue measure and length to mean Jordan length, the total variation of rectifiable curves. Then the usual line integral gives the area [G].

It is an important fact that both the area and the length of the boundary are semicontinuous with respect to convergence in the appropriate sense. If the boundary curves are parameterized, then the parameterizations have to be controlled. The Frechet distance $\Phi(\gamma_n, \gamma)$ between two parameterized curves γ_n and γ is the smallest $\varepsilon \geq 0$ such that there is a homeomorphism $h : \mathbf{S}^1 \to \mathbf{S}^1$ such that $\max_{x \in \mathbf{S}^1} |\gamma_n(x) - \gamma(h(x))| \leq \varepsilon$ [C].

Theorem. *Suppose Γ_n is a uniformly bounded sequence of closed piecewise C^1 curves in the plane, enclosing compact regions T_n (or the T_n are convex) which are converging to a compact set T bounded by a closed piecewise C^1 curve Γ (or T is convex) in the sense of Hausdorff convergence. Suppose also that $\mathrm{L}(\Gamma_n) \leq L$ and $\mathrm{L}(\Gamma) \leq L$ are uniformly bounded by $L < \infty$. Then*

$$\lim_{n \to \infty} \operatorname{Area}(T_n) = \operatorname{Area}(T).$$

Suppose furthermore that Γ_n is converging to Γ in the sense that Fréchet distance of the parametrizations $\Phi(\gamma_n,\gamma)\to 0$. ($\Gamma_n=\gamma_n(\mathbf{S}^1)$.) Then

$$\liminf_{n\to\infty} \mathrm{L}(\Gamma_n)\ge \mathrm{L}(\Gamma).$$

Proof. Let $r=\eta(T_n,T)$. First approximate Γ_n and Γ by polygons that are close in Hausdorff distance, length and area ([B]. p. 28.) There is a closed polygonal curve q enclosing Q with $|\operatorname{Area}(Q)-\operatorname{Area}(T)|<r$, $|\mathrm{L}(q)-\mathrm{L}(\Gamma)|<r$ and $\eta(Q,T)<r$. Similarly there is a closed polygonal curve q_n enclosing the region Q_n with $|\operatorname{Area}(Q_n)-\operatorname{Area}(T_n)|<r$, $|\mathrm{L}(q_n)-\mathrm{L}(\Gamma_n)|<r$ and $\eta(Q_n,T_n)<r$. Note that $T\subset Q_r$ so that $T_n\subset T_r\subset Q_{2r}$. Similarly, $T\subset (Q_n)_{2r}$. Then, since Steiner's Inequality applies to polyhedra by the Corollary to the First Variation Formula,

$$\begin{aligned}
\operatorname{Area}(T_n)-\operatorname{Area}(T) &\le \operatorname{Area}(Q_{2r})-\operatorname{Area}(T)\\
&\le \operatorname{Area}(Q)+4\,\mathrm{L}(q)r+4\pi r^2-\operatorname{Area}(T)\\
&\le r+4Lr+(4+4\pi)r^2\\
\operatorname{Area}(T_n)-\operatorname{Area}(T) &\ge \operatorname{Area}(T_n)-\operatorname{Area}((Q_n)_{2r})\\
&\ge \operatorname{Area}(T_n)-\operatorname{Area}(Q_n)-4\,\mathrm{L}(q_n)r-4\pi r^2\\
&\ge -r-4Lr-(4+4\pi)r^2
\end{aligned}$$

Thus if $n\gg 1$ so that $\eta(T_n,T)\le 1$,

$$|\operatorname{Area}(T_n)-\operatorname{Area}(T)|\le (5+4L+4\pi)\eta(T_n,T)\to 0 \qquad \text{as} \qquad n\to\infty.$$

The length formula follows by the semicontinuity of Jordan length [C], [R] with respect to Fréchet convergence of curves. $\square$

If we are willing to restrict ourselves to simple closed curves (Jordan curves without self intersections), then the statement is much simpler. For piecewise C^1 simple curves the arclength of the boundary curve of $\Omega\subset\mathbf{R}^2$ is given by the perimiter

$$\operatorname{Per}(\Omega,\mathcal{O})=\int_{\mathcal{O}}|D\chi_\Omega|$$

where the characteristic function $\chi_\Omega(x)=1$ if $x\in\Omega$ and $\chi_\Omega(x)=0$ otherwise and where the integral means the variation. Thus for $\mathcal{O}\subset\mathbf{R}^2$ an open set and $f\in L^1(\mathcal{O})$ we define

$$\int_{\mathcal{O}}|Df|=\sup\left\{\int_{\mathcal{O}} f\,(g_x+h_y)\,dxdy : g,h\in C_0^1(\mathcal{O}),\ g^2+h^2\le 1\right\}.$$

If $\operatorname{Per}(\Omega,\mathcal{O})<\infty$ for all bounded open sets $\mathcal{O}$ then we say Ω is a Cacciopoli set. Then $\operatorname{Per}(\Omega,\mathcal{O})=\mathcal{H}^1(\mathcal{O}\cap\partial\Omega)$ is the Hausdorff measure of the boundary[Gi]. The subspace $BV(\Omega)=\{f\in L^1(\Omega):\int_\omega|Df|<\infty\}$ with the norm

$$\|f\|_{BV(\Omega)}=\|f\|_{L^1(\Omega)}+\int_\Omega|Df|$$

is a Banach space. Semiconinuity holds for perimeter [Gi].

Lemma. Semicontinuity for BV. *If $\Omega \subset \mathbf{R}^2$ is open and $\{f_j\}$ a sequence in $BV(\Omega)$ which converges in $L^1_{\text{loc}}(\Omega)$ to a function f then*

$$\int_\Omega |Df| \le \liminf_{j\to\infty} \int_\Omega |Df_j|.$$

This implies that if T_n are uniformly bounded Jordan domains such that $T_j \to T$ in Hausdorff distance (so $\chi_{T_j} \to \chi_T$ in $L^1_{\text{loc}}(\Omega)$) then $\operatorname{Per}(T, \mathcal{O}) \le \liminf \operatorname{Per}(T_j, \mathcal{O})$. Equivalently, this is a statement about the semicontinuity of the mass of a 1-current with respect to the flat norm [M].

We deduce some other facts that are being used in the discussion above. We would like to find the rate of change of arclength for a one parameter family of curves.

First Variation Formula for closed curves in the plane. *Suppose $f(s,t)$ is a one parameter family of closed continuous, piecewise C^2 curves in the plane defined on C^2 pieces $f_i : \{(s,t) : -\varepsilon < s < \varepsilon, l_i(s) \le t \le u_i(s)\} \to \mathbf{R}^2$ where $l_i(s), u_i(s)$ are differentiable functions on $(-\varepsilon, \varepsilon)$ such that $f_{i-1}(s, u_{i-1}(s)) = f_i(s, l_i(s))$ and $f_1(s, l_1(s)) = f_k(s, u_k(s))$ for all $i = 1, \dots, k$. The length is $L(s) = \sum_i \mathrm{L}(f_i(s,\cdot))$. Let $T = (f_i)_t/|(f_i)_t|$ be the tangent vector, N be the outward normal, τ the arclength along $f(0,t)$, and κ the curvature. Then*

$$\begin{aligned} L'(0) = &\sum_{i=1}^k \int_{l_i(s)}^{u_i(s)} \langle \kappa N, V\rangle \, d\tau + \\ &+ \sum_{i=1}^k \left(|(f_{i-1})_t(s, u_{i-1})| \frac{d\,u_{i-1}}{ds} - |(f_i)_t(s, l_i)| \frac{d\,l_i}{ds} + \langle T, V\rangle \Big|_{l_i}^{u_i} \right). \end{aligned}$$

Proof. The length of the piecewise C^2 curve is given by

$$L(s) = \sum_{i=1}^k \int_{l_i(s)}^{u_i(s)} |(f_i)_t(s,t)| \, dt.$$

Set $f_0 = f_k$ and $u_0 = u_k$. Differentiating,

$$\begin{aligned} L'(s) &= \sum_{i=1}^k \left(|(f_i)_t(s,u_i)| u_i' - |(f_i)_t(s, l_i)| l_i' + \int_{l_i(s)}^{u_i(s)} \frac{\langle f_t, f_{st}\rangle}{|f_t(s,t)|} \, dt \right) \\ &= \sum_{i=1}^k \Bigg(|(f_i)_t(s,u_i)| u_i' - |(f_i)_t(s, l_i)| l_i' \\ &\qquad - \int_{l_i(s)}^{u_i(s)} \left\langle \frac{d}{dt}\left(\frac{f_t}{|f_t|}\right), f_s \right\rangle dt + \frac{\langle f_t, f_s\rangle}{|f_t|}\Big|_{l_i}^{u_i} \Bigg). \end{aligned}$$

Thus at $s = 0$, writing the variation vector field as $V(t) = f_s(0,t)$,

$$L'(0) = \sum_{i=1}^{k} \int_{l_i(s)}^{u_i(s)} \langle \kappa N, V \rangle |(f_i)_t| \, dt + \\ + \sum_{i=1}^{k} \left(|(f_{i-1})_t(s, u_{i-1})| u'_{i-1} - |(f_i)_t(s, (l_i))| l'_i + \langle T, V \rangle \Big|_{l_i}^{u_i} \right). \qquad \square$$

We now provide the estimate for the growth of the boundary curve needed for Steiner's inequality. Let $P \subset \mathbf{R}^2$ be a compact bounded by a polygonal curve with finitely many vertices. Then the boundary of the parallel set $P_r = P \boxplus rB$ consists of finitely many line segments and arcs of circles, thus a piecewise C^2 curve. Suppose that going around the polygon consists of vertices X_i and closed line segments corresponding to the edges $e_i = X_i X_{i+1}$ where $i = 1, \ldots, k$ and $X_{k+1} = X_1$. For every point in the plane $X \in \mathbf{R}^2 - P$ has a certain distance to P given by $d = \operatorname{dist}(X, P)$. The closest point on P to X may be one point that is interior to an edge or it may be a vertex of P, or it may be several closest points, each of which is in some edge or some vertex. Label each X by the set of closest edges and vertices. The exterior region is then subdivided into zones with the same labelling. These may be single points, curves, or open regions. They may be infinite or finite, but there are only finitely many zones in the plane. The zones with more than two closest elemets are points. The zones with two labels occur on the are curves, for example the curve equidistant between two points (a straight line) or the locus of points equidistant from a vertex and a line (a straight line if the vertex is an endpoit of the line, or a parabola, in case the point is not an endpint of the edge.) In both cases, the boundary of the parallel set is transverse to the locus of two closest elements. It follows that for only finitely many $r \in \{r_1 < r_2 < \cdots < r_q\}$ such that the boundary of the parallel set contains points that are closest to more than two elements. Thus if $r_j < r < r_{j+1}$ (which we call a *generic* distance) then ∂P_r consists of $k < \infty$ segments which are line segments or arcs of circles, and such that the endpoints of the segments depend in a C^1 fashion on r since they are given by the intersection of ∂P_r with parabolas and lines (the two element zones.) For convenience let t be the arclength parameter, N the outward normal and $f(r,t)$ the parameterization of ∂P_r. Then $f_r = N$ and $\kappa = 0$ on line segments in ∂P_r which correspond to "one edge zones" and circles whose curvatures are always $\kappa = 1/r$, that correspond to "one vertex zones." Note that ∂P_r contains no concave circular zones. Thus we have the corollary.

Corollary. Steiner's Inequality for Polygons. *Suppose P is the compact region of area A bounded by a closed polygonal curve p of length L. Let $r > 0$ be a generic distance and $L(r) = \mathrm{L}(\partial P_r)$ be the length of the boundary of the equidistant domain. Then $L'(r) \le 2\pi$. Hence, for all $r \ge 0$,*

$$\mathrm{L}(\partial P_r) \le L + 2\pi r \qquad \textit{and} \qquad \operatorname{Area}(P_r) \le A + Lr + \pi r^2,$$

with equality for convex sets.

Proof. Let each circular or linear piece of the boudary be parameterized by arc-length $l_i(r) \le t \le u_i(r)$ so $T = f_t$, $V = N$. By the first variation formula,

$$L'(0) = \sum_{i=\text{circle}} \int_{l_i(s)}^{u_i(s)} \frac{1}{r}\, dt + \sum_{i=1}^{k} \left(\frac{d\, u_{i-1}}{dr} - \frac{d\, l_i}{dr} \right). \tag{19}$$

The segments of ∂P_r don't necessarily correspond to all vertices and edges of P. It may happen that at some edpoint of a circular arc, two vertices X_i and X_j of P are the closest points. This may happen if there is a "concave section" of ∂P between X_i and X_j. Let θ_i denote the "total curvature" corresponding to the arc f_i of ∂P. If f_i is a line, then $\theta_i = 0$. If f_i is a circular arc corresponding to a "convex" vertex, say X_j, then $\theta_i = \angle(e_{j-1}, e_j)$ is the angle deficit form the edge e_{j-1} to e_j. If one takes the outward normal N_{j-1} of the E_{j-1} edge and N_j corresponding to the e_j edge, then the zone labelled X^j is the sector in the plane whose vertec is X_j and whose sides are rays in the directions N_{j-1} and N_j and whose angle $\angle(N_{j-1}, N_j) = \theta_i$. The points of the arc f_i are a distance r from X_j in this sector.

The remaining circular arcs f_i of ∂P_r correspond to points X_j that are not convex, in the sense that at an endpoint of f_i is not closest to an adjacent edge of X_j. Say the endpoint u_i is closest to the vertex X_j and the vertex X_m of the edge e_m not adjacent to X_j. The total angle of such an arc is not the angle between adjacent edges at X_j but rather some part of it. The remaining angle of the vertex X_j has to be counted in the "total curvatuere" of the "concave part of ∂P." Thus, the angle corresponding to the endpoint u_i (same as l_{i+1}) is the remainder of the angle at X_j plus the total curvature (the sum of the angle deficits of intervening vertices and whatever angle there is left between N_{m-1} and $X_m X$. Call the total ν_i which is negative. The total curvature around P is thus

$$\sum_{\theta_i > 0} \theta_i + \sum_{\nu_i < 0} \nu_i = \int_{\partial P_r} \kappa\, dt + \sum_i \nu_i = 2\pi.$$

which is just the Gauß Bonnet Theorem for ∂P_r. It remains to compute u_i' and l_i'. At a convex vertex, the arc f_i is either a sector or a line segment whose normal at the endpoint u_i coincides with N_j of N_{j-1} which is the normal of one of the adjacent edges of X_j, one of the closest points to $f_i(u_i)$. It follows that $u_i' = l_{i+1}' = 0$. Suppose that the endpoint $f_{i-1}(u_{i-1})$ is closest to a pair of vertices X_j and X_m of P. In that case, the curve $r \to Y(r) = f_{i-1}(r, u_{i-1}(r)) = f_i(r, l_i(r))$ is on the perpendicular bisector of $X_j X_m$, and the arcs on both sedes of this point are decreasing by the same amount. The angle between the perpendicular bisector and the vector $X_j Y(r)$ is $-\nu_i/2$. It follows that

$$u_{i-1}'(r) = \tan(\nu_i/2) \qquad \text{and} \qquad l_i'(r) = -\tan(\nu_i/2). \tag{20}$$

The final case is if $Y(r)$ is closest to a vertex, say X_j and a line segment e_m. Let $Z_m \in e_m$ be the interior point where $\operatorname{dist}(Y(r), X_j) = \operatorname{dist}(Y(r), Z_m)$. Now the locus os points equidistant from X_j and e_m is a parabola. By the geometry of parabolas, the tangent line to the parabola at $Y(r)$ bisects the angle $\angle(Y(r)X_j, Y(r)Z_m)$. It follows that (20) holds in this case as well. Finally, it may happen that $Y(r)$ is equidistant to two line segments e_j and e_m. The locus of points equidistant to two lines is the angle bisector between the lines. Thus if $X_j \in e_j$ $Z_m \in e_m$ are interior points where $\operatorname{dist}(Y(r), e_j) = \operatorname{dist}(Y(r), e_m) = \operatorname{dist}(Y(r), X_j) = \operatorname{dist}(Y(r), Z_m)$ then $-\nu_i/2 = \angle(Y(r)X_j, Y(r)Z_m)$ and therefore (20) holds in this case as well.

Thus we can complete the estimate using $dt = r\,d\theta$ on circular arcs in (19) so

$$L'(r) = \sum_{\theta_i > 0} \theta_i + \sum_{\nu_i < 0} 2\tan\left(\frac{\nu_i}{2}\right) < \sum_{\theta_i > 0} \theta_i + \sum_{\nu_i < 0} \nu_i = 2\pi.$$

If P was convex, then there are no zones corresponding to the concave parts of P so no ν_1 terms and this inequality is equality. Suppose $0 = r_0 \le r_1 < \ldots < r_j \le r$, then

$$\begin{aligned} L(r) =& L(0) + \sum_{i=1}^{j} \int_{r_{i-1}}^{r_j} L'(r)\,dr + \int_{r_j}^{r} L'(r)\,dr \\ \le& L + \sum_{i=1}^{j} \int_{r_{i-1}}^{r_j} 2\pi\,dr + \int_{r_j}^{r} 2\pi\,dr = L + 2\pi r. \end{aligned}$$

Finally, since

$$A'(r) = \int_{\partial P_r} \langle V, N\rangle\,dt$$

it follows that since $V = f_r = N$ that

$$A(r) = A(0) + \int_0^r L(r)\,dr \le A + Lr + \pi r^2$$

and the proof is complete. □

Remarks. These notes were inspired by my talk "My Favorite Proofs of the Isoperimetric Inequality" given in the Undergraduate Colloquium at the University of Utah on Nov. 20, 2001. They are offered as an instructional module, which may be useful for the course on Curves and Surfaces, geometric methods in education, beginning analysis, advanced calculus or for an introduction to proofs. The notes represent a self contained discussion of the isoperimetric inequality. One of the objectives of the author was to show by example that that many of the arguments are interconnected. Methods that prove one inequality also can be used to prove the others. We mention as exercises specific instances of these interconnections. For example, in one problem the student is asked to show that the Wirtinger inequality implies the Brunn-Minkowski inequality.

A second objective of the author is to show the importance of inequalities. However, we deduce our inequalities by producing a sequence of equalities and then discarding the obviously negative terms to obtain the desired inequality. This gives an easy way to treat the case of equalty. It sometimes makes for a slightly circuitous route to a result for the sake of self-containedness. Thus we don't invoke the Schwarz inequality when it is just as easy to complete the square and discard a negative term.

We have made a selection of proofs which are usually not presented together in the same reference. For example, in Chern's beautiful book *Global Differential Geometry*, [C], the proofs of Schmidt and Hurwitz are presented. We select one of these proofs. do Carmo's undergraduate text [dC] on Curves and Surfaces presents Schmidt's, Hurwitz's and Santalo's proof of the isoperimetric inequality. Blashke's text [B] on the Circle and Sphere presents the Steiner symmetrization proof and the hinge proof, of which we borrow the hinge proof, but we don't deduce the Brunn-Minkowski inequality using symmetrization as Blaschke does. Blaschke & Reichardt's text [BR] on differential geometry presents Crone & Frobenius's proof, Knothe's proof, Hurwitz's proofs, Sz. Nagy's and Bol's proof and Santalo's proofs, from which we borrow Knothe's (which is attributed to Blaschke in [S].) Crone and Frobenius's method is used instead to prove the Brunn inequality in Ljusternik's text [L] on convex figures, from which the isoperimetric inequality is deduced in turn. In Blaschke and Leichtweiß's differential geometry text [BL], the calculus of variations argument is given, as well as Schmidt's and Hurwitz's proofs of the isoperimetric inequality and Groß's proof of the convergence of Steiner symmetrization. The quadrilateral and Ptolemy's inequalities are one of hundreds of inequalities given in [BD]. Fourrey gives delightful historical commentary [Fo] as well as several proofs of the Pythagorean Theorem and Heron's Formula.

We have included references for most of the facts that are quoted. Space prohibits a really careful discussion, particularly about the meaning of area, length and the regularity needed to make these notions rigorous. Although we hope to have provided ample references to where these issues are discussed. Many of the sources are accessible to undergraduate students.

There are several places where the isoperimetric inequality is discussed in detail. We mention specifically Osserman's article [O] and the book of Burago-Zalgaller [BZ].

References

[B] W. Blaschke, *Kreis und Kugel*, Chelsea, New York, 1949, Originally published by Teubner, Leipzig, 1916.

[BL] W. Blaschke & K. Leichtweiß, *Elementare Differentialgeometrie 5th. ed.*, Springer, Berlin, 1973.

[BR] W. Blaschke & H. Reichardt, *Einführung in die Differentialgeometrie, 2nd. ed.*, Springer, Berlin, 1960.

[BF] T. Bonnesen & W. Fenchel, *Theorie der Konvexen Körper*, Springer, Berlin, 1934.

[BD] O. Bottema, R. Djordjević, R. Janić, D. Mitrinović, P. Vasić, *Geometric Inequalities*, Wolters Noordhoff Publishing, Groningen, 1968.

[BZ] J. Burago & V. Zalgaller, *Geometric Inequalities*, Springer, Berlin, 1988., Originally published Nauka, Leningrad, 1980.

[C] L. Cesari, *Rectifiable Curves and the Weierstrass Integral*, Amer. Math. Monthly **65** (1958), 485-500.

[dC] M. do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice Hall, Englewood Cliffs, 1976.

[F] H. Federer, *Geometric Measure Theory*, Springer, Berlin, 1969, p. 183.

[Fo] E. Fourrey, *Curiosités Géométriques*, Librairie Vuibert, Paris, 1938.

[GT] D. Gilbarg & N. Trudinger, *Elliptic Partial Differential Equations of the Second Order, 2nd ed.*, Springer, Berlin, 1977, p. 157.

[Gi] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäser, Boston, 1984.

[G] H. Guggenheimer, *Differential Geometry*, Dover, Mineola, 1977, Originally published by McGraw-Hill, New York 1963.

[G2] H. Guggenheimer, *Applicable Geometry: Global and Local Convexity*, Krieger, Huntington, 1977.

[H] H. Hadwiger, *Vorlesungen Über Inhalt, Oberfläche und Isoperimetrie*, Springer, Berlin, 1957, p. 153.

[H2] H. Hadwiger, *Altes und Neues über Konvexe Körper*, Birkhäuser, Basel, 1955.

[LT] P. Li & A. Treibergs, *Applications of Eigenvalue Techniques to Geometry*, Contemporary Geometry: J.-Q, Zhong Memorial Volume, H. H. Wu, ed., Plenum Press, New York, 1991, pp. 22–95.

[L] Ljusternik, *Convex Figures and Polyhedra*, Heath & Co., Boston, 1966.

[M] F. Morgan, *Geometric Measure Theory: a Beginners Guide*, Academic Press, San Diego, 1988, p. 58.

[O] R. Osserman, *The isoperimetric inequality*, Bull. American Math. Soc, **84** (1978), 1128–1238.

[P] G. Pólya, *Mathematics of Plausible Reasoning*, Volume I: Induction and Analogy in Mathematics, Princeton U. Press, Princeton, 1954.

[PS] G. Pólya & G. Szegö, *Isoperimetric Inequalities in Mathematical Physics*, Annals of Math. Studies, vol. 27, Princeton U. Press, Princeton, 1951.

[R] T. Radó, *Length and Area*, Colloquium Publications, vol. 30, American Mathematical Society, New York, 1948.

[S] L. Santaló, *Integral Geometry and Geometric Probability*, Encyclopedia of Mathematics and its Applications, Vol. 1, Addison Wesley Publ. Co., Reading, 1976.

[St] D. Struik, *A Concise History of Mathematics, 3rd. ed.*, Dover, New York, 1967.

[Wa] W. Wade, *An Introduction to Analysis, 2nd. ed.*, Prentice Hall, Upper Saddle River, N. J., 2000, p. 263.

[W] H. Weinberger, *A First Course in Partial Differential Equations*, Dover, Mineola, 1997, Originally published by Xerox College Publishing, Lexington, 1965.

University of Utah,
Department of Mathematics
155 South 1400 East, Rm 233
Salt Lake City, UTAH 84112-0090
E-mail address: treiberg@math.utah.edu