The isoperimetric inequality in the plane

Andrew Putman

Let γ be a simple closed curve in $\mathbb{R}^2$. Choose a parameterization $f : [0,1] \to \mathbb{R}^2$ of γ and define

$$\text{length}(\gamma) = \sup_{0=t_0<t_1<\cdots<t_k=1} \sum_{i=1}^{k} d_{\mathbb{R}^2}(f(t_{i-1}), f(t_i)) \in \mathbb{R} \cup \{\infty\}; \tag{0.1}$$

here the supremum is taken over all partitions of $[0,1]$. This does not depend on the choice of parameterization. By the Jordan curve theorem, the simple closed curve γ encloses a bounded region in $\mathbb{R}^2$; define $\text{area}(\gamma)$ to be the Lebesgue measure of this bounded region. The classical isoperimetric inequality is as follows.

Isoperimetric Inequality. *If γ is a simple closed curve in $\mathbb{R}^2$, then* $\text{area}(\gamma) \leq \frac{1}{4\pi}\text{length}(\gamma)^2$ *with equality if and only if γ is a round circle.*

This inequality was stated by Greeks but was first rigorously proved by Weierstrass in the 19th century. In this note, we will give a simple and elementary proof based on geometric ideas of Steiner. For a discussion of the history of the isoperimetric inequality and a sample of the enormous number of known proofs of it, see [1] and [3].

Our proof will require three lemmas. The first is a sort of "discrete" version of the isoperimetric inequality. A polygon in $\mathbb{R}^2$ is *cyclic* if it can be inscribed in a circle.

Lemma 1. *Let P be a noncyclic polygon in $\mathbb{R}^2$. Then there exists a cyclic polygon P' in $\mathbb{R}^2$ with the same cyclically ordered side lengths as P satisfying* $\text{area}(P) < \text{area}(P')$.

Proof. All triangles are cyclic, so P has at least 4 sides. The set of all polygons in $\mathbb{R}^2$ with the same cyclically ordered side lengths as P and with one vertex at the origin is compact. It follows that there exists a polygon P' in $\mathbb{R}^2$ with the same cyclically ordered side lengths as P whose area is maximal among all such polygons. We will prove that P' is cyclic. It is clear that P' is convex. There are now two cases.

Case 1. *The polygon P has 4 sides.*

We remark that this case could be deduced immediately from Bretschneider's formula for the area of a convex quadrilateral (see [2]), but we will give a self-contained proof.

Let a, b, c, and d be the side lengths of P (cyclically ordered). Consider a convex polygon Q with the same cyclically ordered side lengths as P. Let

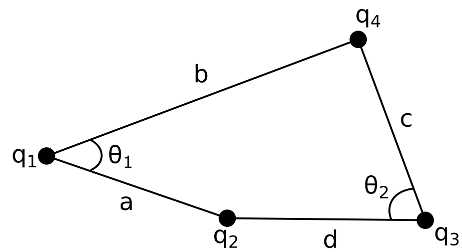

Figure 1: *The quadrilateral Q in Step 1 of the proof of Lemma 1.*

$q_1, \ldots, q_4$ be the vertices and let θ_1 and θ_2 be the angles labeled in Figure 1. Since any three non-colinear points determine a circle, there are circles containing $\{q_1, q_2, q_4\}$ and $\{q_2, q_3, q_4\}$. These circles will be the same (and hence Q will be cyclic) exactly when $\theta_1 + \theta_2 = \pi$.

It is clear that the isometry class of Q is determined by θ_1 and θ_2. However, not all pairs of angles are possible; indeed, computing the length of the diagonal from q_2 to q_4 using the law of cosines in two ways, we see that

$$a^2 + b^2 - 2ab\cos(\theta_1) = c^2 + d^2 - 2cd\cos(\theta_2). \tag{0.2}$$

Conversely, any angles θ_1 and θ_2 satisfying (0.2) and $0 \leq \theta_1, \theta_2 \leq \pi$ can be realized by some convex polygon as above. The area of Q is $\frac{1}{2}ab\sin(\theta_1) + \frac{1}{2}cd\sin(\theta_2)$. Letting $f(\theta_1, \theta_2) = ab\sin(\theta_1) + cd\sin(\theta_2)$ and $g(\theta_1, \theta_2) = a^2 + b^2 - 2ab\cos(\theta_1) - c^2 - d^2 + 2cd\cos(\theta_2)$, our goal therefore is to show that among all angles satisfying $0 \leq \theta_1, \theta_2 \leq \pi$ and $g(\theta_1, \theta_2) = 0$, the function $f(\theta_1, \theta_2)$ is maximized when $\theta_1 + \theta_2 = \pi$.

It is clear that this maximum will occur when $0 < \theta_1, \theta_2 < \pi$, so using Lagrange multipliers we see that at this maximum, there will exist some $\lambda \in \mathbb{R}$ such that $\nabla f = \lambda \nabla g$, i.e. such that

$$ab\cos(\theta_1) = 2ab\lambda\sin(\theta_1) \quad \text{and} \quad cd\cos(\theta_2) = -2cd\lambda\sin(\theta_2).$$

Since $0 < \theta_1, \theta_2 < \pi$, we have $\sin(\theta_1) \neq 0$ and $\sin(\theta_2) \neq 0$, so we can manipulate the above formulas and see that $\cot(\theta_1) = -\cot(\theta_2)$. This implies that $\theta_1 + \theta_2 = \pi$, as desired.

Case 2. *The polygon P has more than 4 sides.*

Assume that P' is not cyclic. This implies that there exist four vertices $q_1, \ldots, q_4$ of P' that do not lie on a circle. Let Q be the quadrilateral with these four vertices. Using Case 1, there exist a cyclic quadrilateral Q' with the same side lengths as Q but with $\text{area}(Q) < \text{area}(Q')$. Let $X_1, \ldots, X_4$ be the components of $P' \setminus Q$ adjacent to the four sides of of Q (possibly some of the X_i are empty), so $\text{area}(P') = \text{area}(Q) + \text{area}(X_1) + \cdots + \text{area}(X_4)$. As is shown in Figure 2, we can attach the X_i to Q' to form a polygon P'' whose cyclically ordered side lengths are the same as those of P' but whose area equals $\text{area}(Q') + \text{area}(X_1) + \cdots + \text{area}(X_4)$. But this implies that $\text{area}(P'') > \text{area}(P')$, contradicting the maximality of the area of P'. □

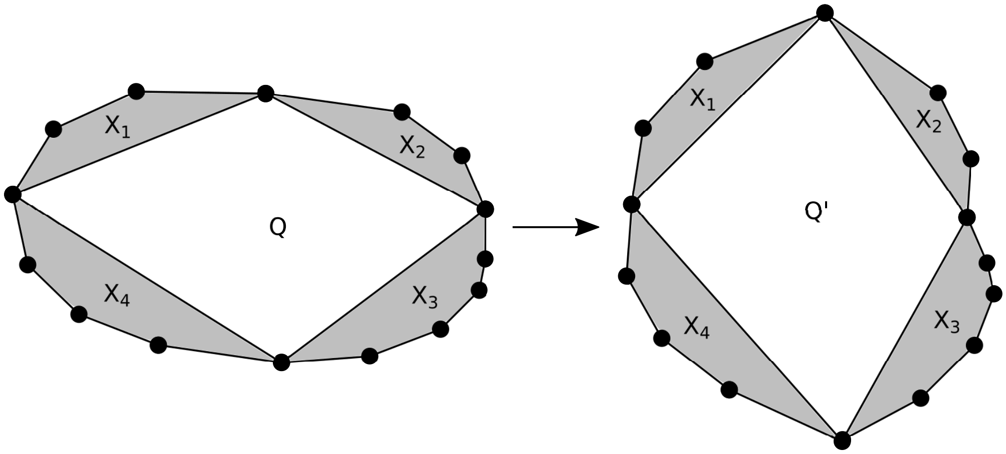

Figure 2: *Changing the quadrilateral Q in P to Q' (without changing the side lengths of P) increases the area since* $\text{area}(Q) < \text{area}(Q')$ *but the four pieces $X_1, \ldots, X_4$ making up the rest of P just are rotated without their area changing.*

For the second lemma, say that a simple closed curve in $\mathbb{R}^2$ is *convex* if it encloses a convex region.

Lemma 2. *Let γ be a simple closed curve in $\mathbb{R}^2$ and let γ' be the boundary of the convex hull of the closed region enclosed by γ. Then γ' is a convex simple closed curve satisfying* $\text{length}(\gamma') \leq \text{length}(\gamma)$.

Proof. Parameterize γ as $f : [0,1] \to \mathbb{R}^2$ and define $\Lambda = f^{-1}(\gamma \cap \gamma')$. Choose f such that $f(0) \in \gamma'$, and hence $0, 1 \in \Lambda$. The set Λ is nonempty and closed, so its complement consists of at most countably many disjoint open intervals $\{I_\alpha\}_{\alpha \in A}$. For $\alpha \in A$, write $\partial I_\alpha = \{x_\alpha, y_\alpha\} \subset \Lambda$ with $x_\alpha < y_\alpha$. Define $f' : [0,1] \to \mathbb{R}^2$ to equal f on Λ and to parameterize a straight line from $f(x_\alpha)$ to $f(y_\alpha)$ on I_α for all $\alpha \in A$. The function f is then a parameterization of γ'. Let $\mathcal{P}$ be the set of all partitions of $[0,1]$. For $P \in \mathcal{P}$ written as $0 = t_0 < t_1 < \cdots < t_k = 1$, define

$$\ell(f, P) = \sum_{i=1}^{k} d_{\mathbb{R}^2}(f(t_{i-1}), f(t_i)) \quad \text{and} \quad \ell(f', P) = \sum_{i=1}^{k} d_{\mathbb{R}^2}(f'(t_{i-1}), f'(t_i)).$$

Our goal is to show that $\sup_{P \in \mathcal{P}} \ell(f', P) \leq \sup_{P \in \mathcal{P}} \ell(f, P)$.

Define $\mathcal{P}_1$ to be the set of partitions P of $[0,1]$ such that if a point of I_α appears in P for some $\alpha \in A$, then both x_α and y_α appear in P. Since every partition can be refined to a partition in $\mathcal{P}_1$, we have

$$\sup_{P \in \mathcal{P}} \ell(f', P) = \sup_{P \in \mathcal{P}_1} \ell(f', P). \tag{0.3}$$

Next, define $\mathcal{P}_2$ to be the set of partitions P of $[0,1]$ that contain no points of I_α for any $\alpha \in A$. For $P \in \mathcal{P}_1$, define $\widehat{P} \in \mathcal{P}_2$ to be the result of deleting all points that lie in I_α for some $\alpha \in A$. The key observation is that

$$\ell(f', P) = \ell(f', \widehat{P}) = \ell(f, \widehat{P}) \qquad (P \in \mathcal{P}_1).$$

This implies that

$$\sup_{P\in\mathcal{P}_1} \ell(f', P) = \sup_{P\in\mathcal{P}_2} \ell(f', P) = \sup_{P\in\mathcal{P}_2} \ell(f, P) \leq \sup_{P\in\mathcal{P}} \ell(f, P). \tag{0.4}$$

Combining (0.3) and (0.4), the lemma follows. □

Lemma 3. *Let γ be a convex simple closed curve in $\mathbb{R}^2$. Then for all $\epsilon > 0$, there exists a polygon P inscribed in γ satisfying* $\text{area}(P) > \text{area}(\gamma) - \epsilon$.

Proof. Translating γ, we can assume that 0 lies in its interior. For $0 < \delta < 1$, define $\gamma_\delta = \{\delta \cdot x \mid x \in \gamma\}$. Then γ_δ is a convex simple closed curve contained in the interior of the region bounded by γ satisfying

$$\text{area}(\gamma_\delta) = \delta^2 \cdot \text{area}(\gamma).$$

Choose δ sufficiently close to 1 such that $\text{area}(\gamma_\delta) > \text{area}(\gamma) - \epsilon$. We can then find a polygon P inscribed in γ such that γ_δ lies in the interior of P, and hence $\text{area}(P) > \text{area}(\gamma_\delta) > \text{area}(\gamma) - \epsilon$. □

Proof of the isoperimetric inequality. The theorem is trivial if $\text{length}(\gamma) = \infty$, so assume without loss of generality that $\text{length}(\gamma) < \infty$. Assume first that γ is not convex. Let γ' be the boundary of the convex hull of the region bounded by γ. Lemma 2 says that $\text{length}(\gamma') \leq \text{length}(\gamma)$, and it is clear that $\text{area}(\gamma') > \text{area}(\gamma)$. It is therefore enough to prove the theorem for γ'. Replacing γ with γ', we can therefore assume that γ is convex.

Fix some $\epsilon > 0$. Use Lemma 3 to find a polygon P inscribed in γ such that $\text{area}(P) > \text{area}(\gamma) - \epsilon$. Since P is inscribed in γ, we have $\text{length}(P) \leq \text{length}(\gamma)$. Lemma 1 ensures that there exists a cyclic polygon P' with the same cyclically ordered side lengths as P satisfying $\text{area}(P') \geq \text{area}(P)$. Let C be the circle in which P' is inscribed. Since P' is inscribed in C, we have $\text{area}(P') < \text{area}(C)$. Adding more vertices to P, we can ensure that $\text{length}(P') > \text{length}(C) - \epsilon$. We now combine all of the our estimates to deduce that

$$\begin{aligned}
\text{area}(\gamma) &< \text{area}(P) + \epsilon \leq \text{area}(P') + \epsilon < \text{area}(C) + \epsilon \\
&= \frac{1}{4\pi}\text{length}(C)^2 + \epsilon < \frac{1}{4\pi}(\text{length}(P') + \epsilon)^2 + \epsilon \\
&= \frac{1}{4\pi}(\text{length}(P) + \epsilon)^2 + \epsilon \leq \frac{1}{4\pi}(\text{length}(\gamma) + \epsilon)^2 + \epsilon.
\end{aligned}$$

Since $\text{area}(\gamma) < \frac{1}{4\pi}(\text{length}(\gamma)+\epsilon)^2+\epsilon$ for all $\epsilon > 0$, we conclude that $\text{area}(\gamma) \leq \frac{1}{4\pi}\text{length}(\gamma)$, as desired.

To finish the proof, we must show that $\text{area}(\gamma) < \frac{1}{4\pi}\text{length}(\gamma)$ when γ (still assumed to be convex) is not a round circle. Since γ is not a round circle, we can find four points $q_1, \ldots, q_4 \in \gamma$ that do not lie on a circle. Let Q be the quadrilateral inscribed in γ with the vertices $q_1, \ldots, q_4$. By Lemma 1, we can find a cyclic quadrilateral Q' with the same side lengths as Q but with

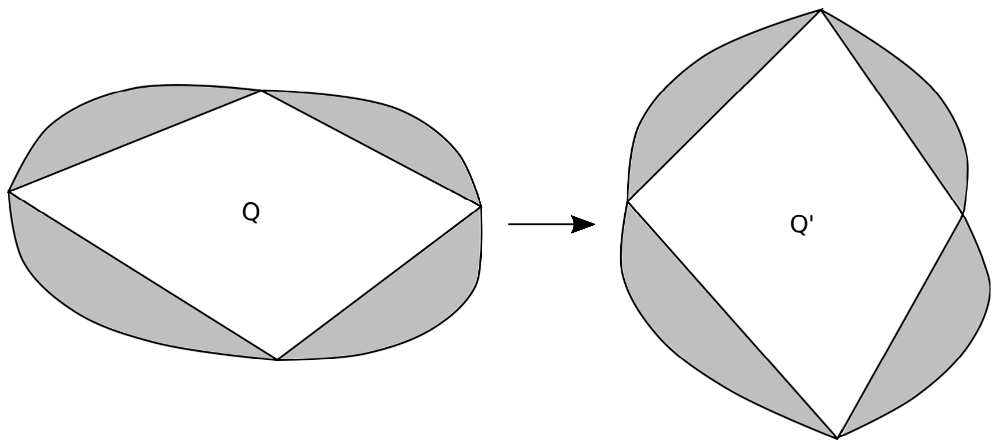

Figure 3: *Just like in the second step of the proof of Lemma 1, we change Q to Q' without changing the length of γ; each of the four shaded regions is merely rotated and glued onto Q'.*

$\text{area}(Q') > \text{area}(Q)$. Just like in Case 2 of the proof of Lemma 1, we can use Q' to find a simple closed curve γ' with $\text{length}(\gamma') = \text{length}(\gamma)$ but with $\text{area}(\gamma') > \text{area}(\gamma)$ (see Figure 3). This implies that

$$\text{area}(\gamma) < \text{area}(\gamma') \leq \frac{1}{4\pi}\,\text{length}(\gamma')^2 = \frac{1}{4}\,\text{length}(\gamma)^2,$$

as desired. □

References

[1] V. Blåsjö, The isoperimetric problem, Amer. Math. Monthly **112** (2005), no. 6, 526–566.

[2] J. L. Coolidge, Questions, Discussions, and Notes: A Historically Interesting Formula for the Area of a Quadrilateral, Amer. Math. Monthly **46** (1939), no. 6, 345–347.

[3] A. Treibergs, Inequalities that imply the isoperimetric inequality, notes available at `http://www.math.utah.edu/~treiberg/isoperim/isop.pdf`.

Andrew Putman
Department of Mathematics
Rice University, MS 136
6100 Main St.
Houston, TX 77005
andyp@math.rice.edu