The Isoperimetric Problem: Geometry and Analysis

Mark: 85%

Contents

1 Introduction

Putting in here, at this very place where you now see enormous
Ramparts, the rising castle of Carthage's new town, they purchased
Land, just as much as one could mark off with the hide of a single
Ox. And they called it Byrsa, 'The Hide', to recall the transaction.
- Aeneid Book 1, lines 365-368 [Ver07]

When Queen Dido - as per the tales of Roman mythology - founded the ancient North-African coastal city of Carthage, she was allowed by the local King only as much land as could be enclosed by the hide of one ox and the coast. To a first approximation, the problem is to find a curve with endpoints on a flat line that encloses (along with the flat line) the greatest area when compared to all other curves of the same length.

Somehow, it seems obvious that the solution to Dido's problem is a semi-circle - yet the point is to prove it. In this essay, however, we consider a directly related problem which is somehow more aesthetic:

Among all shapes with equal perimeter, which are the shapes that enclose the greatest area?

This problem is called the *isoperimetric problem* and often comes with associated inequalities - with characterisations of equality being the solution to the problem. Similarly to Dido's problem, it is somehow obvious that the solution is a circle, and yet again the point is to prove this conjecture.

Results about the isoperimetric problem have been known since Greek antiquity, and a fellow called Zenodorus, at some date between 200 B.C.E and

90 C.E., proved certain results which may indeed be regarded as satisfactory solutions to the isoperimetric problem [Hea14]:

Theorem 1. *Among all polygons with equal perimeter and equal number of sides, it is the regular polygon that encloses the largest area.*

Theorem 2. *The circle encloses greater area than any polygon of the same perimeter.*

In this essay we state these above theorems without proof and instead focus on the work of the previous few centuries, as this work gives associated inequalities which reveal some connections of the isoperimetric problem with topics in analysis that are, at first glance, disconnected from geometry.

In the first part of the essay, we will prove the isoperimetric problem for all C^1 curves using elementary techniques learned in the first year in Geometry and Motion.

In the second part of the essay, we will derive and explore a version of the Sobolev inequality, which is *a priori* a purely analytic inequality and does not seem related to geometry in the slightest. We will then, however, sketch a proof that (a slightly stronger version of) the Sobolev inequality is in fact equivalent to the isoperimetric inequality.

In the last part of the essay we will state some solutions to the isoperimetric problem for curves on the surface of a sphere, and discuss their implications for obtaining new insights on the isoperimetry of planar curves.

2 Parametrised Curves

In this section we will use the theory of parametrised curves that we learned in the first year in order to solve the isoperimetric problem for C^1 curves.

We begin this section with a very brief proof of a classical inequality, which the reader may have met in high school. We will use this "Arithmetic Mean - Geometric Mean" inequality later in the section.

Theorem 3 (The AM-GM Inequality)**.** *Let $a, b \in [0, \infty)$. Then*

$$\sqrt{ab} \leq \frac{a+b}{2},$$

with equality if and only if $a = b$.

Proof. The inequality follows by noticing the following fact:

$$0 \leq (a-b)^2$$

and so $0 \leq a^2 + b^2 - 2ab$, whence follows

$$4ab \leq a^2 + b^2 + 2ab,$$

and so $ab \leq (a+b)^2/4$. The inequality follows by taking square roots.

Now if $a = b$, then clearly equality holds. Conversely, if equality holds, then

$$4ab = a^2 + b^2 + 2ab$$

and so $(a-b)^2 = 0$, and so $a = b$. $\square$

Now, let us begin with the geometry. Recall, from Multivariable Calculus, Green's Theorem:

Theorem 4. *Let $\Omega \subset \mathbb{R}^2$ be a domain with C^1 boundary and let $(L, M) : U \to \mathbb{R}^2$ be a planar vector field on $U \supset \overline{\Omega}$. Furthermore, let $\gamma(s)$ be a regular arc-length parametrisation of $\partial\Omega$. Then*

$$\int\int_{\Omega} \partial_1 M - \partial_2 L \, \mathrm{d}A = \int_{\partial\Omega} (L, M) \cdot \gamma' \, \mathrm{d}s\,. \tag{1}$$

Choosing M and L cleverly, we have the following useful lemma, giving us a formula for the area of the region Ω:

Lemma 5. *Let $\gamma = (x, y) : [0, L) \to \mathbb{R}^2$ be a regular, positively oriented, arc-length, C^1 parametrisation of a closed, non-intersecting curve Γ. Let the region bounded by Γ be Ω, so that $\Gamma = \partial\Omega$. Then the area of Ω is given by*

$$A = \int_0^L xy' \, \mathrm{d}s = -\int_0^L yx' \, \mathrm{d}s = \frac{1}{2}\int_0^L (xy' - y'x) \, \mathrm{d}s\,.$$

Proof. For the first equality, apply Green's theorem with $L(x, y) = 0$ and $M(x, y) = x$. Then the left hand side of (1) is the area of Ω, and the right hand side is

$$A = \int_\Gamma (0, x) \cdot (x', y') \, \mathrm{d}s = \int_0^L xy' \, \mathrm{d}s,$$

as required.

For the second equality, we again apply Green's theorem but this time with $L(x, y) = -y$ and $M(x, y) = 0$. Then once again, the left hand side of (1) is the area of Ω, and the right hand side is

$$A = \int_\Gamma (-y, 0) \cdot (x', y') \, \mathrm{d}s = -\int_0^L yx' \, \mathrm{d}s,$$

as required.

For the last equality, we use the previous two equalities to obtain

$$A = \frac{1}{2}(A + A) = \frac{1}{2}\int_0^L (xy' - y'x) \, \mathrm{d}s,$$

as required. □

Now we may state and prove our solution to the isoperimetric problem.

Theorem 6 (The Isoperimetric Inequality)**.** *Let Γ be a closed, non-intersecting, C^1 curve with length L, and let A be the area of the region bounded by Γ. Then*

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

with equality if and only if C is a circle.

The following proof is taken from [Car76]. The main idea is to project a given convex curve onto a circle, and then measure how different they are using the Cauchy-Schwartz inequality. We then apply the AM-GM inequality (Theorem 3) to obtain the final inequality. After this, all that is left to do is to analyse the conditions of equality.

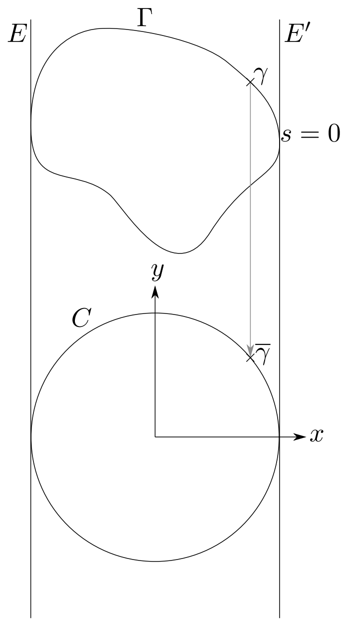

Figure 1: Our curve Γ, the two tangential lines E and E', the circle C, and the arc-length parametrisation γ, with the "projected" parametrisation $\bar{\gamma}$ of the circle. This figure was reproduced from [Car76].

Proof. Firstly, let us note that we may assume that Γ is convex, for if not then the convex hull of Γ would decrease the perimeter while increase the area.

Consider two vertical parallel lines E and E' on the left and right of Γ, and move them towards each other until they just touch Γ, being tangent to it. As in Figure 1, construct a circle C below curve Γ (such that Γ and C do not intersect) which has E and E' as tangents. Then the perpendicular distance between E and E' is $2r$, where r is the radius of C.

Impose a standard coordinate system with origin at the centre of the circle. Let $\gamma = (x, y) : [0, L) \to \mathbb{R}^2$ be a C^1 counterclockwise arc-length parametrisation of Γ with $\gamma(0) \in E'$, and so we may parametrise C with $\overline{\gamma} = (x, \overline{y}) : [0, L) \to \mathbb{R}^2$, again as in Figure 1. We therefore have the unit tangent to Γ being γ', and we obtain the inward unit normal by rotating γ' by $\pi/2$ to obtain $\nu = (-y', x')$.

Let the area bounded by Γ be A, and clearly the perimeter of Γ is L. Let also the area of C be $\overline{A} = \pi r^2$.

Now we observe the key idea of this proof. How are we going to obtain the inequality? Notice that if Γ is a circle, then the unit normal $\nu(s)$ is collinear to the position vector of the circle at that point $\bar{\gamma}(s)$. We recall that we measure collinearity using the Cauchy-Schwartz inequality, and so with this motivation behind us we apply this inequality to ν and $\bar{\gamma}$ to obtain

$$|\bar{\gamma} \cdot \nu| = \bar{\gamma} \cdot -\nu \leq |\bar{\gamma}||\nu| = \sqrt{x^2 + \bar{y}^2} = r, \tag{2}$$

with equality if and only if $\bar{\gamma}$ and ν are collinear. We also evaluate the dot product $\bar{\gamma} \cdot -\nu = xy' - \bar{y}x'$, and so we integrate both sides of the inequality (2) to obtain

$$\int_0^L xy' \,\mathrm{d}s - \int_0^L \bar{y}x' \,\mathrm{d}s \leq \int_0^L r \,\mathrm{d}s = Lr.$$

Now from Lemma 5 we have that

$$A = \int_0^L xy' \,\mathrm{d}s \qquad \text{and} \qquad \overline{A} = \pi r^2 = -\int_0^L \bar{y}x' \,\mathrm{d}s,$$

and so

$$A + \pi r^2 = \int_0^L xy' \,\mathrm{d}s - \int_0^L \bar{y}x' \,\mathrm{d}s \leq Lr, \tag{3}$$

with equality holding if and only if, as in inequality (2), ν and $\overline{\gamma}$ are collinear.

Now, by the AM-GM inequality (Theorem 3), we have that

$$\sqrt{A}\sqrt{\pi r^2} \leq \frac{1}{2}(A + \pi r^2) \leq \frac{1}{2}Lr, \tag{4}$$

and so

$$L^2 - 4\pi A \geq 0. \tag{5}$$

Now we analyse the conditions of equality. Equality in (5) is achieved if and only if equality is achieved in both inequalities of (4), which is achieved if and only if *both* on the one hand $A = \pi r^2$ *and* on the other hand ν and $\overline{\gamma}$ are collinear. It remains to show that these two conditions are equivalent to Γ being a circle.

Firstly suppose that Γ is a circle. Then clearly $A = \pi r^2$, and ν and $\overline{\gamma}$ are collinear.

Now suppose that $A = \pi r^2$, and that ν and $\overline{\gamma}$ are collinear. Then, since $|\nu| = 1$ and $|\overline{\gamma}| = r$, we have that $\overline{\gamma} = -r\nu$ (where we have the minus sign because ν is pointing in the opposite direction of $\overline{\gamma}$); i.e. $(x, \overline{y}) = -r(y', x')$. So $x = -ry'$. But $A = \pi r^2$, and since A is independent of the choice of E and E', then so is r. So rotating the whole setup (except for, of course, the curve Γ) by $\pi/2$, we have the change of coordinates $(x, y) \mapsto (-y, x)$. But since $x = -ry'$, under a change of coordinates we have that $-y = -rx'$, and so

$$x^2 + y^2 = r^2\left((x')^2 + (y')^2\right) = r^2$$

whence we conclude that Γ is a circle, as required. $\square$

In the above proof, we have assumed that the curve was C^1. It is worth noting, however, that the isoperimetric property of the circle holds even amongst *piecewise* C^1 curves. This is because we can approximate any such piecewise curve by a sequence of C^1 curves arbitrarily well.

3 Relations to PDE Theory: the Sobolev Inequality

In this section we will explore an analytic inequality due to Sobolev that holds for all continuously differentiable functions $f \in C^1(\mathbb{R}^2, \mathbb{R})$ with compact support. We first prove a weak form of the inequality that does not have a characterisation of equality. Using this, we motivate a connection between the Sobolev and isoperimetric inequalities and prove a version of the Sobolev inequality from the isoperimetric inequality, carrying with us a sharp constant and a characterisation of equality. Finally, we prove that this sharp[1] version of the Sobolev inequality is in fact equivalent to the isoperimetric inequality.

We will now present our first version of the Sobolev inequality. The following form of the Sobolev inequality is presented without a characterisation of equality, which is a necessary trade-off for the aesthetically simple proof of Nirenberg. The original proof is for a Sobolev equality in high dimensions, but here we keep the dimensions to two.

Theorem 7 (Sobolev, Nirenberg)**.** *If $f \in C^1(\mathbb{R}^2, \mathbb{R})$ with compact support, then*

$$\left(\int_{\mathbb{R}^2} |Df|\right)^2 \geq \int_{\mathbb{R}^2} |f|^2.$$

The following proof begins with a strange rewriting of a function f in the form of an integral, and by making some estimates and then integrating over the plane we are able to reduce the integrals to the desired result. Here we present a simplified version of the proof found in [Eva10].

Proof. As in the hypothesis, let $f \in C^1(\mathbb{R}^2, \mathbb{R})$ with compact support. Since f has compact support, all the integrals below are finite. Then by the Fundamental Theorem of Calculus, we may strangely write

$$f(x, y) = \int_{-\infty}^{x} D_x f(t, y) \mathrm{d}t,$$

and so we may impose the following estimates,

$$|f(x, y)| = \left| \int_{-\infty}^{x} D_x f(t, y) \mathrm{d}t \right| \leq \int_{-\infty}^{x} |D_x f(t, y)| \, \mathrm{d}t \leq \int_{-\infty}^{\infty} |Df(t, y)| \, \mathrm{d}t,$$

where the first inequality is the integral triangle inequality, and the second inequality comes from the fact that we are integrating a non-negative function over the whole plane and that the magnitude of a whole vector $|Df|$ is greater than or equal to the magnitude of just one component $|D_x f|$.

Now a similar inequality holds for the second variable,

$$|f(x, y)| \leq \int_{-\infty}^{\infty} |Df(x, s)| \, \mathrm{d}s \, .$$

Therefore

$$|f(x, y)|^2 \leq \int_{-\infty}^{\infty} |Df(x, s)| \, \mathrm{d}s \int_{-\infty}^{\infty} |Df(t, y)| \, \mathrm{d}t.$$

Now we integrate both sides of the above inequality twice - once with respect to each variable - giving

$$\begin{aligned} &\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)|^2 \mathrm{d}x \mathrm{d}y \\ &\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Df(x, s)| \mathrm{d}s \mathrm{d}x \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Df(t, y)| \mathrm{d}t \mathrm{d}y \end{aligned}$$

[1]A sharp constant in an inequality is the best possible constant that keeps the inequality true.

where the rearrangement after the equality is possible because of considerations about dependence of variables. With nicer notation, we have

$$\left(\int_{\mathbb{R}^2} |Df|\right)^2 \geq \int_{\mathbb{R}^2} |f|^2$$

which is the required result. $\square$

Now we see that the isoperimetric and Sobolev inequalities are in fact equivalent. Firstly we prove that the isoperimetric inequality implies the Sobolev inequality, carrying the sharp constant with us as we go along. This gives us, unlike in Theorem 7, an interesting characterisation of equality. We then prove that the Sobolev inequality with this sharp constant implies the isoperimetric inequality. The ideas in the following sketches of proofs are taken almost entirely from the reference [Oss78], but here we present them in a more accessible way. For the proof that the isoperimetric inequality implies the Sobolev inequality, we also sketch out a proof of a simplified version of the *co-area formula*, taking inspiration from [Maz03]. We will not be giving fully rigorous proofs as these are beyond the scope of this essay. An outline of the main ideas, however, is indeed possible.

Theorem 8. *Let $\mathcal{C}$ be the set of al closed, non-intersecting, C^1 plane curves, and let $\mathcal{F}$ be the set of all functions in $C^1(\mathbb{R}^2, \mathbb{R})$ with compact support. Then*

$$\inf_{f \in \mathcal{F}} \frac{\left(\int_{\mathbb{R}^2} |Df|\right)^2}{\int_{\mathbb{R}^2} |f|^2} = \inf_{\Gamma \in \mathcal{C}} \frac{L^2(\Gamma)}{A(\Gamma)}.$$

The key takeaway here is that the constants of the inequalities are the same. We prove the above theorem by showing that the isoperimetric and Sobolev inequalities are equivalent; that is, we assume the isoperimetric inequality, and from that deduce a Sobolev inequality *with the required constant.* We then assume the Sobolev inequality, and deduce the isoperimetric *with the same required constant.*

Sketch of proof (isoperimetric $\implies$ Sobolev): Let $f \in C^1(\mathbb{R}^2, \mathbb{R})$ have compact support in some $\Omega \subset \mathbb{R}^2$. As in Figure 2, let us, for each $t \geq 0$, consider the level sets

$$\Gamma(t) := \{x \in \mathbb{R}^2 : |f(x)| = t\}.$$

Firstly we wish to specify for which $t \geq 0$, $\Gamma(t)$ is a curve. Let $t \geq 0$. From Multivariable Calculus, we know that in order to have the level set $\Gamma(t)$ being a one-dimensional submanifold of $\mathbb{R}^3$, we need $\text{rank}(Df(t)) = 1$; that is, $Df(t) \neq 0$. By a theorem of Sard (which we shall not prove here), almost all values of t are not critical points of f. Therefore almost all values of t define a one-dimensional submanifold; that is for almost all values of t, $\Gamma(t)$ is a piecewise-C^1 level set. Sard's theorem effectively lets us ignore these "bad" values of t (that do not produce a one-dimensional submanifold) when we integrate over all values of t. From now on, let us only consider the "good" values of t (that *do* produce one-dimensional submanifolds).

Now, for each $t \geq 0$ we consider the domains, again as in Figure 2,

$$\Omega(t) := \{x \in \mathbb{R}^2 : |f(x)| > t\};$$

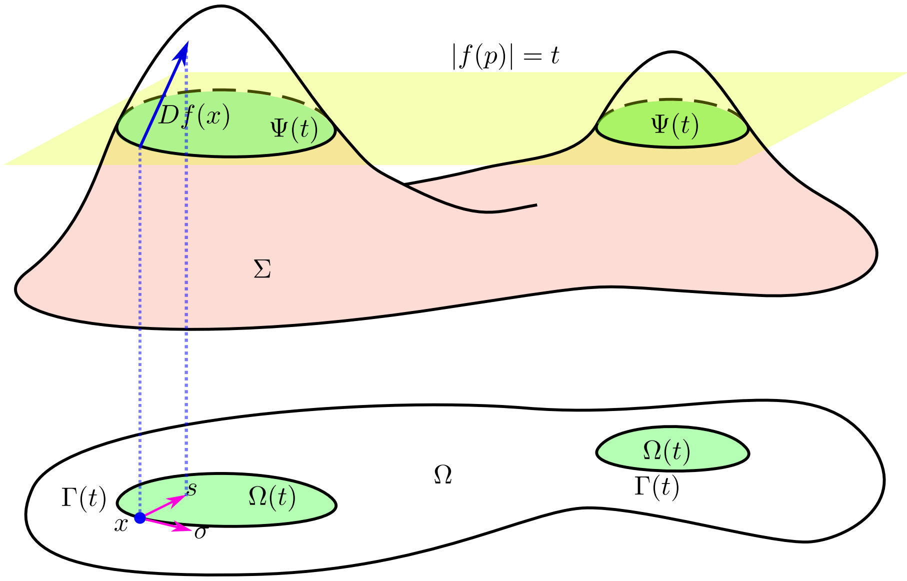

Figure 2: The graph of f with domain Ω, the horizontal plane $f(p) = t$, the level sets $\Gamma(t)$, the parameters s and σ, the vector of greatest ascent $Df(x)$, and the region Σ. This figure was the author's own work.

it is clear that for each $t \geq 0$, $\Omega(t)$ is the interior of $\Gamma(t)$. We denote by $A(t)$ the area of $\Omega(t)$, and by $L(t)$ the length of $\Gamma(t)$. We have now finished describing all the geometrical objects that we need to consider in this proof.

In the first part of the proof, we wish to show that

$$\int_{\Omega} |Df| = \int_0^{\infty} L(t)\,\mathrm{d}t\,.$$

This is a very simplified version of a more general result known as the *co-area formula* that holds for more general functions with domain and range in higher dimensions. The co-area formula is a result of geometric measure theory and can be studied in, for example, [Mor16].

Now, let $x \in \Omega$. Consider the level set $\Gamma(t)$ that contains x. We parametrise $\Gamma(t)$ by some arc-length parameter $\sigma(t)$ that depends on t. Consider the vector of greatest ascent of the graph of f at x; recall that this vector is given by the derivative of f at x, $Df(x)$. As in Figure 2, we introduce the line element $\mathrm{d}s(\sigma)$ (s being a function of σ) which points in the direction of the projection of $Df(x)$ down onto $\Omega(t)$. Therefore s is the parameter in Ω along the line of greatest ascent along f projected down to Ω, and so the line element $\mathrm{d}s$ is orthogonal to the line element $\mathrm{d}\sigma$. Therefore the magnitude of greatest ascent is given by the magnitude of the derivative of f with respect to this parameter s:

$$|Df| = \left|\frac{\mathrm{d}f}{\mathrm{d}s}\right| = \left|\frac{\mathrm{d}f}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}s}\right| = \left|\frac{\mathrm{d}t}{\mathrm{d}s}\right|,$$

where the second equality follows by the chain rule and the third equality follows from the fact that f increases with t at the same rate, and so $|\frac{\mathrm{d}f}{\mathrm{d}t}| = 1$. Now, as s increases, t increases, and so in fact we have $|Df| = \frac{\mathrm{d}t}{\mathrm{d}s}$; that is,

$$\mathrm{d}s = |Df|^{-1}\,\mathrm{d}t\,.$$

Now, since $\mathrm{d}s$ and $\mathrm{d}\sigma$ are orthogonal, we have the area element at x being given by

$$\mathrm{d}A = \mathrm{d}s\,\mathrm{d}\sigma = |Df|^{-1}\,\mathrm{d}t\,\mathrm{d}\sigma,$$

and so we have

$$\int_\Omega |Df|\,\mathrm{d}A = \int_\Omega \mathrm{d}t\,\mathrm{d}\sigma\,.$$

Now we wish to evaluate the integral on the right hand side. Since our choice of parameter σ depends on t, we have to integrate with respect to σ first, and then we may integrate with respect to t. So we have

$$\int_\Omega |Df| = \int_\Omega \mathrm{d}\sigma\,\mathrm{d}t = \int_0^\infty \int_{\Gamma(t)} \mathrm{d}\sigma\,\mathrm{d}t = \int_0^\infty L(t)\,\mathrm{d}t,$$

which concludes the first part of the proof.

In the second part of the proof, we use the isoperimetric inequality to deduce the Sobolev inequality with the required constant.

Now we apply the isoperimetric inequality to each $\Omega(t)$; we have $L(t) \geq \sqrt{4\pi A(t)}$ with equality if and only if $\Omega(t)$ are circles. Therefore

$$\int_\Omega |Df| = \int_0^\infty L(t)\,\mathrm{d}t \geq \sqrt{4\pi}\int_0^\infty \sqrt{A(t)}\,\mathrm{d}t, \tag{6}$$

with equality if only if $\Omega(t)$ are circles.

So we have expressed the right-hand side of the Sobolev inequality. Now let us express the left-hand side of the Sobolev inequality. We integrate f^2 over the domain Ω to obtain

$$\begin{aligned}\int_\Omega f^2 &= \int\int_\Omega \left[\int_0^{|f(x,y)|} 2t\,\mathrm{d}t\right]\mathrm{d}x\,\mathrm{d}y \\ &= \int_0^\infty 2t\left[\int\int_{t\leq|f(x,y)|} 1\,\mathrm{d}x\,\mathrm{d}y\right]\mathrm{d}t = \int_0^\infty 2tA(t)\,\mathrm{d}t\,.\end{aligned}$$

Now before moving on, let us quickly explain the second and third integral expressions above. The first equality we find by integrating the derivative of f^2. We may then view this integral as an integral over the domain $\Sigma \subset \mathbb{R}^3$ as in Figure 2 defined as

$$\Sigma = \left\{(x,y,z) \in \mathbb{R}^2 : (x,y) \in \Omega \text{ and } z \in [0, |f(x,y)|]\right\},$$

which is precisely the domain in $\mathbb{R}^3$ that is bounded by the xy-plane and the graph of $|f|$. So we have two ways of integrating over this volume. The second integral expression firstly integrates over permissible values of z for each (x,y), and then integrates over permissible values of (x,y). On the other hand, the third integral expression integrates over the set

$$\{(x,y) \in \Omega : |f(x,y)| \geq t\}$$

for each t. This set is precisely the closure of the domain $\overline{\Omega(t)}$. So now think of t going from 0 to ∞, and for each t we have the surface $\Psi(t)$, again as in Figure 2, given by

$$\Psi(t) = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in \overline{\Omega(t)} \text{ and } z = t\}$$

which is just the closure of the domain $\overline{\Omega(t)}$ translated in the z-axis to corresponding value of t. It is clear then that the union over all t of $\Psi(t)$ is just Σ. So the second and the third integral expressions are integrating the same integrand over the same domain, and so they are equal.

We therefore have found that

$$\int_\Omega f^2 = \int_0^\infty 2tA(t)\,\mathrm{d}t\,. \tag{7}$$

Now since A is a decreasing function, we have that

$$t\sqrt{A(t)} \leq \int_0^t \sqrt{A(\tau)}\,\mathrm{d}\tau,$$

where equality holds if $A(t)$ is constant on some finite interval $[0, M]$ and zero thereafter. Multiplying both sides by $\sqrt{A(t)}$, we have that

$$tA(t) \leq \sqrt{A(t)} \int_0^t \sqrt{A(\tau)}\,\mathrm{d}\tau = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t \sqrt{A(\tau)}\,\mathrm{d}\tau\right]^2,$$

with the same condition of equality. Integrating both sides, we have that

$$\int_0^\infty 2tA(t)\,\mathrm{d}t \leq \left[\int_0^\infty \sqrt{A(t)}\,\mathrm{d}t\right]^2, \tag{8}$$

with equality if A is constant on some interval $[0, M]$ and zero thereafter. Combining the inequalities and equalities (6), (7), and (8) we have that

$$\left(4\pi \int_\Omega f^2\right)^{\frac{1}{2}} \leq \int_\Omega |Df|,$$

with equality if for every $t \geq 0$, $\Omega(t)$ is a circle, and the function A is constant on some interval $[0, M]$ and zero thereafter. Since f has compact support on Ω, we may extend the domain of the above integrals to the whole of $\mathbb{R}^2$ without changing the result. Since we have a condition of equality, we have that

$$\inf_{f \in \mathcal{F}} \frac{\left(\int_{\mathbb{R}^2} |Df|\right)^2}{\int_{\mathbb{R}^2} |f|^2} = 4\pi.$$

□

Before we move on to a sketch of a proof of the converse, let us first discuss this constant of 4π. The reader may notice that the characterisation of equality is impossible, because if for any t the function $A(t)$ is constant and $\Omega(t)$ is a circle, then the graph of f would have to look like a cylinder. This is clearly impossible for a graph of a function $\mathbb{R}^2 \supset \Omega \to \mathbb{R}$. We can, however, approximate this graph by a one-parameter family of truncated cones as in Figure 3. If the graph of a function is a truncated cone, then this function is only piecewise C^1;

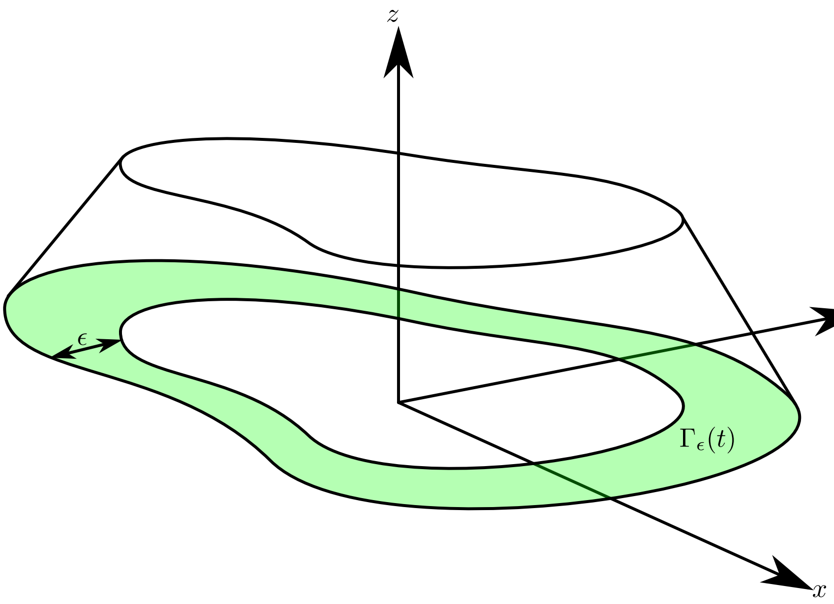

Figure 3: The graph of the function f_ϵ, along with the region $\Gamma_\epsilon(t)$. As we decrease ϵ, the graph of f_ϵ approximates ever-better a single cylinder. This figure was the author's own work.

but again we may approximate it arbitrarily well by a family of C^1 functions. It is tempting, therefore, to make the conjecture that this family of truncated cones can get arbitrarily close to equality in the Sobolev inequality. We will not prove this conjecture in this essay, but we may treat it as a plausible result.

In the proof of the converse, we do indeed use this idea of approximating a cylinder with a truncated cone. The following proof is also taken from [Oss78].

Sketch of proof (Sobolev $\implies$ isoperimetric): Let $\Omega \subset \mathbb{R}^2$ be a compact domain with C^1 boundary $\partial\Omega =: \Gamma$. The idea is to apply the Sobolev inequality to the characteristic function χ_Ω of Ω, but as this is not a continuous function (let alone C^1), we approximate it by a family of piecewise-C^1 functions (which can themselves be approximated by C^1 functions) $f_\epsilon : \mathbb{R}^2 \to \mathbb{R}$ with compact support in Ω such that, for small $\epsilon > 0$ and $p \in \Omega$,

$$f_\epsilon(p) := \begin{cases} 1 & \text{if } d(p,\Gamma) \geq \epsilon \\ \frac{d(p,\Gamma)}{\epsilon} & \text{if } d(p,\Gamma) < \epsilon, \end{cases}$$

where $d(p,\Gamma)$ is the distance from p to the boundary Γ, as in Figure 3.

Now, as $\epsilon \to 0$, $f_\epsilon \to \chi_\Omega$, and so we find an expression for the right-hand side of the isoperimetric inequality in terms of the right-hand side of the Sobolev inequality:

$$\int_\Omega f_\epsilon^2 \to \int_\Omega 1 = A$$

where A is the area of Ω.

For the left hand side, we have that

$$|Df_\epsilon| = \begin{cases} \frac{1}{\epsilon} & \text{in } \Omega \cap \Gamma_\epsilon \\ 0 & \text{in } \Omega \setminus \Gamma_\epsilon \end{cases}$$

where $\Gamma_\epsilon := \{p \in \Omega : d(p, \Gamma) < \epsilon\}$. Therefore

$$\int_\Omega |Df_\epsilon| = \frac{\text{Area}(\Omega \cap \Gamma_\epsilon)}{\epsilon},$$

and as $\epsilon \to 0$ it is tempting to conflate the right hand side with the length of Γ, which we may call L. This can be formalised using a notion of measure called *Minkowski content* [Oss78].

Now the Sobolev inequality implies that for each $\epsilon > 0$,

$$\left(4\pi \int_\Omega f_\epsilon^2\right)^{\frac{1}{2}} \leq \int_\Omega |Df_\epsilon|$$

and so taking the limit as $\epsilon \to 0$ we may be tempted to surmise that $\sqrt{4\pi A} \leq L$, or in other words, that

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

But equality in the Sobolev inequality holds if Ω is a circle, and so we have equality in the isoperimetric inequality if Ω is a circle. And so we have that

$$\inf_{\Gamma \in \mathcal{C}} \frac{L^2(\Gamma)}{A(\Gamma)} = 4\pi.$$

□

The reader may notice our use of words and phrases such as "tempted" and "surmise" and "plausible results". This use of language is to lay bare the lack of rigour in the argument, but full rigour - which is outside the scope of this essay - may be found through studying Minkowski content and the argument presented in the reference [Oss78]. Yet the key ideas are outlined above; all that is necessary is to fill in the gaps.

4 For the Sphere

In this section, we discuss various isoperimetric inequalities for the sphere and their implications for the plane. We will not give proofs of these theorems as they are outside the scope of this essay, but references will be of course provided.

At the beginning of the essay, we motivated the isoperimetric problem using the story of Queen Dido from ancient mythology. Dido, however, did unfortunately not have her problem solved by the results thus far in this essay. Earth is after all a sphere, and these theorems hold only in the plane. The theorem Dido was looking for was

Theorem 9 ([Oss78])**.** *Let Γ be a closed non-intersecting C^1 curve on the sphere S_R^2, with perimeter L and area A. Then*

$$L^2 \geq 4\pi A - \frac{A^2}{R^2},$$

with equality if and only if Γ is a circle.

An interesting version of this theorem was proved by Bernstein in a paper in 1905 [Ber05]; he gave the inequality

$$L^2 - 4\pi A + \frac{A^2}{R^2} \geq 4R^2 \sin^2\left(\frac{d}{4R(1+2\pi)}\right)\left(2\pi + \sin^2\left(\frac{d}{4R(1+2\pi)}\right)\right)$$

where d is smallest of all widths of circular annuli that contain the curve Γ. Interestingly, taking the limit as $R \to \infty$, we obtain the following isoperimetric inequality for curves on the plane:

$$L^2 - 4\pi A \geq \frac{\pi d^2}{2 + 4\pi}.$$

The quantity $L^2 - 4\pi A$ is called the *isoperimetric deficit* of the curve ([Oss78]) as it measures how far off the curve Γ is from the circle in terms of bounding the largest area with the smallest perimeter. This inequality, therefore, gives us a lower bound on how "un-isoperimetric" a given curve is. This lower bound, however, may not be the best possible lower bound.

In the 1920s, Bonnesen proved the following improvement on that of Bernstein [Bon29]:

$$L^2 - 4\pi A \geq 4\pi d^2,$$

in which the constant 4π Bonnesen showed to be the optimal one.

5 Concluding Remarks

In this essay, we solved the isoperimetric problem in the plane using elementary techniques from first-year Geometry and Motion. We have also seen that the Sobolev inequality, which is firstly derived in a purely analytic way, has deep connections to geometry in that it is equivalent to the isoperimetric inequality.

The isoperimetric inequality is an ancient inequality, but it is still a motivator for contemporary research. It is used as a litmus test for new techniques, and any contemporary mathematics that can shine a light on a theorem first studied millennia ago has surely a promising future.

References

[Ber05] Felix Bernstein. "Über die isoperimetrische Eigenschaft des Kreises auf der Kugeloberfläche und in der Ebene". In: *Math. Ann.* 60 (1905), pp. 117–136.

[Bon29] Tommy Bonnesen. *Les problèmes des Isopérimètres et des Isépiphanes.* Paris: Gauthier-Villars, 1929.

[Car76] Manfredo P. do Carmo. *Differential Geometry of Curves and Surfaces.* First Edition. Englewood Cliffs, New Jersey: Prentice-Hall, 1976.

[Eva10] Lawrence C. Evans. *Partial Differential Equations.* First Edition. Providence, Rhode Island: American Mathematical Society, 2010.

[Hea14] Thomas L. Heath. *A History of Greek Mathematics, Volume 2.* Cambridge: Cambridge University Press, 2014.

[Maz03] Vladimir Maz'ya. "Lectures on isoperimetric ande isocapacitary inequalities in the theory of Sobolev spaces." In: *Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002).* Vol. 338. Contemp. Math. Providence, RI: Amer. Math. Soc., 2003, pp. 307–340.

[Mor16] Frank Morgan. *Geometric Measure Theory.* Fifth Edition. London: Academic Press, 2016.

[Oss78] Robert Osserman. "The Isoperimetric Inequality". In: *Bull. Amer. Math. Soc.* 84.6 (1978), pp. 1182–1238.

[Ver07] Vergil. *Aeneid.* Trans. by Frederick Ahl. Oxford: Oxford University Press, 2007.