A Beautiful Inequality by Saint-Venant and Pólya Revisited

Arno Berger

Abstract. In mathematical physics and beyond, one encounters many beautiful inequalities that relate geometric or physical quantities describing the shape or size of a set. Such *isoperimetric* inequalities often have a long history and many important applications. For instance, the eponymous and most classical of all isoperimetric inequalities was known already in antiquity. It asserts that among all closed planar curves of a given length, the circles with perimeter equal to that length, and only they, enclose the largest area. Though not nearly as well-known, an isoperimetric inequality conjectured by Saint-Venant in the 1850s and first proved by Pólya almost a century later, is also very beautiful and important. By presenting a short proof as well as two simple physical interpretations, this article illustrates why the result deserves to be cherished by every student of applied analysis.

1. INTRODUCTION. The inequality of the title is an example of an *isoperimetric* inequality, that is, a sharp inequality that relates geometric or physical quantities describing the shape or size of a set [**20**, Sec. 1]. Despite its beauty and importance, the result does not seem to be altogether well-known. This article attempts to change that.

To set the scene, fix an integer $d \geq 2$ and a non-empty open subset U of the Euclidean space $\mathbb{R}^d$. For convenience, say that the set U is **nice** if it is bounded, and its boundary ∂U is sufficiently well-behaved for the Dirichlet boundary value problem

$$\Delta w := \frac{\partial^2 w}{\partial x_1^2} + \ldots + \frac{\partial^2 w}{\partial x_d^2} = -1 \text{ in } U\,, \quad w = 0 \text{ on } \partial U\,, \tag{1}$$

to have a unique (classical) solution $w = w(x) = w(x_1, \ldots, x_d)$. As one learns in an introductory course on partial differential equations (PDE), this is rather a mild assumption. For example, U is nice provided that each point in ∂U is the tip of a small cone contained entirely in the complement $\mathbb{R}^d \setminus U$; in particular, U is nice whenever ∂U is a smooth surface [**12**, Sec. 2.8]. The maximum principle, another fixture in every introductory PDE course, implies that $w > 0$ in U. Therefore, the integral

$$J(U) := \int_U w(x)\,\mathrm{d}x\,,$$

the main protagonist of this article, is positive (and finite). Notice that just like the d-dimensional volume (or Lebesgue measure) $m(U)$, the value of $J(U)$ does not change if U is subjected to any rotation, translation, or reflection. Also, for $a > 0$ and the scaled copy $aU := \{ax : x \in U\}$ of U, clearly $m(aU) = a^d m(U)$ whereas $J(aU) = a^{d+2}J(U)$. Thus it is natural to consider the ratio $J(U)/m(U)^{1+2/d} > 0$, the value of which does not change if U is scaled either. To develop a quantitative sense for this ratio, first consider a simple example.

Example 1. Given $0 < a < b$, let U be the annular region $\{x \in \mathbb{R}^d : a < |x| < b\}$, hence $m(U) = (b^d - a^d)\omega_d$. Here and throughout, $|\cdot|$ and ω_d denote the Euclidean length and the volume of the unit ball in $\mathbb{R}^d$, respectively. From multi-variable calculus, the reader may recall that $\omega_d = \pi^{d/2}/\Gamma(1 + d/2)$ where Γ is the Euler Gamma function. The solution of (1) is radially symmetric; more precisely, $w(x) = W(|x|)$ with $W = W(r)$ being the unique solution of

$$\frac{\mathrm{d}^2W}{\mathrm{d}r^2} + \frac{d-1}{r}\frac{\mathrm{d}W}{\mathrm{d}r} = -1\,, \quad W(a) = W(b) = 0\,.$$

From this, a fun calculus exercise yields a smooth positive function C_d such that

$$J(U) = C_d\left(\frac{a}{b}\right) m(U)^{1+2/d}\,.$$

While the general expression for $C_d(s)$ is somewhat cumbersome to write down, it is readily checked that, for instance,

$$C_2(s) = \frac{1}{8\pi}\left(\frac{1+s^2}{1-s^2} + \frac{1}{\log s}\right) \quad \text{and} \quad C_4(s) = \frac{1}{12\pi\sqrt{2}}\left(\frac{1-s^2}{1+s^2}\right)^{3/2}. \tag{2}$$

Just as in these examples, C_d is decreasing in $0 < s < 1$ for every $d \geq 2$, with

$$c_d := \lim_{s\downarrow 0} C_d(s) = \frac{1}{\omega_2^{2/d} d(d+2)} = \frac{\Gamma(1+d/2)^{2/d}}{\pi d(d+2)}\,,$$

and $\lim_{s\uparrow 1} C_d(s) = 0$; see also Figure 1. Clearly, then, $J(U)/m(U)^{1+2/d}$ never is larger than c_d. Moreover, it seems plausible that $c_d = J(U)/m(U)^{1+2/d}$ whenever U is a ball. This indeed is the case; see Lemma 6 below. However, notice that $J(U)$ is not defined for $a = 0$ since U, a punctured ball, is not a nice open set.

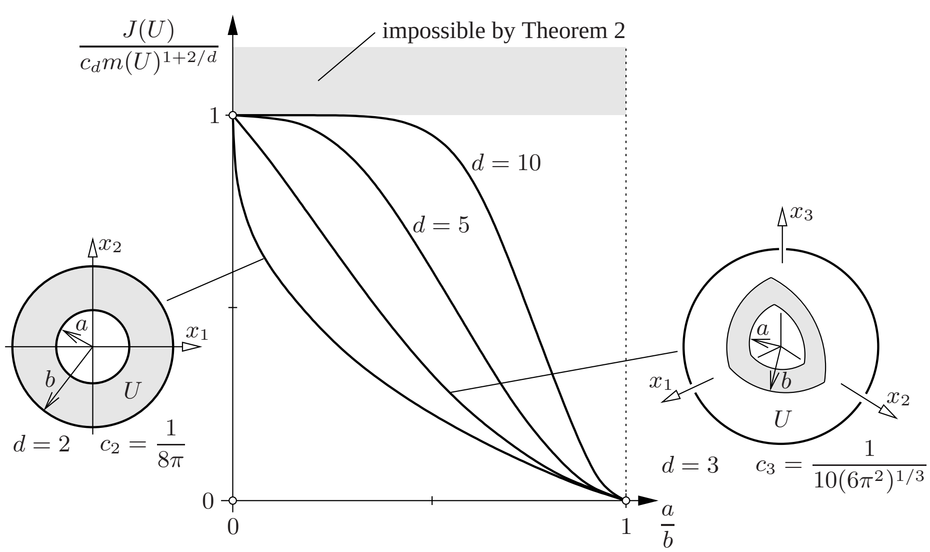

Figure 1. The ratio $J(U)/(c_d m(U)^{1+2/d})$ for the annular region $U = \{x \in \mathbb{R}^d : a < |x| < b\}$ is decreasing in $s = a/b$ and, as asserted by Theorem 2, is smaller than 1 for all $0 < a < b$.

Example 1 demonstrates that the ratio $J(U)/m(U)^{1+2/d}$ can be arbitrarily small. It also suggests that, by contrast and perhaps somewhat surprisingly, this ratio cannot be arbitrarily large. On first sight, it may seem as if the upper bound gleaned from Example 1 was specific for the particularly nice open sets considered there. This, however, is not the case. In fact, it is the purpose of this article to advertise, and present a proof of, the following inequality which identifies c_d as a universal sharp upper bound.

Theorem 2. *Let $d \geq 2$ be an integer, and $U \subset \mathbb{R}^d$ a nice open set. Then*

$$J(U) \leq c_d m(U)^{1+2/d}\,; \tag{3}$$

moreover, equality holds in (3) *if and only if U is a ball.*

As the reader may suspect right away, and will soon see explicitly, the key step in the proof of Theorem 2 is to establish (3) for *connected* U, and to address possible equality in this case. For $d = 2$ and *simply connected* U, that is, for planar connected open sets "without holes," Theorem 2 was, in essence, conjectured in 1856 by J.B. de Saint-Venant [**24**] who also supported it with ample experimental evidence based on a physical interpretation of $J(U)$; see Section 3 below. Saint-Venant's conjecture was first proved in 1948 by G. Pólya [**21**].

As far as the author has been able to ascertain, Theorem 2 appears, more or less explicitly, only in the specialized PDE literature (such as, e.g., [**6, 9, 11**]) where it is proved in much greater generality by means of advanced techniques that may be difficult to appreciate for the non-expert. As formulated in this article, however, Theorem 2 follows directly from a simple form of Talenti's inequality, an important basic tool in PDE theory. Section 2 recalls a tailor-made version of this classical inequality, and then gives a short proof of Theorem 2. Section 3 discusses two pertinent physical interpretations of $J(U)$, as well as variations on Theorem 2, for the case $d = 2$.

Remark. The definition of $J(U)$ makes sense for $d = 1$ as well. Every bounded (nonempty) open set $U \subset \mathbb{R}$ is nice, and a short calculus exercise, aided perhaps by Lemma 5 below, confirms that the conclusion of Theorem 2, with $c_1 = \frac{1}{12}$, remains correct in this case also.

2. PROOF OF THEOREM 2. As indicated earlier, utilizing the appropriate tools makes the proof of Theorem 2 very short indeed. For the task at hand, arguably the most appropriate tools are *symmetrizations*, also referred to as *rearrangements*; see, e.g., [**3, 8, 14, 16**] for friendly, informative introductions to the subject. As it turns out, a few simplified concepts and facts are all that is needed here. They are reviewed first.

Given a nice open set $U \subset \mathbb{R}^d$, or indeed any set with finite volume, denote by U^* the unique (open) ball centered at the origin with $m(U^*) = m(U)$, that is, let

$$U^* := \left\{ x \in \mathbb{R}^d : |x| < (m(U)/\omega_d)^{1/d} \right\}.$$

For every bounded continuous function $f : U \to \mathbb{R}$ and every real number t, denote the level set $\{x \in U : f(x) > t\}$ of f simply by $\{f > t\}$. The function $f^* : U^* \to \mathbb{R}$ given by

$$f^*(x) := \sup\left\{ t \in \mathbb{R} : x \in \{f > t\}^* \right\} \tag{4}$$

is the (**Schwarz**) **symmetrization** or **symmetric decreasing rearrangement** of f. Although the definition (4) may look a bit daunting on first sight, observe that, informally put, f^* simply is a more symmetric version of f which, however, retains certain key aspects of the latter function, notably its range, distribution of values, and integral; see Proposition 3 below. In the suggestive words of Pólya and Weinstein in [**23**], if f

describes a "hill" over U then f^* describes a rearranged "round hill" over U^*, where each level set has been converted into a ball of the same volume, and all these balls are concentric; see also Figure 2. A few basic properties of symmetrizations are easily deduced from (4), but can also be found, among many other interesting facts, e.g., in [**3, 8, 16**].

Proposition 3. Let $d \geq 2$ be an integer, and $U \subset \mathbb{R}^d$ a nice open set. Assume that $f : U \to \mathbb{R}$ is bounded and continuous. Then:

(i) f^* is radially symmetric and decreasing, i.e., $f^*(x) = f^*(y)$ whenever $x, y \in U^*$ with $|x| = |y|$, and $f^*(x) \leq f^*(y)$ if $|x| \geq |y|$;

(ii) f^* is bounded, with $\inf_{U^*} f^* = \inf_U f$ and $\sup_{U^*} f^* = \sup_U f$;

(iii) f^* is continuous, provided that U is connected;

(iv) $\{f^* > t\} = \{f > t\}^*$ for every $t \in \mathbb{R}$, and hence also $m(\{f^* > t\}) = m(\{f > t\})$;

(v) $\displaystyle\int_{U^*} f^*(x)\,\mathrm{d}x = \int_U f(x)\,\mathrm{d}x$.

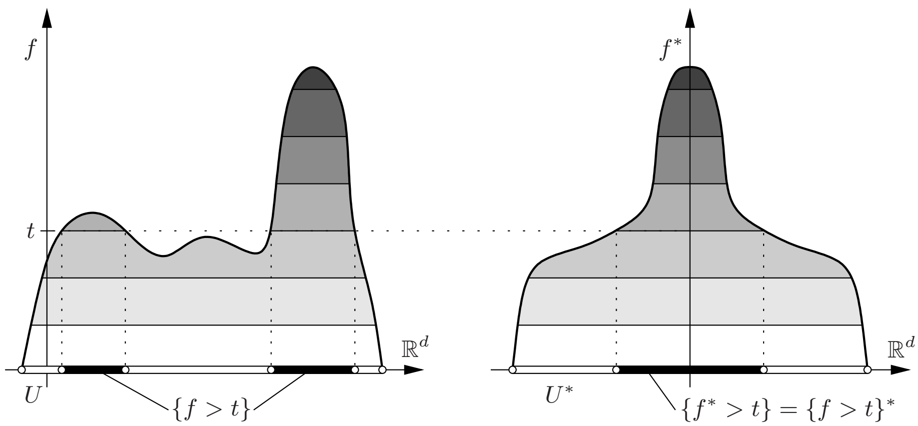

Figure 2. Illustrating the symmetrization $f^* : U^* \to \mathbb{R}$ of a function $f : U \to \mathbb{R}$.

With regard to Theorem 2, a key question is whether or not symmetrization interacts in a fruitful way with solving the Dirichlet problem (1). More specifically, is the solution w of the latter related in any way to the solution v of

$$\Delta v = -1 \ \text{ in } U^*, \quad v = 0 \ \text{ on } \partial U^* \ ? \tag{5}$$

Note that $(-1)^* = -1$, and keep in mind that U^* simply is a ball centered at the origin. It is a most fortunate feature of the Laplace operator Δ, and indeed of more general elliptic second-order differential operators that this question has an affirmative answer. Formulated specifically with Theorem 2 in mind, the following is a simple special case of *Talenti's inequality*. (After having witnessed this powerful tool contribute decisively to the proof presented below, a reader yearning for an in-depth study may want to turn, e.g., to [**2, 15**] for details.) Recall that a statement holds for *almost all* $x \in U^*$ if all x for which the statement does fail are contained in a subset of U^* with volume zero.

Proposition 4. Let $d \geq 2$ be an integer, and $U \subset \mathbb{R}^d$ a connected nice open set. Assume that w and v are the solutions of the Dirichlet problems (1) and (5), respectively. Then $w^*(x) \leq v(x)$ for almost all $x \in U^*$; moreover, $w^*(x) = v(x)$ for almost all $x \in U^*$ if and only if U is a ball.

In addition to this powerful fact, two elementary observations are helpful. A first observation is about the quantity $N_p(z) := \left(\sum_{n=1}^{\infty} |z_n|^p\right)^{1/p}$, where $z = (z_n)$ is any real sequence, and $p > 0$. For $p \geq 1$, and restricted to the appropriate space of sequences, the reader may recognize N_p as the p-norm familiar from linear analysis.

Lemma 5. Let $z = (z_n)$ be a real sequence. If $N_{p_o}(z) < \infty$ for some $p_o > 0$ then $N_p(z) < \infty$ for all $p \geq p_o$, and $p \mapsto N_p(z)$ is strictly decreasing on $[p_o, \infty[$ unless $z_n = 0$ for all but at most one n.

Proof. If $N_{p_o}(z) < \infty$ then $\lim_{n\to\infty} z_n = 0$ and hence $|z_n|^p \leq |z_n|^{p_o}$ for every $p \geq p_o$ and all sufficiently large n, so $N_p(z) < \infty$ as well. Since the assertion is correct for $z = 0$, assume that $z_n \neq 0$ for at least one n. Pick any $q > p \geq p_o$, and assume first that $N_p(z) = 1$. Then $|z_n|^q \leq |z_n|^p$ for all n, and in fact $|z_n|^q < |z_n|^p$ unless $z_n \in \{-1, 0, 1\}$. Summing over n yields $N_q(z) < N_p(z)^{p/q} = 1$ unless $|z_n| = 1$ for some, necessarily unique n. In general, that is, for $N_p(z) \neq 1$ simply apply this argument to the sequence $\big(z_n/N_p(z)\big)$ instead of (z_n). This yields $N_q(z)/N_p(z) < 1$ unless $z_n = 0$ for all but one n. ■

A second observation is that (3) is an equality whenever $U \subset \mathbb{R}^d$ is a ball. This, the "if" part of the assertion regarding equality in Theorem 2, could be established by carefully considering the case $a \downarrow 0$ in Example 1. Still simpler is a direct calculation.

Lemma 6. Let $d \geq 2$ be an integer, and $U \subset \mathbb{R}^d$ a ball. Then $J(U) = c_d m(U)^{1+2/d}$.

Proof. By the translation invariance of $J(U)$, it can be assumed that U equals $\{x \in \mathbb{R}^d : |x| < a\}$ for some $a > 0$. Then $m(U) = a^d \omega_d$, as well as $w(x) = (a^2 - |x|^2)/(2d)$, and hence

$$J(U) = \int_{|x|<a} \frac{a^2 - |x|^2}{2d}\,\mathrm{d}x = \frac{\sigma_d}{2d}\int_0^a (a^2 - r^2) r^{d-1}\,\mathrm{d}r = \frac{a^{d+2}\sigma_d}{d^2(d+2)}\,,$$

where σ_d denotes the $(d-1)$-dimensional volume of $\{x \in \mathbb{R}^d : |x| = 1\}$, the unit sphere in $\mathbb{R}^d$, and the second equality is an application of spherical coordinates or, more grandiosely put, the *co-area formula*; see, e.g., [**16**, Sec. 2.2]. As the reader may recall from multi-variable calculus, $\sigma_d = d\omega_d$, and so indeed

$$J(U) = \frac{a^{d+2}\omega_d}{d(d+2)} = \frac{m(U)^{1+2/d}}{\omega_d^{2/d} d(d+2)} = c_d m(U)^{1+2/d}\,.$$

(When interpreted with caution, this calculation is correct for $d = 1$ also.) ■

The scene is now set for a very short *Proof of Theorem 2*. Assume first that U is connected, and let w, v be as in Proposition 4. Since $w^*(x) \leq v(x)$ for almost all $x \in U^*$,

$$J(U) = \int_U w(x)\,\mathrm{d}x = \int_{U^*} w^*(x)\,\mathrm{d}x \leq \int_{U^*} v(x)\,\mathrm{d}x = J(U^*)\,,$$

where the second equality is due to Proposition 3(v). By Proposition 4, the inequality is strict unless U is a ball. Since U^* is a ball in any case, Lemma 6 yields

$$J(U) \leq J(U^*) = c_d m(U^*)^{1+2/d} = c_d m(U)^{1+2/d}\,,$$

and again this inequality is strict unless U is a ball. This proves the theorem in the case where U is connected. In general, let (U_n) be the (at most countably many) connected components of U. Then $m(U) = \sum_n m(U_n)$, and by what has already been proved,

$$\begin{aligned} J(U) &= \sum\nolimits_n J(U_n) \leq c_d \sum\nolimits_n m(U_n)^{1+2/d} \leq c_d \left(\sum\nolimits_n m(U_n)\right)^{1+2/d} \\ &= c_d m(U)^{1+2/d}\,; \end{aligned}$$

here the second inequality follows from Lemma 5 and is strict unless $m(U_n) = 0$ for all but one n, that is, unless U is connected. ■

It may be worth noting that even without making use of sophisticated PDE tools such as Proposition 4, it is possible to give an explicit upper bound on the ratio $J(U)/m(U)^{1+2/d}$. The following elegant argument was shown to the author by G. Huisken [**13**]; it only relies on advanced calculus tools (and tacitly assumes U to be sufficiently well-behaved for these tools to all be applicable). Denote by $\lambda_1(U)$ the smallest eigenvalue of the Dirichlet Laplacian on U, that is, the smallest number $\lambda > 0$ such that the boundary value problem

$$\Delta u + \lambda u = 0 \;\text{ in } U\,, \quad u = 0 \;\text{ on } \partial U\,,$$

has a solution other than $u = 0$. With this, deduce from, respectively, the Cauchy–Schwarz inequality, the Poincaré inequality, and the divergence theorem applied to $w\Delta w$ using (1), that

$$\left(\int_U w\,\mathrm{d}x\right)^2 \leq m(U)\int_U w^2\,\mathrm{d}x \leq \frac{m(U)}{\lambda_1(U)}\int_U |\nabla w|^2\,\mathrm{d}x = \frac{m(U)}{\lambda_1(U)}\int_U w\,\mathrm{d}x\,,$$

and hence

$$J(U) = \int_U w\,\mathrm{d}x \leq \frac{m(U)}{\lambda_1(U)}\,.$$

Now, the Faber–Krahn inequality [**8, 16, 22**] says that $\lambda_1(U)$ is minimal precisely when U is a ball of volume $m(U)$. It is well-known, and also easy to check by direct calculation, that for a ball with radius $a > 0$ simply $\lambda_1 = j_{d/2-1}^2/a^2$, where $j_{d/2-1}$ denotes the smallest positive zero of the Bessel function (of the first kind) of order $d/2-1$; see, e.g., [**22**, Note F]. Consequently,

$$\lambda_1(U) \geq \frac{j_{d/2-1}^2}{(m(U)/\omega_d)^{2/d}} = \frac{\omega_d^{2/d} j_{d/2-1}^2}{m(U)^{2/d}}\,,$$

and putting everything together,

$$J(U) \leq \frac{m(U)^{1+2/d}}{\omega_d^{2/d} j_{d/2-1}^2} = c_d \frac{d(d+2)}{j_{d/2-1}^2} m(U)^{1+2/d}\,. \tag{6}$$

This shows that $J(U)/m(U)^{1+2/d}$ indeed is bounded above by a constant independent of U. Note, however, that $j_{d/2-1}^2 < d(d+2)$ for every positive integer d. For instance, $j_0^2 = 5.783 < 2 \cdot 4$ and $j_{1/2}^2 = \pi^2 < 3 \cdot 5$. Thus the bound (6), though arrived at with very little effort and by relying on nothing more than advanced calculus tools, is *not* sharp, quite unlike (3).

3. PHYSICAL INTERPRETATIONS FOR $d = 2$. Its inherent mathematical beauty aside, Theorem 2 is important for its real-world applications as well. In fact, it was precisely these types of applications that suggested the result in the first place. This concluding section briefly discusses two physical interpretations of $J(U)$ as well as (im)possible improvements to Theorem 2. For simplicity and concreteness, assume $d = 2$ throughout. In this case, Theorem 2 says that for every nice open set $U \subset \mathbb{R}^2$,

$$J(U) \leq \frac{m(U)^2}{8\pi}\,; \tag{7}$$

moreover, equality holds in (7) if and only if U is a disc.

For a first physical interpretation, imagine the stationary laminar flow of an ideal viscous incompressible fluid through a pipe with constant cross-section U; see also Figure 3. Assume the pipe to be so long that the flow remains essentially unaffected by whatever disturbances may occur at either end. Under these idealized but not altogether unreasonable assumptions, basic fluid dynamics [**17**] yields the following relation between the volume flow rate Q through the pipe and the pressure difference $p_1 - p_2$ between its ends:

$$Q = (p_1 - p_2)\frac{J(U)}{L\mu}\,; \tag{8}$$

here L and μ are the length of the pipe and the dynamic viscosity of the fluid, respectively. Informally put, (8) identifies $J(U)$ as a kind of *fluid-dynamical conductivity* of a pipe with cross-section U. By (7),

$$Q \leq (p_1 - p_2)\frac{m(U)^2}{8\pi L\mu}\,. \tag{9}$$

Note that (9) can be read in two meaningful ways: Given a specific pipe (length L, cross-section U) and fluid (dynamic viscosity μ), it imposes either an upper bound on the volume flow rate attainable with a given $p_1 - p_2$, or a lower bound on the pressure difference required to attain a given Q. By Theorem 2, the maximal volume flow rate, given $p_1 - p_2$, and the minimal pressure difference, given Q, both occur precisely when the pipe has a circular cross-section. In this case, and with $r = \sqrt{m(U)/\pi}$,

$$Q = (p_1 - p_2)\frac{m(U)^2}{8\pi L\mu} = (p_1 - p_2)\frac{\pi r^4}{8L\mu}\,,$$

which the reader may recognize as the famous *Hagen–Poiseuille law* first formulated in 1838.

For a second, quite different interpretation of $J(U)$ assume that U is simply connected, and consider a rod made from perfectly linear-elastic homogeneous material with cross-section U. One end of the rod is free while the other end is fixed, e.g., welded to a wall; see also Figure 3. Subject the rod to a twist about its longitudinal axis by applying a torque T. Similarly to before, assume the rod to be so long that whatever disturbances may occur at the fixed end do not noticeably affect the resulting elastic deformation. Linear elasticity theory [**10**] then yields the twist angle φ at the free end, caused by the torque T:

$$\varphi = \frac{TL}{4GJ(U)}\,; \tag{10}$$

here L and G are the length of the rod and the shear modulus of its material, respectively. Thus (10) identifies $J(U)$ as the *torsion constant* of a rod with cross-section U. The product $4GJ(U)$ is often referred to as *torsional rigidity*. Again, by (7),

$$\varphi \geq 2\pi \frac{TL}{Gm(U)^2}\,. \tag{11}$$

Given a rod with a specific geometry (length L, cross-section U) and made from a specific linear-elastic homogeneous material (shear modulus G), note that (11) provides a lower bound on the twist angle caused by T. All other parameters being equal, Theorem 2 says that φ is minimal precisely if U is circular. In other words, among all (nice open) cross-sections U with area $m(U)$, the unique cross-section with maximal torsion constant is the disc with radius $\sqrt{m(U)/\pi}$. It was in this form that Theorem 2 for $d = 2$ was first conjectured by Saint-Venant.

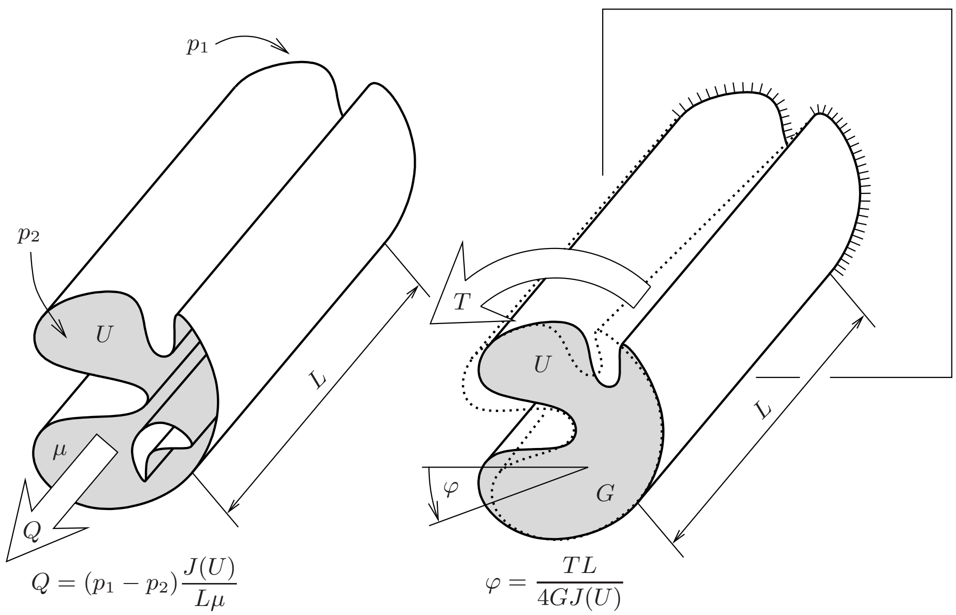

Figure 3. The quantity $J(U)$ can be interpreted as the fluid-dynamical conductivity of a pipe with cross-section U (left) and, if U is simply connected, also as the torsion constant of a homogeneous rod.

The physical interpretations just described raise the practically important question as to whether or not the bounds (9) and (11), and thus ultimately (7) also, can be improved any further. Naturally, this question can be adressed in various ways. For instance, [**18**] shows that if U is *convex*, and $\delta(U)$ denotes the largest radius of any (open) disc contained in U, then

$$J(U) \leq \frac{\delta(U)^2 m(U)}{3}\,. \tag{12}$$

Notice that (12) improves (7) whenever $8\pi\delta(U)^2 < 3m(U)$. For a concrete example, let U be an ellipse with semi-axes $a, b > 0$: Here $m(U) = \pi ab$, $\delta(U) = \min\{a, b\}$, and consequently (12) improves (7) unless $3/8 \leq a/b \leq 8/3$. However, by an explicit calculation in the spirit of Example 1,

$$\frac{J(U)}{\delta(U)^2 m(U)} = \frac{a^2b^2}{4(a^2+b^2)\min\{a,b\}^2} < \frac{1}{4}\,,$$

which shows that, unlike in (7), equality *never* holds in (12) if U is an ellipse.

If U is not convex, perhaps not even simply connected, then (12) may fail, as shown, e.g., by the annulus $\{x \in \mathbb{R}^2 : a < |x| < b\}$ from Example 1, where

$$\frac{J(U)}{\delta(U)^2 m(U)} = \frac{1}{2(a-b)^2}\left(a^2 + b^2 - \frac{a^2-b^2}{\log(a/b)}\right) > \frac{1}{3}.$$

To get an idea what an improvement of (7) might look like in this case, it is worth recalling for a moment the most classical of all isoperimetric inequalities [**7**, Ch. 3]. Known already in antiquity (though not proved rigorously until the late nineteenth century), it asserts that for every nice open set $U \subset \mathbb{R}^2$ with piecewise smooth boundary of total length $\ell(\partial U)$,

$$m(U) \leq \frac{\ell(\partial U)^2}{4\pi}\,; \tag{13}$$

moreover, equality holds in (13) if and only if U is a disc. (Surely, the strong similarity between (7) and (13) is not lost on the reader.) In light of the fundamental importance of (13), improvements and variations thereof have long been of interest, not least to authors and readers of the MONTHLY [**1, 5, 19, 25**]. For a very simple improvement, assume that U has a finite number of disjoint holes. More specifically, denote by H_∞ the unbounded component of $\mathbb{R}^2 \setminus U$, and assume that $\mathbb{R}^2 \setminus (U \cup H_\infty)$ is the disjoint union of $\overline{H_1}, \dots, \overline{H_n}$ where each H_i itself is a nice open set, and $\overline{H_i}$ is the closure of H_i; see also Figure 4. With this, replacing U by $U \cup \overline{H_i}$ (or by $\mathbb{R}^2 \setminus H_\infty$) means, informally put, that the hole H_i (or all holes) are being "removed" or "filled up." Since removing a hole increases the left side in (13) but reduces the right side, the bound $1/(4\pi)$ will typically be very inaccurate, that is, far larger than the actual ratio $m(U)/\ell(\partial U)^2$. By means of the **porosity** of U, defined as

$$p := \sqrt{\frac{\sum_{i=1}^n m(H_i)}{m(\mathbb{R}^2 \setminus H_\infty)}} = \sqrt{\frac{\sum_{i=1}^n m(H_i)}{m(U) + \sum_{i=1}^n m(H_i)}}\,,$$

it is easy to establish a more accurate bound. To this end, notice that

$$\ell(\partial U)^2 = \left(\ell(\partial H_\infty) + \sum\nolimits_{i=1}^n \ell(\partial H_i)\right)^2$$

$$\geq 4\pi \left(\sqrt{m(U) + \sum_{i=1}^{n} m(H_i)} + \sum_{i=1}^{n} \sqrt{m(H_i)} \right)^2$$

$$\geq 4\pi \left(\sqrt{m(U) + \sum_{i=1}^{n} m(H_i)} + \sqrt{\sum_{i=1}^{n} m(H_i)} \right)^2$$

$$= 4\pi \left(\sqrt{\frac{m(U)}{1-p^2}} + p\sqrt{\frac{m(U)}{1-p^2}} \right)^2 = 4\pi m(U) \frac{1+p}{1-p} .$$

Here, the first inequality is due to (13) applied to $\mathbb{R}^2 \setminus H_\infty$ and $H_1, \ldots, H_n$ individually, whereas the second inequality is due to Lemma 5. Thus an improved form of (13) in the presence of holes, with the total size of the latter being measured by the porosity $0 < p < 1$, is

$$m(U) \leq \frac{\ell(\partial U)^2}{4\pi} \cdot \frac{1-p}{1+p} . \tag{14}$$

Moreover, it is clear that equality holds in (14) if and only if U is a disc containing a single circular hole, the radius of the hole being p times the radius of U with the hole removed. To appreciate the improvement over (13) that (14) represents, note for instance that for $p = 0.9$ the bound provided by the former is 19 times the bound provided by the latter!

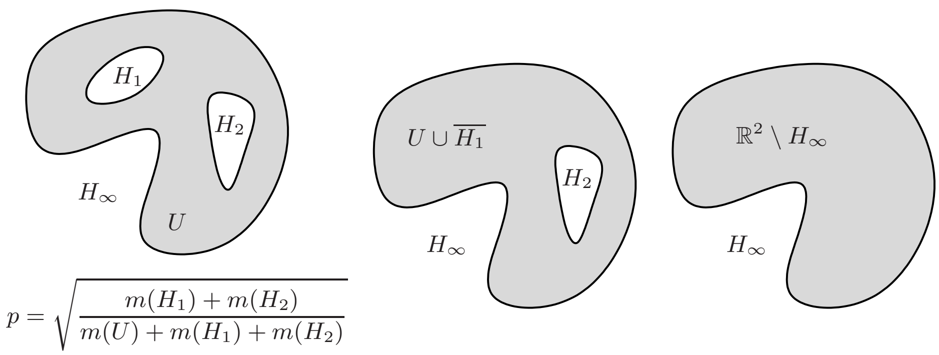

Figure 4. In the presence of holes, the isoperimetric inequality (13) takes the improved form (14).

Returning now to $J(U)$, the reader may wonder whether or not it is possible to achieve a similar improvement for (7). As it turns out, this is *not* possible. Given any $0 < p < 1$, one may utilize Example 1 to design a nice open set U with a single hole and with porosity equal to p such that the ratio $J(U)/m(U)^2$ is as close to $1/(8\pi)$ as one wishes; see also Figure 5. This shows that, in a way, the inequality (7), and thus Theorem 2 as well, are the best possible even in the presence of holes.

Finally, the reader may be curious to know how (10) changes if U is *not* simply connected. Think for instance of a hollow shaft. In this case, $J(U)$ has to be replaced by a quantity $\widetilde{J}(U)$, the precise definition of which is slightly more complicated; see, e.g., [**10, 16, 23**] for details. To mention but one example, if U again is the annulus $\{x \in \mathbb{R}^2 : a < |x| < b\}$ with $0 < a < b$ then

$$\widetilde{J}(U) = \frac{\pi}{8}(b^4 - a^4) = \frac{m(U)^2}{8\pi} \cdot \frac{1+p^2}{1-p^2} ,$$

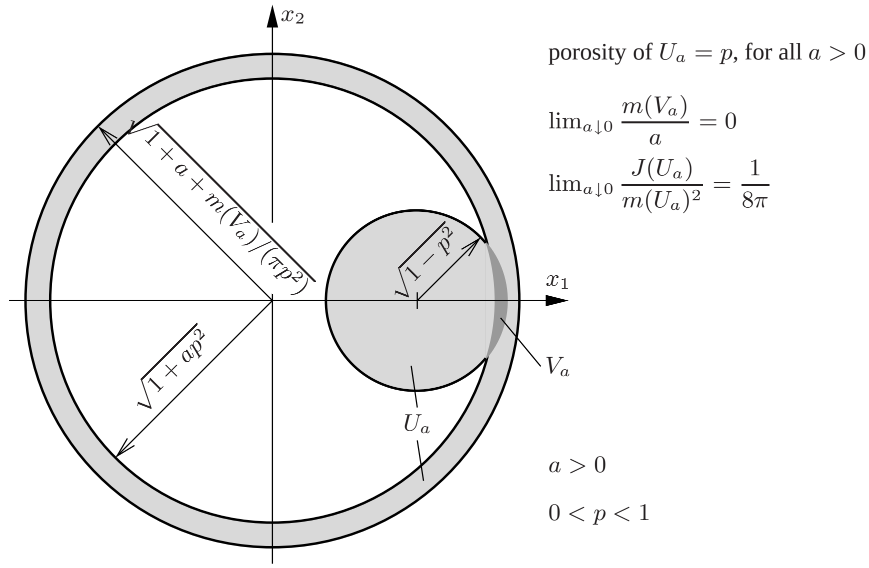

Figure 5. Why (7) cannot be improved in the presence of holes, quite unlike (13).

where the second equality is due to U having porosity $p = a/b < 1$. For comparison, recall from (2) that

$$J(U) = \frac{m(U)^2}{8\pi}\left(\frac{1+p^2}{1-p^2} + \frac{1}{\log p}\right) .$$

A classical theorem by Pólya and Weinstein [**23**] says that the ratio $\widetilde{J}(U)/m(U)^2$ cannot ever be larger than in this example. More precisely,

$$\widetilde{J}(U) \leq \frac{m(U)^2}{8\pi} \cdot \frac{1+p^2}{1-p^2} \tag{15}$$

for every nice open set $U \subset \mathbb{R}^2$ with porosity $0 \leq p < 1$. Although the factor $\dfrac{1+p^2}{1-p^2}$ in (15) may appear reminiscent of the factor $\dfrac{1-p}{1+p}$ in (14), do notice that, in stark contrast to the latter, the former is unbounded as $p \uparrow 1$. Thus, given $m(U)$, it is possible to design cross-sections with arbitrarily large torsion constant—simply take U to be an annulus with area $m(U)$ and porosity p close to 1. Though interesting theoretically, this observation has but little practical value: If p is close to 1, that is, if the set U is very porous or "thin" then it becomes useless as a cross-section of a rod. On the one hand, the resulting diameter of U may be too large for the purpose at hand. On the other hand, and perhaps more stringently, the rod may undergo a loss of stability (buckling), the mere possibility of which has not been taken into account at all in the derivation of (10).

Having seen, in this article, the beauty and importance of the inequality by Saint-Venant and Pólya, the reader may want to further explore the vast and wondrous subject of isoperimetric inequalities through, e.g., [**3, 4, 19, 20, 22, 25**].

ACKNOWLEDGMENTS. The author was supported by an NSERC Discovery Grant. He is grateful to T. Hillen and G. Huisken for insightful communications. He also wishes to thank two anonymous referees who brought the recent preprint [**9**] to his attention and made suggestions that improved the exposition.

REFERENCES

1. Adler, C. F. (1945). An isoperimetric problem with an inequality. *Amer. Math. Monthly*. 52(2): 59–69.
2. Alvino, A., Lions, P. L., Trombetti, G. (1986). A remark on comparison results via symmetrization. *Proc. Roy. Soc. Edinburgh Sect. A*. 102(1-2): 37–48.
3. Baernstein, A., Drasin, D., Laugesen, R. S. (2019). *Symmetrization in Analysis*. New Mathematical Monographs 36. Cambridge: Cambridge University Press.
4. Bandle, C. (1980). *Isoperimetric inequalities and applications*. Monographs and Studies in Mathematics 7. Boston, Mass.-London: Pitman (Advanced Publishing Program).
5. Benson, D. C. (1970). Sharpened forms of the plane isoperimetric inequality. *Amer. Math. Monthly*. 77(1): 29–34.
6. Brasco, L. (2014). On torsional rigidity and principal frequencies: an invitation to the Kohler–Jobin rearrangement technique. *ESAIM Control Optim. Calc. Var.* 20(2): 315–338.
7. Burago, Y. D., Zalgaller, V. A. (1988). *Geometric Inequalities*. Grundlehren der mathematischen Wissenschaften 285. Berlin: Springer-Verlag.
8. Burchard, A. (2009). A Short Course on Rearrangement Inequalities. Pdf lecture notes available at `www.math.toronto.edu/almut/rearrange.pdf`, last accessed: 15 November 2021.
9. Chen, D., Li, H. (2021). Talenti's comparison theorem for Poisson equation and applications on Riemannian manifold with nonnegative Ricci curvature. `arXiv:2104.05568`.
10. Flügge, S. (1958). *Encyclopedia of Physics: Elasticity and Plasticity*, Vol. 6. Berlin-Heidelberg-New York: Springer-Verlag.
11. Gamara, N., A. Hasnaoui, A., Makni, A. (2015). Torsional rigidity on compact Riemannian manifolds with lower Ricci curvature bounds. *Open Math.* 13(1): 557–570.
12. Gilbarg, D., Trudinger, N. S. (1983). *Elliptic Partial Differential Equations of Second Order*, 2nd ed. Grundlehren der mathematischen Wissenschaften 224. Berlin: Springer-Verlag.
13. Huisken, G. (2021). Private communication.
14. Kawohl, B. (1985). *Rearrangements and Convexity of Level Sets in PDE*. Lecture Notes in Mathematics 1150. Berlin: Springer-Verlag.
15. Kesavan, S. (1991). On a comparison theorem via symmetrisation. *Proc. Roy. Soc. Edinburgh Sect. A*. 119(1-2): 159–167.
16. Kesavan, S. (2006). *Symmetrization and applications*. Series in Analysis 3. Hackensack, NJ: World Scientific Publishing.
17. Landau, L. D., Lifshitz, E. M. (1987). *Course of Theoretical Physics: Fluid Mechanics*, Vol. 6. Oxford: Pergamon Press.
18. Makai, E. (1966). A proof of Saint-Venant's theorem on torsional rigidity. *Acta Math. Hungar.* 17(3-4): 419–422.
19. Osserman, R. (1979). Bonnesen-style isoperimetric inequalities. *Amer. Math. Monthly*. 86(1): 1–29.
20. Payne, L. E. (1967). Isoperimetric inequalities and their applications. *SIAM Rev.* 9(3): 453–488.
21. Pólya, G. (1948). Torsional rigidity, principal frequency, electrostatic capacity and symmetrization. *Quart. Appl. Math.* 6(3): 267–277.
22. Pólya, G., Szegö, G. (1951). *Isoperimetric Inequalities in Mathematical Physics*. Annals of Mathematical Studies 27. Princeton, NJ: Princeton University Press.
23. Pólya, G., Weinstein. A. (1950). On the torsional rigidity of multiply connected cross-sections. *Ann. Math.* 52(1): 154–163.
24. Saint-Venant, J. B. de (1856). *Mémoire sur la torsion des prismes*. Mémoires présentés par divers savants à l'Académie des Sciences. 14: 233–560.
25. D. Vukotić, D. (2003). The isoperimetric inequality and a theorem of Hardy and Littlewood. *Amer. Math. Monthly*. 110(6): 532–536.

ARNO BERGER is a professor of mathematics at the University of Alberta. He is interested primarily in dynamical systems, analysis, and probability theory. Before settling in Canada, he held academic positions in Austria and New Zealand, and spent extended research visits to the U.S., U.K., and Germany. He received his PhD from the Technische Universität Wien in 1997.
Mathematical and Statistical Sciences, University of Alberta, Edmonton AB T6G 2G1, Canada
berger@ualberta.ca