THE PROBLEM OF QUEEN DIDO

Overview of the Subject of Isoperimetry

by

Mark Ashbaugh & Rafael Benguria [1]

The subject of isoperimetry has a long and eventful history, both for its impact on people's imaginations and society in general and for the impetus it has given to the study of various mathematical subjects.

Isoperimetry began with the problem confronted by Queen Dido, which was to find the shape of the boundary that should be laid down (using strips of oxhide) to enclose maximum area. If one assumes a straight coastline, then the answer, which was by all appearances discovered by Queen Dido, is to lay down the hide in the shape of a semi-circle.

One finds the problem of Queen Dido colorfully described, including various embellishments of the basic problem, in the expository account that Lord Kelvin gave in 1893 (see `http://math.arizona.edu/~dido/lord-kelvin1894.html`). If one takes account that land may vary in value, or that the coastline may be irregular, one can arrive at various more complicated problems. In a much more recent exposition, Hildebrandt and Tromba, in their book *The Parsimonious Universe: Shape and Form in the Natural World* (originally published as *Mathematics and Optimal Form*), give a much more detailed account of isoperimetric problems and their recurrence throughout history. In particular, it is interesting to see how many walled cities in the Middle Ages were constructed to have a nearly circular perimeter, or to see in general that the growth of many cities gave them a nearly circular form.

On the mathematical side, we find already in Euclid (around 300 BC) the proof that among rectangles of a given perimeter the one having the greatest area is the square. Also, various writers from antiquity speculated on optimal properties of the honeycombs of bees. When Thomas Hales proved in 2001 that regular hexagons provide a least-perimeter way to partition the plane into unit areas, it was the longest standing open problem in mathematics. As for 3D, Lord Kelvin proposed a solution consisting of relaxed, 14-sided, truncated octahedra. In 1994 D. Weaire and R. Phelan disproved Kelvin's conjecture by providing a new candidate using both 12- and 14-sided shapes.

[1] with contributions from Lennie Friedlander, Evans Harrell, Lotfi Hermi, and Frank Morgan.

Among the ancient Greeks who worked on the isoperimetric problem we mention Zenodorus (c. 200 - c. 140 BC) who wrote a now-lost treatise *On Isoperimetric Figures* and Ptolemy (c. 90 - c. 168 AD). It is thanks to Theon of Alexandria (c. 335 - c. 405 AD) who wrote a commentary on the work of Ptolemy that we know the results of Zenodorus. Al-Kindi, an Arab mathematician and the son of one of their kings, wrote in the 9th century *A Treatise on Isoperimetric Figures and Isepiphanies*, that is solids of given surface. There is also a lost treatise by al-Hasan ibn al-Haytham (965 - c. 1039). Abu Ja'far al-Khazin, commenting on Ptolemy's Almagest in the 10th century, generalized earlier works. Johannes de Sacrobosco (John of Holywood, c. 1195 - c. 1256 AD), an English scholar and astronomer, wrote *Tractatus de Sphaera*. A commentary on this treatise, dealing specifically with isoperimetry, can be found in *Two New Sciences* of Galileo Galilei published in 1638.

The mathematical study of the isoperimetric problem and related problems really began to take off with the advent of calculus, when people like Newton, Leibniz, the Bernoullis, and others developed systematic ways of attacking optimization problems based on the calculus, and within a few short years were attacking problems in the calculus of variations (that is, the problem of finding an optimizing path or shape of curve from among some class of curves). For example, the brachistochrone problem was formulated by Johann Bernoulli and solved by Newton and both Bernoulli brothers, Jakob (James) and Johann (John). In the same period, the problem of the shape of a hanging chain (the catenary) was posed and solved, and Newton considered the shape of projectile which would give the least air resistance (the question of designing the optimal shape for the nose-cone of a rocket or missile), but without reaching definitive conclusions. Others, including US President Thomas Jefferson, considered questions such as the optimal shape for ploughshares.

In the century following the early development of calculus by Newton, Leibniz, the Bernoulli brothers, and others, the calculus of variations was brought to a relatively advanced state, especially from the point of view of direct solutions of problems, by Euler and Lagrange. The explicit solution of the classical isoperimetric problem could be derived in those terms (using variational theory with a constraint), and many other problems could be formulated and solved. Euler and Lagrange had shown that all of mechanics could be put into this framework, and that various physical and mathematical problems could be understood from the point of view of various optimization or variational principles (recall Fermat's principle of least time, or, more generally, the d'Alembert/Maupertuis principle of least action, for which Euler gave the definitive formulation). Almost a century later, Jacobi and Hamilton also made important contributions to this area, especially as regards mechanics.

In the nineteenth century Jakob Steiner attacked the classical isoperimetric problem using direct geometrical tools, which were very suggestive and instructive and led to many further developments. Around this time, however, Weierstrass realized that there could be subtle problems involved with attacking certain extremization problems, since it might be that no extremizer exists. Since that time it has been recognized that the existence question is where one must begin in attacking many problems from geometry and the calculus of variations. This led to various existence and uniqueness results, and to the so-called direct methods of the calculus of variations, wherein one tries to prove existence directly using extremizing sequences and various mathematical tools (developed by Weierstrass, Schwarz, Poincaré, Hilbert, and their contemporaries, and also more modern contributors, up to the present time).

A very useful development that came around the turn of the 20th century was Hurwitz's realization that the classical isoperimetric problem could be solved relatively simply in terms of Fourier series and some of their basic properties (e.g., Wirtinger's inequality). The Fourier analysis approach to the isoperimetric inequality gave rise to further studies in higher dimensions where spherical harmonics take the place of Fourier series. This field is nicely summarized from a modern perspective in Groemer's book, *Geometric Applications of Fourier Series and Spherical Harmonics*.

Also in the nineteenth century the Belgian physicist J. Plateau experimented with soap films and conjectured that any wire loop (nice closed curve) bounds a soap film or minimal surface (of mean curvature 0). In 1936 J. Douglas won an inaugural Fields Medal for proving that every such loop bounds an immersed minimal disc, though his solution admitted self-intersections of a type which never occur in real soap films. Only with the advent of geometric measure theory with work of L. C. Young, E. De Giorgi, E. R. Reifenberg, H. Federer, W. Fleming, F. Almgren, J. Taylor, R. Hardt, L. Simon and others was the general existence of certain soap films established. It remains an open question today whether a smooth Jordan curve bounds a least-area soap film ("$(M, 0, \delta)$-minimal set").

With a round soap bubble proved by Schwarz in 1884 to be the least-perimeter way to enclose a given volume of air, the next question was whether the double bubble that forms when two soap bubbles come together is the least-perimeter way to enclose and separate two given volumes of air. Years of progress by many mathematicians and undergraduates culminated in the 2002 proof by M. Hutchings, F. Morgan, M. Ritoré, and A. Ros.

From the point of view of engineering and design, perhaps the first truly interesting isoperimetric problem was to consider "the shape of the strongest column," a problem formulated by Lagrange in 1773 (but not fully solved until much later). In the mid 1800's T. Clausen was able to make his way around some of the points that Lagrange had stumbled over, though some questions have remained about the problem and its resolution up to recent times. See Steve Cox's Mathematical Intelligencer article, "The shape of the ideal column" to get a sense of where things stand currently. Several of the most pertinent recent contributors include J. Keller, I. Tadjbakhsh, M. Overton, and S. Cox. This problem has to do with the buckling of columns, and similar problems can be considered for horizontal beams under a variety of loads, and for plates and other structural members having greater geometrical complexity.

Also in the mid 1800's, J. C. B. St. Venant put forward the question of finding the cross-section of a uniform beam or column that would be most resistant to twisting (the so-called "problem of torsional rigidity"). He conjectured that for a given cross-sectional area, assumed to be a simply-connected region (and with all other physical parameters held fixed), the shape giving the greatest torsional rigidity was the circular one. This problem was finally resolved by George Pólya in 1948 (in the sense that St. Venant had conjectured). Much work has been done on torsion problems since that time, since it is also of interest to consider non-simply connected regions and other variations of the basic problem.

A few years after St. Venant considered the torsion problem, Lord Rayleigh set forth (and formulated conjectures for) (1) the shape of drum that would minimize its fundamental (or "base") tone for fixed area (with other physical parameters held fixed), (2) in static electricity, the shape of capacitor among simply-connected bodies of finite extent that would minimize capacity for given volume, and (3) the shape of clamped plate that would minimize its fundamental frequency for given area. In each case Rayleigh conjectured that the minimizing shape was circular (or spherical, in the case of the 3-dimensional capacitor problem).

Other related problems include the question of what shape minimizes heat loss (described colorfully by Pólya as the explanation for why a cat curls itself into a ball on a cold winter's night) and the shape of a body that minimizes its (gravitational) potential energy.

All of the aforementioned physical problems can be formulated as variational problems, with many leading directly to eigenvalue problems. In the early part of the 20th century there was interesting progress on several of these problems, the most spectacular being the solution of the problem of minimizing the fundamental tone of a drum by Faber and Krahn in independent papers in the early to mid 1920's (the answer is that one should take a circular drum of the given area). Somewhat before Faber and Krahn, Courant had obtained a weaker version of the result, that for fixed perimeter the way to minimize the fundamental tone was to take a circular drum. Earlier Poincaré had made progress on the capacity problem, with the full solution due to Gabor Szegő coming in 1930.

Around 1950, Pólya and Szegő took on the job of studying and systematizing prior works on physical isoperimetric problems, and of advancing the field on a wide front. Their book *Isoperimetric Inequalities in Mathematical Physics*, published at that time, is a classic of the field. The techniques that they put at the forefront included Steiner symmetrization, and, generally, rearrangement inequalities. It could quite justifiably be said that all modern work on isoperimetric inequalities for physical quantities builds on the work of Pólya and Szegő and their collaborators. Pólya and Szegő's book contains, for example, the solutions to the St. Venant and capacity problems mentioned above.

Pólya and Szegő's interest in the subject stimulated interest by others and led to many important and interesting develoments in the field. Perhaps foremost among the early contributors to these developments are Payne, Hersch, and Weinberger, who participated in many of the advances and inspired their students and others to enter the field. Thus we find Payne, Pólya, and Weinberger obtaining very simple and nice *universal inequalities* for combinations of eigenvalues in the mid-50's, and conjecturing what the sharp forms of certain of these inequalities might be. This leads one into the subject of isoperimetric inequalities for eigenvalue ratios, which attracted considerable interest (particularly the ratio λ_2/λ_1) and was finally solved by Ashbaugh and Benguria in 1990. Following a significant advance in work of H. C. Yang in the early 90's, the subject of universal eigenvalue inequalities has taken off, with many papers contributing to and advancing the subject, and with much work continuing to the present day. The work of Yang has allowed researchers to make fundamental connections between the field of universal eigenvalue inequalities and the subject of eigenvalue asymptotics, as begun by Hermann Weyl around 1910. This, too, is a burgeoning field, with key recent contributors including Q. M. Cheng and H. C. Yang, E. M. Harrell and L. Hermi, E. M. Harrell and J. Stubbe, and several others.

Conformal methods play an important role in the study of two-dimensional problems. Szegő used them to prove that the disk minimizes $\mu_1(\Omega)^{-1} + \mu_2(\Omega)^{-1}$ in the class of simply connected planar domains of given area (here the μ_j's are the positive eigenvalues of the Neumann Laplacian). Hersch proved that the smallest positive eigenvalue of the Laplace–Beltrami operator on a two-sphere cannot exceed the one of the operator for the round metric of the same area. The crucial observation is that the numerator in the Rayleigh quotient, $\int |\nabla u|^2 dx$, is conformally invariant when the dimension equals 2. P. C. Yang and S.-T. Yau proved that the first positive eigenvalue on a surface of genus g of given area has an upper bound; moreover, they gave a precise bound. In the case $g = 2$, Jacobson, Levitin, Nadirashvili, Nigam, and Polterovich proved that Yang and Yau's bound is sharp, and it is saturated on a singular metric on a surface of conformal type of $y^2 = x^5 - x$. Their proof relied on some numerics. It would therefore be interesting to have a numerics-free proof. In the case when the dimension of a manifold is higher than 2, Urakawa proved that, in the class of metrics of fixed volume, the first positive eigenvalue of the Laplacian can be arbitrarily large. However, within a given conformal class, it is bounded, and these upper bounds are bounded from below when one varies conformal classes (Friedlander, Nadirashvili). Recently, Colbois, Dryden, and El Soufi studied bounds for eigenvalues of the Laplacian for G-invariant metrics in a certain conformal class. Here G is a Lie group acting on a manifold.

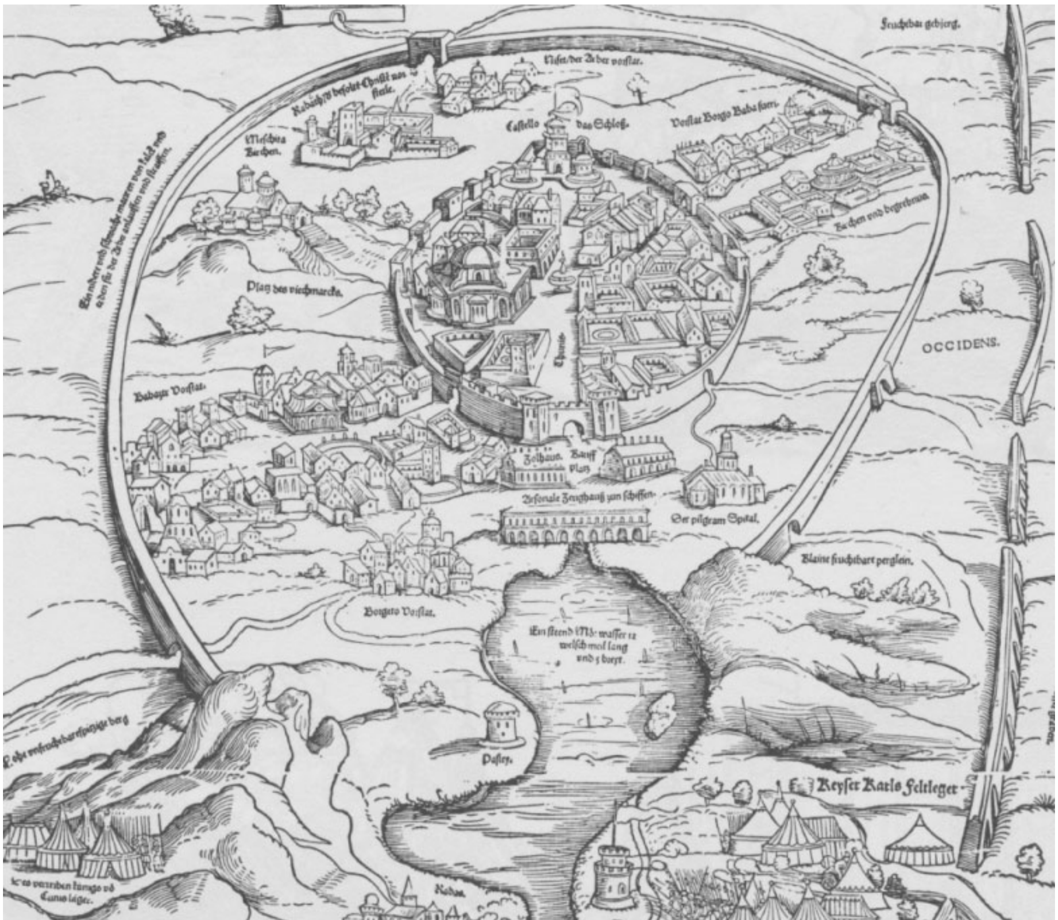
Obviously there are many other topics that figure in the history of isoperimetric problems and related areas and the most we could do here was point out some of the highlights. To help make up for the deficiencies of such coverage, we conclude with a brief summary of some of the relevant literature, which it is hoped can be used to widen the coverage and give hints of other worthy topics in the general area. For historical orientation, we recommend the article by Lord Kelvin and the expository book by Hildebrandt and Tromba (both mentioned earlier).

For further background on the classical isoperimetric problem one cannot do better than to consult the book of Burago and Zalgaller, *Geometric Inequalities*, and the 1978 review article in the Bulletin of the American Mathematical Society by Robert Osserman, "The isoperimetric inequality." Other books and articles of interest include Tikhomirov's *Stories of Maxima and Minima*, Pólya's *Mathematics and Plausible Reasoning* (in 2 vols.; the most relevant sections of this can be found at the conference website, `http://math.arizona.edu/~dido/polya1954.html`), and Pólya's article "Circle, sphere, symmetrization, and some classical physical problems", or D. Pedoe's *Circles: A Mathematical View* and N. Kazarinoff's *Geometric Inequalities.* For aspects of the isoperimetric problem occurring in the setting of Riemannian geometry one can consult the books of Chavel (*Eigenvalues in Riemannian Geometry*, *Riemannian Geometry: A Modern Introduction*, and *Isoperimetric Inequalities: Differential Geometric and Analytic Perspectives*) and of Marcel Berger (*A Panoramic View of Riemannian Geometry*; or see his books *Geometry I and II* for much useful related information, mostly in the classical setting). Chavel's book *Eigenvalues in Riemannian Geometry* includes topics that extend well into the domain of isoperimetric inequalities for physical quantities.

For modern developments in minimal surface theory and much more, one cannot do better than to consult Almgren's *Plateau's Problem: An Invitation to Varifold Geometry,* and Morgan's *Geometric Measure Theory: A Beginner's Guide* (fourth edition, 2009). Beyond that one has Federer's classic *Geometric Measure Theory.* For the more classical background in minimal surface theory, there are a number of books and articles, among which we mention Osserman's *Survey of Minimal Surfaces* (updated edition, Dover, 1986).

On the side of isoperimetric inequalities for physical quantities one can find much of interest in the works of Pólya and Chavel already mentioned. In the 1960's and beyond, a key role was filled by Payne's SIAM Review paper, "Isoperimetric inequalities and their applications." This paper provides the background and setting for many physical isoperimetric problems (including their mathematical formulation), and also states a variety of open problems and conjectures. In 1991, Payne updated his discussion of many of these problems in his contribution “Some comments on the past fifty years of isoperimetric inequalities” to the book *Inequalities: Fifty Years on from Hardy, Littlewood and Pólya*, edited by W. N. Everitt. Beyond that, one has the books of C. Bandle (*Isoperimetric Inequalities and Applications*), R. Sperb (*Maximum Principles and Their Applications*), and B. Kawohl (*Rearrangements and Convexity of Level Sets in PDE*), dating to the early to mid 80’s, and the more recent books of D. Bucur and G. Buttazzo (*Variational Methods in Shape Optimization Problems*), A. Henrot (*Extremum Problems for Eigenvalues of Elliptic Operators*), and S. Kesavan (*Symmetrization and Applications*).

Finally, we mention the excellent book by Lieb and Loss, *Analysis*, second edition, which covers much of interest in the field of symmetrization and rearrangements in the context of the classical inequalities of analysis and mathematical physics, as well as much else besides. In particular, the book covers the problems of minimizing capacity and gravitational potential energy, and has a full discussion of Lieb–Thirring inequalities and their relation to the question of the stability of matter.



Detail from a sketch made in commemoration of Carlos Quintos’ campaign on the doubled-walled city of Tunis, clearly satisfying the isoperimetric property of the circle. (31 August 1535).