

# Mathematics of the Gateway Arch

Robert Osserman

One of the oldest problems to have been solved using the calculus of variations was to find the equation for the shape formed by a hanging chain or flexible cord. Most often it is said that Galileo was the first to pose the problem and that he claimed that the resulting curve was a parabola. (See, for example, Goldstine [5], p. 32, footnote 42, and Truesdell [18], pp. 43–44.) Both of these statements are at best half true. Galileo’s exact words (in a classic English translation of his book *Two New Sciences* [4], p. 290) are

...a cord stretched more or less tightly assumes a curve which closely approximates the parabola...the coincidence is more exact in proportion as the parabola is drawn with less curvature or, so to speak, more stretched; so that using parabolas described with elevations less than  $45^\circ$  the chain fits its parabola almost perfectly....

Nowhere, to my knowledge, does Galileo address the question of finding the exact shape of the curve formed by the hanging chain. And he is exactly right that for a parabola with elevation less than  $45^\circ$  the curve formed by the chain is an extremely close approximation to the parabola. (See Figure 1.)

This passage follows immediately after an extended discussion during which Galileo presents one of his most famous discoveries: that the path of a projectile in a vacuum under the influence of gravity is a parabola ([4], pp. 244–290). In the course of the discussion, he points out that for a fixed initial velocity, the range will be maximized by an initial elevation of  $45^\circ$ . Finally he addresses the question of how to quickly draw a number of parabolas, leading to the passage cited above.

The reason that one might mistake Galileo’s intentions and conclusions is that elsewhere in *Two New Sciences* ([4], p. 149) Galileo addresses the same problem and says simply that the “chain will

assume the form of a parabola.” The reason that virtually everyone refers to this passage rather than the one in which Galileo spells out his conclusions in far greater detail and complete accuracy is not as clear.

Some of Galileo’s earliest experiments with hanging chains apparently date from the summer of 1602, when he was working with Guidobaldo del Monte. Had he thought to pose and try to answer the question of the exact equation for the shape of a hanging chain, he would have encountered two serious obstacles. First, the essential tool needed—the calculus—had not yet been invented. Second, the equation for the curve involves logarithms, which had also not yet been invented.

After the problem was finally explicitly stated, it was treated frequently by a number of authors between 1690 and 1720 and the correct answer derived, although some of the initial reasoning used was decidedly suspect. The curve became known as a *catenary*. Its equation is  $y = \cosh x$ , the hyperbolic cosine, up to change of coordinates.

One method used to arrive at this result was the calculus of variations, finding the shape of a curve whose center of gravity is the lowest among all curves having a prescribed length and prescribed endpoints.

## The Gateway Arch

It was Robert Hooke who in 1675 made the connection between the ideal shape of an arch and that of a hanging chain in an aphorism that says, in abbreviated form, “As hangs the chain, so stands the arch.” In other words, the geometry of a standing arch should mirror that of a hanging chain. The horizontal and vertical forces in a hanging chain must add to a force directed along the chain, since any component perpendicular to the chain would cause it to move in that direction to gain equilibrium. Similarly, one wants the combined forces at each point of an arch to add up to a vector tangent to the arch. In both cases, the horizontal component of the force is constant and simply transmitted along the arch or chain, while the vertical forces are mirror images.

---

Robert Osserman is professor emeritus of mathematics at Stanford University and special projects director at MSRI. His email address is [osserman@math.stanford.edu](mailto:osserman@math.stanford.edu).



Figure 1. Left: 45° parabola and catenary; Right: 30° parabola and catenary.

At least in part for these reasons, the shape of the Gateway Arch is often described mistakenly as a catenary (when not even more mistakenly as a parabola). In fact, the equation on which the arch is based is

$$(1) \quad y = A \cosh Bx + C,$$

which is a catenary only if  $A = 1/B$ . For the Gateway Arch,  $A, B$ , and  $C$  are numerical constants, with  $A$  approximately equal to  $.69(1/B)$ . As a result, those who wish to be more accurate describe the shape of the Arch as a “modified catenary” or, more often, a “weighted catenary”. The idea of the weighted catenary is to use either a chain with literal weights attached at various points or else a continuous version, such as a cord of varying density. (Both of those methods, incidentally, were used by Saarinen to find a shape for the Gateway Arch that appealed to him esthetically.)

Some natural questions to ask are:

1. Is there a density function that yields the curve defined in equation (1) above, and if so, what is that function?
2. How descriptive is the term “weighted catenary”? In other words, what are the curves that fall in this category?

In order to answer these questions, we start with some precise definitions.

**Definitions.** (1) • A *weighted chain*  $C$  is a pair  $(f(x), \rho(s))$  where  $f(x)$  is a suitably smooth function on an interval  $x_1 \leq x \leq x_2$ ,  $s$  is arclength along the curve  $y = f(x)$ , and  $\rho(s)$  is a positive continuous function of  $s$ .

• The function  $\rho(s)$  is called the *density function* of  $C$ .

• A *weighted catenary* is a weighted chain  $C$  in which

- (i)  $-\infty < f'(x_1) < 0 < f'(x_2) < \infty$ ,
- (ii)  $f''(x)$  is continuous and positive,
- (iii) for  $x_1 \leq X \leq x_2$ ,

$$(2) \quad f'(X) = V(X)/H,$$

with

$$(3) \quad V(X) = \int_{x_0}^X \rho(s(x)) \frac{ds}{dx} dx$$

where  $f'(x_0) = 0$ , and  $H$  is a constant that depends on both  $f$  and  $\rho$ :

$$(4) \quad H = V(x_2)/f'(x_2).$$

There is a certain amount of arbitrariness in these definitions. The smoothness of the function  $f(x)$  is deliberately left vague, since one might well want to allow different classes of functions in different situations. Similarly, one could allow the density function to have point masses or to be even more general.

Nonetheless, conditions (i)–(iii) are natural ones based on the physics underlying the problem. As noted earlier, for equilibrium we want the sum of the horizontal and vertical forces acting on the chain to be a vector tangent to the curve. The horizontal force is constant and simply transmitted along the curve, while the vertical force is due to gravity acting on the part of the chain from the lowest point to the point that we are considering on the curve and is equal to the total weight of that part of the chain.

Note that we do not require that the endpoints be at the same height—in other words, that  $f(x_1) = f(x_2)$ —although that will be true in all the examples of interest to us. In that case, condition (i) would follow from the physics of the situation. However, condition (i) does guarantee the existence of an interior point  $x_0$  of the interval where  $f' = 0$ , while condition (ii) implies that there is a unique such point.

Note also that the fact that there is a constant nonzero horizontal component to the force acting on the curve implies that the slope cannot be infinite at the endpoints. For example, a weighted catenary cannot be in the form of a semicircle.

Note finally that if  $\rho(s)$  is multiplied by a positive constant, then both  $H$  and  $V(x)$  are multiplied by the same constant. It follows that if the pair  $(f(x), \rho(s))$  defines a weighted catenary, then so does the pair  $(f(x), c\rho(s))$  for any positive constant  $c$ .

We are now able to state our first result.

**Theorem 1.** *If  $f(x)$  satisfies conditions (i) and (ii) above, then there is a function  $\rho(s)$ , unique up to a multiplicative constant, such that the pair  $(f(x), \rho(s))$  is a weighted catenary. Namely*

$$(5) \quad \rho(s(x)) = \frac{Hf''(x)}{\sqrt{1 + f'(x)^2}},$$

where  $H$  is an arbitrary positive constant. The value of  $H$  is then given by

$$(6) \quad H = \frac{W}{|f'(x_1)| + |f'(x_2)|},$$

with

$$W = \int_{x_1}^{x_2} \rho(s(x)) \frac{ds}{dx} dx$$

equal to the total weight of the chain.

*Proof.* Suppose first that there exists a function  $\rho(s)$  such that the pair  $(f(x), \rho(s))$  is a weighted catenary. Then by equations (2) and (3), for  $X \geq x_0$ ,

$$Hf''(X) = V'(X) = \rho(s(X))(ds/dx)|_{x=X}.$$

Since  $ds/dx = \sqrt{1 + f'(x)^2}$ , it follows that  $\rho(s)$  must have the form given in equation (5). For  $X < x_0$ ,

$$V(X) = - \int_X^{x_0} \rho(s(x))(ds/dx) dx$$

and

$$Hf''(X) = V'(X) = \rho(s(X))(ds/dx)|_{x=X}$$

as before. Finally, the total weight  $W$  is

$$W = \int_{x_1}^{x_2} \rho(s) ds = \int_{x_1}^{x_0} + \int_{x_0}^{x_2} = V(x_2) - V(x_1).$$

But  $V(x_2) = Hf'(x_2) > 0$ , and  $V(x_1) = Hf'(x_1) < 0$ , so that  $H$  takes the form indicated.

This proves that the density function  $\rho$  is uniquely determined by the function  $f(x)$  up to the choice of the constant  $H$ .

For the existence of a suitable density function  $\rho(s)$ , given  $f(x)$ , define  $\rho(s)$  by equation (5). Then

$$\begin{aligned} V(X) &= \int_{x_0}^X \rho(s(x)) \frac{ds}{dx} dx = \int_{x_0}^X Hf''(x) dx \\ &= Hf'(X) - Hf'(x_0) = Hf'(X), \end{aligned}$$

so that the pair  $(f(x), \rho(s))$  form a weighted catenary, and then the constant  $H$  must have the form indicated in equation (6).  $\square$

**Example 1** (catenary).

$$f(x) = \frac{1}{B} \cosh Bx + C.$$

Then

$$f'(x) = \sinh Bx, \quad \sqrt{1 + f'(x)^2} = \cosh Bx$$

and  $f''(x) = B \cosh Bx$ , so

$$\rho(s(x)) = HB, \quad \text{constant.}$$

*Comment:* This is a kind of inverse to the original catenary problem of asking for the shape taken by a uniformly weighted hanging chain. It tells us that *only* for a uniform chain will the resulting curve be a catenary.

**Example 2** (parabola).

$$f(x) = ax^2 + b, \quad f'(x) = 2ax, \quad f''(x) = 2a.$$

Then

$$\rho(s(x)) = 2aH \left/ \frac{ds}{dx} \right.$$

and

$$\int \rho(s) ds = \int \rho(s(x)) \frac{ds}{dx} dx = 2aH \int dx,$$

and

$$\int_{s(x_1)}^{s(x_2)} \rho(s) ds = 2aH(x_2 - x_1).$$

*Comment:* This is in a sense the answer to Galileo's original question of finding a mechanical method for drawing a parabola. It tells us exactly how we have to weight a chain so that it will hang in the form of a parabola. The key is that the weight distribution has to be uniform in the horizontal direction. For this reason, it is usually stated that the cables on a suspension bridge will hang in the shape of a parabola, since the weight of the cables themselves, together with that of the vertical support cables holding up the roadway, is small compared to the weight of the roadway, which has (in essence) a uniform horizontal weight distribution.

A curious by-product of the equation for the density  $\rho$  in the case of a parabola is that, since  $ds/dx \geq 1$ , with equality only where  $f'(x) = 0$ , it follows that  $\rho(s(x))$  is maximum at the vertex and then decreases monotonically as  $|f'(x)|$  increases. In other words, the chain has to be weighted the most near the vertex and then decrease as the steepness of the curve increases. As a result, if Saarinen had decided that he found a parabolic arch most pleasing esthetically, he would have been faced with the paradox that in order to have the line of thrust be everywhere directed along the arch, the arch would have to be thickest at the top and taper down toward the bottom, which would be both ungainly esthetically and potentially disastrous structurally.

**Example 3** (circular arc).

$$x = R \cos \theta, \quad y = R \sin \theta, \quad \pi < \theta_1 \leq \theta \leq \theta_2 < 2\pi.$$

Then

$$f'(x) = \frac{dy}{d\theta} / \frac{dx}{d\theta} = -\cot \theta, \quad f''(x) = -\frac{1}{R} \csc^3 \theta,$$

while

$$s = R\theta, \quad ds/dx = -\csc \theta > 0.$$

So

$$\rho(s(x)) = \frac{H}{R} \csc^2 \theta = HR/y^2.$$

*Comment:* As noted earlier, a full semicircle cannot be a weighted catenary. We see explicitly that the density function would tend to infinity. However, any circular arc that is short of a semicircle at both ends can be realized (at least in theory) as a weighted catenary. The weighting is simply

$$\rho(s) = \frac{H}{R \cos^2(s/R)}, \quad \pi R < s_1 \leq s = R\theta \leq s_2 < 2\pi R.$$

Although this example may seem a typical "text-book example", of no practical interest but merely a case that can be worked out explicitly, the opposite is true. A monumental "Gateway Arch" had been proposed as an entryway to a planned 1942 International Exposition in Rome, but it was derailed by the war. The arch was to be in the form

of a semicircle, and the question of exactly how much it should be tapered was potentially critical. (See [12] for more on this subject.)

Note that this theorem answers the first part of Question 1 above about the existence of a density function yielding the particular “weighted catenary” shape of the Gateway Arch. It also answers Question 2 concerning the term “weighted catenary” itself, which turns out to have essentially no content beyond “convex curve”. Before turning to the second part of Question 1 regarding the exact form of the density function that yields the curve underlying the Gateway Arch, we prove a strong form of the converse to the theorem.

**Converse of Theorem 1.** *Let  $C$  be a curve with endpoints  $P_1 = (x_1, y_1, 0)$  and  $P_2 = (x_2, y_2, 0)$ , where  $x_1 < x_2$ . Assume that  $C$  is parameterized by arc length  $s$  in the form*

$$\mathbf{X}(s) = (x(s), y(s), z(s)), \quad P_1 = \mathbf{X}(s_1), \quad P_2 = \mathbf{X}(s_2).$$

*Let  $\rho(s) > 0$  be a continuous function along  $C$ , taken as a density function. Assume gravity acts in the negative  $y$ -direction, and the pair  $(C, \rho(s))$  is in equilibrium. Then:*

1.  $C$  lies entirely in the  $x, y$ -plane.
2.  $C$  is in the form of a graph  $y = f(x)$ ,  $x_1 \leq x \leq x_2$ .
3.  $f'(x)$  is finite and monotone increasing for  $x_1 \leq x \leq x_2$ .

The idea underlying the proof is that if we know that  $C$  is a graph and  $(C, \rho(s))$  is in equilibrium, so that equation (2) above holds, where  $V(X)$  is defined by (3), then

$$-f'(x_1) = \frac{1}{H} \int_{x_1}^{x_0} \rho(s(x))(ds/dx)dx < \infty,$$

$$f'(x_2) = \frac{1}{H} \int_{x_0}^{x_2} \rho(s(x))(ds/dx)dx < \infty,$$

and for  $x_1 \leq x_3 < x_4 \leq x_2$ ,

$$f'(x_4) - f'(x_3) = \frac{1}{H} \int_{x_3}^{x_4} \rho(s(x))(ds/dx)dx > 0.$$

Hence condition 3 in the statement of the converse must hold.

In terms of the representation above for  $C$  as  $\mathbf{X}(s)$ , the unit tangent vector to  $C$  will be given by  $\mathbf{T} = d\mathbf{X}/ds$ . We denote the tension at the point  $\mathbf{X}(s)$  by  $\tau$ , and for the case considered here, the tension vector will be

$$(7) \quad \tau\mathbf{T} = (H, V, W),$$

where  $H$  is constant and  $W = 0$ , while  $V$  is given by (3). Then

$$(8) \quad \frac{d}{ds} \tau\mathbf{T} = \left(0, \frac{dV}{ds}, 0\right) = \left(0, \frac{dV}{dx} / \frac{ds}{dx}, 0\right) = (0, \rho(s), 0)$$

by (3), or

$$(9) \quad \frac{d}{ds} (\tau\mathbf{T}) = \rho(s)\mathbf{j},$$

where  $\mathbf{j}$  is the unit vector directed along the positive  $y$ -axis.

The proof of the above converse depends on the fact that equation (9) is the equation of equilibrium for an arbitrary curve joining the points  $P_1, P_2$  in the  $x, y$ -plane, with no special assumptions about its shape. (See [2], Chapter III, sections 1-3, for a detailed discussion of these questions in maximum generality. I would like to thank Joe Keller for providing this reference, as well as further helpful information. Note that in contrast to [2], we take the density  $\rho$  to mean weight, rather than mass, per unit length, so that the acceleration of gravity is absorbed into it.)

*Proof of the converse.* Let  $C$  be defined by

$$\mathbf{X}(s) = (x(s), y(s), z(s)), \quad s_1 \leq s \leq s_2,$$

where

$$\mathbf{X}(s_1) = (x_1, y_1, 0), \quad \mathbf{X}(s_2) = (x_2, y_2, 0), \quad x_1 < x_2,$$

and let  $\rho(s) > 0$  be a continuous density function along  $C$ . Then the pair  $(C, \rho(s))$  will be in equilibrium if and only if (9) holds; that is,

$$(10) \quad \frac{d}{ds} \left( \tau(s) \frac{dx}{ds} \right) = 0, \quad \frac{d}{ds} \left( \tau(s) \frac{dy}{ds} \right) = \rho(s),$$

$$\frac{d}{ds} \left( \tau(s) \frac{dz}{ds} \right) = 0.$$

Hence

$$(11) \quad \tau(s) \frac{dx}{ds}(s) = c,$$

$$(12) \quad \tau(s) \frac{dy}{ds}(s) = \tau(s_1) \frac{dy}{ds}(s_1) + \int_{s_1}^s \rho(\sigma) d\sigma,$$

$$(13) \quad \tau(s) \frac{dz}{ds}(s) = d,$$

where  $c$  and  $d$  are constants. We claim first that  $C$  must lie completely in the  $x, y$ -plane. If not, there must be a value  $s_0, s_1 < s_0 < s_2$ , where  $|z(s)|$  attains its maximum  $|z(s_0)| > 0$ . Then  $(dz/ds)(s_0) = 0$ , and the constant  $d$  in (13) must vanish. But if  $z(s) \neq 0$ , there must be some value  $s_3$  such that  $dz/ds$  is nonzero at  $s_3$  and hence in an interval around  $s_3$ . But then, since  $d = 0$ , it follows from (13) that  $\tau(s) \equiv 0$  on that interval. But that contradicts equation (9), since  $\rho(s) > 0$ . Hence we must have  $z(s) \equiv 0$ , and the curve lies entirely in the  $x, y$ -plane.

Exactly analogous reasoning using equation (11) shows that there cannot be a value  $s_4$  where  $dx/ds = 0$ . In fact, since  $x(s_2) = x_2 > x_1 = x(s_1)$ , there must be a point where  $dx/ds > 0$ . It follows that  $dx/ds > 0$  on the whole interval  $[s_1, s_2]$ , and hence  $x$  is a strictly monotone increasing function of  $s$  on that interval. Thus we have a monotone increasing inverse  $s(x)$ , and we can set

$$y(s(x)) = f(x), \quad x_1 \leq x \leq x_2,$$

defining  $C$  as a graph. It follows from equations (11) and (12) that

$$(14) \quad f'(x) = \frac{dy}{ds} \Big/ \frac{dx}{ds} = \frac{1}{c} \tau(s_1) \frac{dy}{ds}(s_1) + \frac{1}{c} \int_{s_1}^s \rho(\sigma) d\sigma.$$

Hence  $f'(x)$  is a monotone increasing function of  $s$  and therefore of  $x$ . This completes the proof of the converse.  $\square$

Note that it follows from the above that

$$(15) \quad f''(x) = \frac{d}{ds} f'(x) \Big/ \frac{dx}{ds} = \frac{1}{c} \rho(s(x)) \frac{ds}{dx} > 0.$$

Since  $c$  is the horizontal component of the tension vector, it is the quantity we denoted earlier by  $H$ , and equation (15) is the same as equation (5) above.

We now come to our principal example of the theorem.

**Definition 1.** A *flattened catenary* is a curve of the form  $y = f(x) = A \cosh Bx + C$ , or

$$y = Dg(x) + C, \quad D = AB,$$

where  $0 < D < 1$ , and

$$g(x) = \frac{1}{B} \cosh Bx$$

is a catenary.

**Example 4** (flattened catenary). We examine this case in the following theorem.

**Theorem 2.** Let  $C$  be a weighted catenary defined by the pair  $(f(x), \rho(s))$ , and let  $p(x) = \rho(s(x))(ds/dx)$ . Assume that coordinates are chosen so that  $f'(0) = 0$ . Then  $p(x)$  will be of the form

$$p(x) = ay + b = af(x) + b$$

for some constants  $a > 0$  and  $b$  if and only if  $f(x)$  is a flattened catenary:

$$f(x) = A \cosh Bx + C,$$

for constants  $A, B, C$ .

*Proof.* We have for all  $X$  in  $[x_1, x_2]$  that  $f'(X) = V(X)/H$ , where

$$V(X) = \int_{x_0}^X p(x) dx, \quad H = \int_{x_1}^{x_2} \frac{p(x)}{|f'(x_1)| + |f'(x_2)|} dx.$$

Then

$$f''(X) = V'(X)/H = p(X)/H.$$

Suppose first that

$$(16) \quad y = f(x) = A \cosh Bx + C.$$

Then  $f'(x) = AB \sinh Bx$  and

$$f''(x) = AB^2 \cosh Bx = B^2(f(x) - C).$$

Then

$$(17) \quad p(x) = Hf''(x) = af(x) + b,$$

where

$$(18) \quad a = HB^2 > 0, \quad b = -HB^2C.$$

Conversely, suppose that  $p(x) = af(x) + b$ . Then  $f''(x) = (af(x) + b)/H$ . Let

$$g(x) = f(x) + b/a.$$

Then  $g'(0) = f'(0) = 0$ , and

$$g''(x) = f''(x) = ag(x)/H.$$

It follows that  $g(x) = A \cosh Bx$  where  $B^2 = a/H$ . Hence

$$f(x) = g(x) - b/a = A \cosh Bx + C,$$

where

$$(19) \quad A = f(0) + b/a, \quad B = \sqrt{a/H}, \quad C = -b/a. \quad \square$$

### Historical Note

The flattened catenary was studied in the nineteenth century by a number of authors in France, England, Ireland, and Scotland. The first may have been Villarceau [19]; see Heyman [6], p. 48. Others include W. J. M. Rankine [13] and A. M. Howe [9]. (I thank William Thayer for these latter references and for alerting me to the connection between the equation for the Gateway Arch and the nineteenth-century application discussed here.) In those works, the flattened catenary is often referred to as a "transformed catenary", and the term "two-nosed catenary" is used for a highly flattened catenary, for reasons explained below. The context in which it arises is in answer to the question: What is the ideal shape of an arch that supports a horizontal roadway held up by evenly packed dirt on top of the arch? The vertical load applied to each point on the arch is then proportional to the distance from that point up to the level of the roadway. If we write the equation of the arch as  $y = f(x)$ , where  $y$  is the distance from the roadway down to a given point on the arch, then the total vertical thrust acting on the part of the arch from the apex to an arbitrary point of the arch will correspond to the weight of a hanging chain from the minimum point to a given point, and it will be of the form  $\int p(x) dx$ , where  $p(x)$  is a linear function of  $y$ . The same reasoning as in the above theorem then shows that  $f(x)$  should be of the form (16): a flattened catenary. (See, for example, Heyman [6], pp. 48-49.)

We now come to the Gateway Arch itself. At first sight it would appear to be a very different situation from that of an arch supporting a bridge, since it supports nothing except its own weight. On closer look, however, the fact that the Arch is tapered means that the weight supported at each level is not determined just by the length of the arch above it but by an integral with respect to arclength of a function exactly analogous to the density function for a weighted chain. The physical Gateway Arch is reasonably well modeled as one of a class of geometric shapes that are of interest in their own right.

## Tube-like Domains

Two chance meetings in 1939 led to one of the major advances in differential geometry, the intrinsic proof by Chern of an  $n$ -dimensional Gauss-Bonnet theorem. The first occurred when mathematician Hermann Weyl attended a lecture by statistician Harold Hotelling in which Hotelling described a new result of his and posed an open question. Hotelling had been led by a statistical problem to pose the question: Given  $r > 0$ , what is the probability that a point chosen at random on the  $n$ -sphere would lie within a distance  $r$  of a given curve  $C$ ? The answer amounts to computing the volume of the domain consisting of all points whose distance to  $C$  is less than  $r$  and comparing that to the total volume of the sphere. Hotelling was able to answer the question in both Euclidean  $n$ -space and the  $n$ -sphere for  $r$  sufficiently small [8]. The answer in Euclidean space was that the volume is equal to the length of  $C$  times the volume of an  $(n - 1)$ -ball of radius  $r$ , with an analogous result for the  $n$ -sphere. The answer was perhaps a surprise, since one might think that the volume of the “tube” around the curve  $C$  would depend somehow on the twists and turns in the curve, and not simply on its total length.

What Hotelling wanted to know, and was unable to derive, was a formula for the case that the curve  $C$  was replaced by a surface of two or more dimensions. Hermann Weyl [22] provided the answer in spectacular fashion: for a compact smooth submanifold  $S$  of Euclidean space, the volume of the domain consisting of all points lying within a fixed distance  $r$  of  $S$  is given, for  $r$  sufficiently small, by a polynomial in  $r$  whose coefficients are *intrinsic* quantities associated with  $S$ , and do not depend on the way that  $S$  is immersed in space. Most strikingly, the top coefficient does not even depend on the geometry of  $S$ , but only its topology; it is just a constant times the Euler characteristic  $\chi$  of  $S$ .

The other chance meeting in 1939 took place between André Weil and Lars Ahlfors in Finland. Weil first learned about the classical Gauss-Bonnet theorem from Ahlfors, who expressed an interest in having a generalization to higher dimensions in order to develop a higher-dimensional Nevanlinna theory of value distribution ([20], p. 558, [21]). When Weil came to America several years later he met Carl Allendoerfer and learned that both Allendoerfer and Werner Fenchel had independently proved a Gauss-Bonnet theorem for manifolds embedded in Euclidean space, using Weyl’s formulas. Remembering Ahlfors’s desire for a general Gauss-Bonnet theorem, Weil was able, together with Allendoerfer, to prove the Gauss-Bonnet theorem for all Riemannian manifolds ([1], [20], p. 299). Their proof went via Weyl’s formulas together with local embedding theorems. Shortly after, using the appropriate form of the integrand from these

earlier approaches, Chern was able to provide an intrinsic proof of the Gauss-Bonnet theorem without the detour through embeddings.

Our interest here is in a generalization of Hotelling’s result in a different direction. We go back to the case of a curve  $C$  in Euclidean space, but are interested in much more general domains.

**Notation 1.** A curve  $C$  in  $\mathbb{R}^n$  will be denoted by

$$X(s) = (x_1(s), \dots, x_n(s)), 0 \leq s \leq L,$$

$s =$  arc length parameter,

$$e_1(s) = dX/ds$$

the unit tangent vector,

$$e_1(s), e_2(s), \dots, e_n(s)$$

a positively oriented orthonormal frame field along  $C$ .

**Definition 2.** A *tube-like domain*  $T$  with centerline  $C$  is the image under a diffeomorphism  $F$  of a domain  $D$  in  $\mathbb{R}^n$ , with coordinates  $u_1, \dots, u_n$ , such that  $0 < u_1 < L$ , and for any  $t$ ,  $0 < t < L$ , the intersection  $D_t$  of  $D$  with the hyperplane  $u_1 = t$  contains the point  $(t, 0, \dots, 0)$ . The map  $F$  satisfies  $F(u_1, u_2, \dots, u_n) = X(u_1) + u_2 e_2(u_1) + \dots + u_n e_n(u_1)$ , with  $u_1 = s =$  arc length parameter along  $C$ .

For  $0 < s < L$ , let  $T_s$  be the intersection of  $T$  with the normal hyperplane to  $C$  at the point  $X(s)$ . Then  $F : D_s \rightarrow T_s$  is a linear transformation mapping  $D_s$  isometrically to  $T_s$ .

**Theorem 3.** Let  $T$  be a tube-like domain in  $\mathbb{R}^n$  with centerline  $C$ . Then the volume  $V$  of  $T$  is given by

$$(20) \quad V = \int_0^L V(T_s) ds - \int_0^L \left[ \int_{D_s} \vec{k}(s) \cdot (F(s, u_2, \dots, u_n) - X(s)) du_2 \dots du_n \right] ds,$$

where  $\vec{k}(s) = d^2X/ds^2$  is the curvature vector, and for each fixed value of  $s$ , the vector

$$F(s, u_2, \dots, u_n) - X(s) = u_2 e_2(s) + \dots + u_n e_n(s)$$

traces out the normal section  $T_s$  of  $T$  as the point  $(s, u_2, \dots, u_n)$  ranges over the parameter domain  $D_s$ .

*Proof.* Let  $J = \det dF$  be the Jacobian of the map  $F$ . If  $Y = (y_1, \dots, y_n) = F(u_1, \dots, u_n)$ , then

$$\begin{aligned} J &= \frac{\partial(y_1, \dots, y_n)}{\partial(u_1, \dots, u_n)}, \quad V(T) = \int_D J du_1 du_2 \dots du_n \\ &= \int_0^L \left[ \int_{D_s} J du_2 \dots du_n \right] ds. \end{aligned}$$

Now write

$$Y = X(u_1) + u_2 e_2 + \dots + u_n e_n,$$

with  $e_j = e_j(u_1)$ ,  $0 \leq u_1 \leq L$ . Then

$$\frac{\partial Y}{\partial u_1} = e_1 + u_2 \frac{de_2}{du_1} + \cdots + u_n \frac{de_n}{du_1} \quad \text{and} \quad \frac{\partial Y}{\partial u_i} = e_i, \\ i = 2, \dots, n,$$

and

$$\frac{\partial Y}{\partial u_1} \wedge \cdots \wedge \frac{\partial Y}{\partial u_n} = \frac{\partial Y}{\partial u_1} \wedge e_2 \wedge \cdots \wedge e_n = J e_1 \wedge \cdots \wedge e_n.$$

Write

$$\frac{\partial e_i}{\partial u_1} = \sum_{j=1}^n a_{ij} e_j = a_{i1} e_1 + \cdots,$$

where the coefficients are given by

$$a_{i1} = \frac{de_i}{du_1} \cdot e_1 = -\frac{de_1}{du_1} \cdot e_i = -\vec{k} \cdot e_i,$$

since the curvature vector  $\vec{k}(s) = \frac{d^2 X}{ds^2}$  equals  $\frac{de_1}{ds}$ . So

$$\sum_{i=2}^n a_{i1} u_i = -\vec{k} \cdot \sum_{i=2}^n u_i e_i = -\vec{k}(u_1) \cdot (Y - X),$$

where  $a_{ij} = a_{ij}(u_1)$ ,  $X = X(u_1)$ ,  $Y = Y(u_1, \dots, u_n)$ .

Finally,

$$\frac{\partial Y}{\partial u_1} = \left(1 + \sum_{i=2}^n a_{i1} u_i\right) e_1 + \cdots,$$

where the omitted terms involve the vectors  $e_2, \dots, e_n$ , and will vanish after taking the wedge product with  $e_2$  through  $e_n$ . Hence,

$$J = 1 + \sum_{i=2}^n a_{i1} u_i = 1 - \vec{k}(u_1) \cdot (Y - X(u_1)),$$

as asserted.  $\square$

**Note.** Different versions of this theorem have been known at least since 1978, when Michael Raugh [14] studied the case  $n = 3$  in the context of the growth and structure of tendrils. Another discussion of the case  $n = 3$  appears in Chapter 3 (*Cálculo de várias variáveis*) of [3], pp. 202-210. Both of those discussions assume the existence of a Frenet frame field for the centerline of the domain. A recent paper of Raugh [15] also treats the  $n$ -dimensional case.

**Note.** For the proof to be valid, one needs the Jacobian to be positive. If it were to change sign, then there would be cancellations, and the integral would not represent the actual volume. Since our definition of a tube-like domain includes the assumption that the defining map is a diffeomorphism, the Jacobian cannot change sign. However, if one starts with a domain  $D$  and tries to represent it as a tube domain, then one must check the positivity of the Jacobian. We shall return to this point later on.

**Corollary 1.** *If the point  $X_s$  is the centroid of  $T_s$  for each  $s \in [0, L]$ , then*

$$(21) \quad V(T) = \int_0^L V(T_s) ds.$$

Note that it suffices to consider points where  $\vec{k}(s) \neq 0$ .

*Proof.* For each value of  $s$ , the inner integral in the second term of the formula for  $V(T)$  is a sum of terms consisting of integrals over  $D_s$  of terms of the form  $c_j u_j$ , where  $c_j$  is a constant. But if  $X(s)$  is the centroid of  $T_s$ , then the origin is the centroid of  $D_s$ , and each of those integrals must vanish.  $\square$

**Definition 3.** Under the hypotheses of Corollary 1, the centerline  $C$  is called a *centroid curve* of  $T$ .

A given domain may have more than one centroid curve; for example, a Euclidean ball can be viewed as a tube-like domain with respect to any diameter as centerline, and that diameter will be a centroid curve for the ball.

**Corollary 2** (Hotelling's theorem). *If  $T$  is a tube domain in  $\mathbb{R}^n$  of radius  $r$  around a curve  $C$  of length  $L$ , then*

$$V(T) = \omega_{n-1} r^{n-1} L,$$

where  $\omega_k$  is the volume of the ball of radius  $r$  in  $\mathbb{R}^k$ .

*Proof.* In this case  $D$  is a cylinder of radius  $r$  and height  $L$ . For  $0 < s < L$ , the normal section  $T_s$  of  $T$  is an  $(n-1)$ -ball of radius  $r$  centered at the point  $X(s)$  of  $C$ .  $\square$

A number of statisticians returned to questions related to Hotelling's theorem after a lapse of many years. See the paper [11] and further references given there.

### Special Cases

$n = 2$ . Then  $e_2$  is uniquely determined by  $e_1 = dX/ds$ . The orthogonal section  $T_s$  is an interval  $I_s$ , and the second term vanishes if  $X(s)$  is the midpoint of  $I_s$  at all points where the curvature is nonzero.

**Note.** For  $n = 2$ ,  $\vec{k}(s) = \kappa(s)e_2(s)$ , where the curvature  $\kappa$  can be positive, negative, or zero.

$n = 3$ . If  $\vec{k}(s) \neq 0$  at a point, then the same holds in an interval around the point, and we may define  $e_2 = \vec{k}(s)/\kappa(s)$ , where  $\kappa(s) = |\vec{k}(s)|$  is the curvature. Then  $e_2(s)$  is a smooth unit vector field on the interval, and  $e_1, e_2$ , together determine  $e_3$ . Since  $de_1/ds = \vec{k} = \kappa e_2$ , we have  $a_{21} = -\kappa$ ,  $a_{31} = 0$ , and  $J = 1 - \kappa(s)u_2$ . As already noted, this equation also holds whenever  $\kappa = 0$ . Therefore, for  $n = 3$ ,

$$V(T) = \int_0^L \left[ \int_{T_s} (1 - \kappa(s)u_2) du_2 du_3 \right] ds \\ = \int_0^L V(T_s) ds - \int_0^L \kappa(s)M_s ds,$$

where  $M_s$  is the moment of  $T_s$  about the line through  $X(s)$  in the  $e_3$  (binormal) direction. It follows that for (21) to hold, one does not need  $X(s)$  to be the centroid of the orthogonal section, but only that the moment of that section be zero with respect to the line through  $X(s)$  in the binormal direction, at all points where  $\kappa \neq 0$ . (See Raugh [15].)

An analogous statement holds for arbitrary  $n$ .

Another formulation is

$$\begin{aligned} \int_{D_s} J du_2 \dots du_n &= \int_{D_s} \left(1 + \sum_{i=2}^n a_{i1} u_i\right) du_2 \dots du_n \\ &= V(T_s) + \sum_{i=2}^n a_{i1} M_i, \end{aligned}$$

where  $M_i$  is the  $i$ -th moment of  $D_s$ , and  $a_{i1} = a_{i1}(s)$ .

### Gateway Arch Specifications

The *shape of the Gateway Arch* is a polyhedral surface, the piecewise linear approximation by quadrilaterals to the surface  $S$  of the theoretical Gateway Arch, with two of the sides of each quadrilateral lying on designated straight-line segments of  $S$ .

**Note.** The surface just described is an abstract mathematical representation of the surface of the physical Gateway Arch, that surface consisting of the exterior of a set of stainless steel plates subject to a variety of distortions. Among those are the (un-even) expansion and contraction with changes of temperature, deflection due to wind forces, and deformation under the moving weight distribution of the internal transportation system carrying passengers to the top.

The *theoretical Gateway Arch* is a tube-like domain  $T$  whose centerline  $C$  is the reflection in a horizontal plane of a flattened catenary. The curve  $C$  is given by the equation

$$(22) \quad y = 693.8597 - A \cosh Bx,$$

where  $A$  and  $B$  are the numerical constants

$$(23) \quad A = 68.7672, B = .0100333,$$

$$(24) \quad -299.2239 \leq x \leq 299.2239,$$

and  $x, y$  represent distances measured in feet. We may also write the equation of  $C$  as

$$(25) \quad y = 625.0925 - Y$$

where

$$(26) \quad Y = A \cosh Bx + C = A(\cosh Bx - 1)$$

represents the vertical distance down from the vertex of  $C$ .

The orthogonal section  $T(x, y)$  at the point  $(x, y)$  of  $C$  is in the form of an equilateral triangle of area  $Q(x, y)$ , where the sides of  $T(x, y)$  when  $x = 0$  (at the vertex of  $C$ ) have length 17 feet and when  $y = 0$  (at the base of the arch) have length 54 feet. The area  $Q(x, y)$  is defined by interpolating

linearly in  $Y$  between the corresponding areas so defined:

$$Q(x, y) = Q_v + (Q_b - Q_v)Y/Y_b,$$

where  $Q_v$  is the cross-sectional area at the vertex, and  $Q_b$  is the corresponding area at the base. Since the area  $Q$  of an equilateral triangle is equal to

$$Q = \frac{\sqrt{3}}{4} d^2,$$

where  $d$  is the length of a side, and since  $d = 17$  at the vertex  $v$  and  $d = 54$  at the base, we find

$$Q_v = 125.1406, \quad Q_b = 1262.6651,$$

while  $Y_b = Y|_{y=0} = 625.0925$ . It follows that

$$\begin{aligned} Q(x, y) &= 125.1406 + 1.81977Y \\ &= 1262.6651 - 1.81977y. \end{aligned}$$

Using the relation above between the area  $Q$  and the side  $d$  of an equilateral triangle, we obtain the exact dimensions of the cross-sections at each height  $y$ . It remains to specify that the positions of the triangles in the planes orthogonal to the curve  $C$  are symmetric with respect to the  $x, y$ -plane, with one vertex lying in that plane and interior to the curve  $C$ , and that the centroid of each of those triangles lies on  $C$ , so that  $C$  is a centroid curve of the Arch.

It follows from these specifications that the surface of the theoretical Gateway Arch is the union of three surfaces: a pair of ruled surfaces that are symmetric images with respect to the  $x, y$ -plane and meet in a curve lying in the  $x, y$ -plane, called the *intrados* of the Arch, together with a cylindrical surface orthogonal to that plane: the *extrados*. The maximum height of the extrados occurs at the point directly above the vertex of  $C$  where  $x = 0$  and  $y = y_v = 625.0925$ . At this point the orthogonal plane is vertical, and the distance from the vertex of  $C$  to the base of the triangle is one third of the length of the median, or  $17\sqrt{3}/6 = 4.90748$ . Adding this to  $y_v$  gives for the maximum height of the Arch  $h = 629.99998$ .

The width of the Arch is a bit more subtle. We take the points where  $y = 0$  on the centroid curve  $C$  to represent ground level. At those points,  $x = 299.2239$  and  $-299.2239$  so that the width of the curve at ground level is 598.4478. The equilateral triangle representing the cross-section at those points has side 54, and hence the distance to the outer curve of the Arch is

$$54\sqrt{3}/6 = 9\sqrt{3} = 15.58846.$$

Adding this on each side to the width of  $C$  at ground level gives a total width for the Arch of 629.6247. However, since the curve  $C$ , although steep, is not vertical at ground level, the cross-section of the Arch is not horizontal, and the actual outer width is slightly larger. However, one sees that the dimensions of the centroid curve together with the size and shape of the cross-sections produce an arch



that for all practical purposes has exactly the same total height as width. It may be worth noting, however, since it is sometimes a source of confusion, that the centroid curve is distinctly taller than wide, and the same is even more true of the inner curve of the Arch, whose height is 615.3 feet, and width is 536.1. For more on this and related matters, see [12].

### Curvature Computations

Let  $y = f(x) = A \cosh Bx = D(\frac{1}{B} \cosh Bx)$  define a flattened catenary, where  $A, B$  are positive constants, and  $D = AB$ . Then  $f''(x) = B^2y > 0$  so that  $f$  is convex and the curvature  $\kappa > 0$ . Specifically,

$$\kappa = \frac{B^2y}{(B^2y^2 + 1 - D^2)^{3/2}}.$$

For a catenary,  $D = 1$  and  $\kappa = 1/By^2$ , with a maximum of  $B$  at the vertex  $(0, 1/B)$  and decreasing monotonically to zero as  $y \rightarrow \infty$ .

In general, let  $g(y) = (B^2/\kappa)^2 = B^4R^2$ ,  $R = 1/\kappa$ . Then  $dR/dy \geq 0 \Leftrightarrow g'(y) \geq 0$ . But one finds

$$g'(y) = \frac{2}{y^3}(B^2y^2 + 1 - D^2)^2(2B^2y^2 - 1 + D^2)$$

and since  $y \geq A$ ,  $B^2y^2 + 1 - D^2 \geq 1$ , and

$$g'(y) > 0 \Leftrightarrow 2B^2y^2 > 1 - D^2.$$

Case 1:  $D > 1$ . An "elongated" catenary, obtained by stretching a catenary uniformly in the vertical direction. Then the right-hand side is negative,  $g'(y) > 0$  along the whole curve, so that  $R$  has a minimum and  $\kappa$  a maximum at the vertex, where  $x = 0, y = D/B$ , and  $\kappa = BD$ .

Case 2:  $D = 1$ . A catenary, for which, as we have seen,  $R$  has a minimum value of  $1/B$  at the vertex and  $R \rightarrow \infty$  as  $y \rightarrow \infty$ .

Case 3:  $0 < D < 1$ . Then there are two possibilities.

Case 3a:  $D^2 \geq 1/3$ . Since  $y$  takes its minimum value  $A$  at the vertex, we have

$$2B^2y^2 + D^2 \geq 2B^2A^2 + D^2 = 3D^2 \geq 1$$

so that  $g'(y) \geq 0$  along the whole curve, and as in cases 1 and 2, the curvature has its maximum value of  $BD$  at the vertex.

Case 3b:  $D^2 < 1/3$ . Then  $1 - D^2 > 2D^2 = 2B^2A^2$ , and we have  $g'(y) < 0$  in the interval

$$y|_{x=0} = A \leq y < \left(\frac{1 - D^2}{2B^2}\right)^{1/2}.$$

In this case, the vertex of the curve will be a local *minimum* of the curvature, which will increase monotonically until  $y$  reaches the value indicated on the right, where the curvature will reach its maximum. The curve is in this case sometimes referred to as a *two-nosed catenary*.

Figure 2 shows a two-nosed catenary for which  $D = 1/3$ , obtained by stretching a catenary uniformly in the horizontal direction by a factor of

3, which is equivalent, up to a similarity, to compressing it vertically by a factor of 3.

It may be worth observing that when illustrating geometric graphs, one often sees different scales used on the two axes. For a parabola, the resulting curve will still be a parabola, but for a catenary, the result will be a flattened or elongated catenary. As the figure, as well as the above calculation, makes clear, the shape may be very different geometrically from a true catenary. Such differences are obvious when different scales in the horizontal and vertical directions turn a circle into an ellipse, but they are often overlooked for noncompact curves such as the catenary.

Finally, we note that for the centroid curve of the Gateway Arch, we have  $D = AB = .69$ , well above the critical value  $D = \sqrt{3}/3 = .577\dots$ , so that the curvature will take its maximum value at the vertex, where it is equal to  $DB = .0069$  so that the radius of curvature at that point is  $R = 145$  feet, and it grows monotonically from there. Since all points of the cross-sectional equilateral triangles that define the Gateway Arch are less than 54 feet (and in fact,  $< 54\sqrt{3}/3 < 32$  feet) from the centroid curve, it follows that the Jacobians entering into the computations leading up to equation (21) are all positive, where the Gateway Arch is considered as a tube-like domain.

**Note.** This paper is an expanded version of the Roever Lecture given by the author at Washington University in St. Louis on March 24, 2009. William H. Roever was born in St. Louis, and after receiving his Ph.D. at Harvard University and teaching at MIT, he returned to Washington University, where he served as chairman of the mathematics department. He wrote two papers on the curves of light formed by reflection under particular circumstances [16], [17]. He would undoubtedly have been fascinated by the light curve that forms on the Gateway Arch when the sun is at a certain angle.

### Final Note

Although the essence of Hooke's dictum relating the standing arch to the hanging chain seems clear, there are subtle issues that arise, some involving the fact that a standing arch is a three-dimensional body, whereas the hanging chain is modeled just as a curve. In fact, a careful analysis of the stability of an arch and its relation to the geometry of the arch is a subject about which whole books could be (and have been) written. (See, for example, [6].)

Here is an example of a seemingly paradoxical fact that does not seem to have been mentioned anywhere. As we noted at the outset, the fact that the shape of a hanging chain is a catenary is often derived by using the calculus of variations, finding the curve of given length that has the lowest center of gravity. It follows then, according to Hooke, that the ideal shape of an arch would be an upside-down

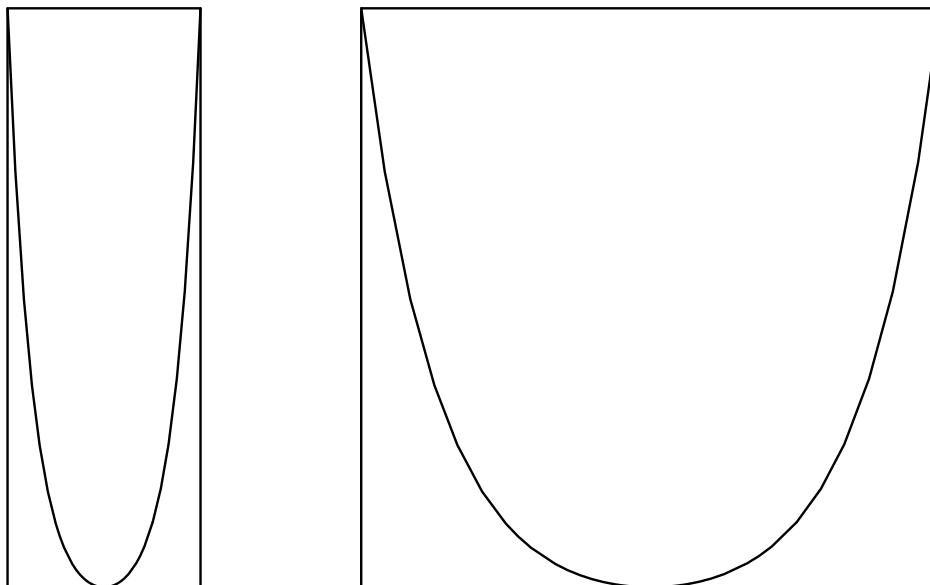


Figure 2. Left: catenary; right: 2-nosed catenary.

catenary, which would have the *highest* center of gravity among all arches of the same length. But that would appear to make it maximally unstable.

Some of these and related issues are discussed in the paper [12]. We hope to return to the subject for a fuller discussion at a later date.

**Acknowledgment.** My thanks to Silvio Levy and Paul Shields for valuable T<sub>E</sub>Xnical assistance.

## References

- [1] CARL ALLENDOERFER and ANDRÉ WEIL, The Gauss-Bonnet theorem for Riemannian polyhedra, *Trans. Amer. Math. Soc.* **53** (1943), 101-29.
- [2] STUART S. ANTMAN, *Nonlinear Problems of Elasticity*, New York, Springer-Verlag 1995.
- [3] VERA L. X. FIGUEIREDO, MARGARIDA P. MELLO, and SANDRA A. SANTOS, *Cálculo com Aplicações: Atividades Computacionais e Projetos*, Coleç ao IMECC, Textos Didáticos 3, Campinas, Instituto de Matemática, Estatística, e Computação Científica 2005.
- [4] GALILEO GALILEI, *Two New Sciences*, Translated by Henry Crew and Alfonso de Salvio, New York, Dover Publications 1954, originally published by The Macmillan Company, 1914.
- [5] HERMAN H. GOLDSTINE, *A History of the Calculus of Variations from the 17th through the 19th Century*, New York, Springer-Verlag 1980.
- [6] JACQUES HEYMAN, *The Masonry Arch*, Chichester, Ellis Horwood Limited, 1982.
- [7] ———, *The Science of Structural Engineering*, London, Imperial College Press, 1999.
- [8] HAROLD HOTELLING, Tubes and spheres in  $n$ -space and a class of statistical problems, *Amer. J. Math.* **61** (1939), 440-60.
- [9] A. MALVERD HOWE, *A Treatise on Arches*, New York, John Wiley & Sons, 1897.
- [10] CHARLES INGLIS, *Applied Mechanics for Engineers*, Cambridge University Press, 1951.
- [11] MARK KNOWLES and DAVID SIEGMUND, On Hotelling's approach to testing for a nonlinear parameter in regression, *International Statistical Review*, **57** (1989), 205-20.
- [12] ROBERT OSSERMAN, How the Gateway Arch got its shape, *Nexus Network Journal*, **12** (2010).
- [13] W. J. MACQUORN RANKINE, *Miscellaneous Scientific Papers*, London, Charles Griffith and Company, 1881.
- [14] MICHAEL R. RAUGH, *Some geometry problems suggested by the shapes of tendrils*, Stanford University thesis, 1978.
- [15] ———, Some differential geometry motivated by coiling tendrils, <http://mikeraugh.interconnect.com/tg2008.html>.
- [16] WILLIAM H. ROEVER, Brilliant points and loci of brilliant points, *Annals of Math.* **3** (1902), 113-28.
- [17] ———, The curve of light on a corrugated dome, *Amer. Math. Monthly* **20** (1913), 299-303.
- [18] CLIFFORD TRUESDELL, *The Rational Mechanics of Flexible or Elastic Bodies, 1638-1788*, Introduction to Leonhard Euler *Opera Omnia* Ser. 2, Vol. X and XI, 1960.
- [19] ANTOINE J. F. YVON VILLARCEAU, *Sur l'établissement des arches de pont*, Paris, Imprimerie Impériale, 1853.
- [20] ANDRÉ WEIL, *Collected Papers*, Vol. 1, Springer-Verlag 1979.
- [21] ———, Letter to the author, March 21, 1990.
- [22] HERMANN WEYL, On the volume of tubes, *Amer. J. Math.* **61** (1939) 461-72.