LETTER FROM THE EDITOR

We have a little something for everyone in this issue of the magazine:

Jadie Adams, Gabriel Lopez, Casey Mann, and Nhi Tran contribute our lead article for this issue. They investigate a remarkable tile devised by Heinz Voderberg in 1936, with the bizarre property that two copies can completely enclose a third. They are led to a remarkable generalization of the Voderberg tile that tesselates the plane. I found it difficult to copy edit this article because I kept getting distracted by the beautiful diagrams.

Jocelyn Bell and Frank Wattenberg take a fresh look at a classic brainteaser: Suppose a dog at the center of a circular lake makes directly for a duck swimming along the edge. Given some information about the relative speeds of the dog and the duck, work out the path traced out by the dog. "Pursuit curve" problems of this sort have long provided entertainment for mathematicians, and Bell and Wattenberg derive some remarkable results.

Russell Gordon investigates the evolutes of cubic polynomials. Just as a smooth function has a tangent line at each point, so too does each point have a circle of best approximation. The curve traced out by the centers of all these circles is called the evolute. Vincent Matsko, for his part, parlays an interest in jump discontinuities into an exploration of quadratic functions involving the greatest integer function. Fans of differential geometry, not to mention horses, will enjoy I. H. Skau's and K. F. Kristensen's discussion of how to build an optimal race track.

The shorter pieces explore a prime number pattern, a novel approach to integration in spherical coordinates, exotic proofs of the fundamental theorem of algebra, and a (relatively) elementary approach to a famous theorem in topology. We round out the issue with problems, reviews, and a proof without words.

Jason Rosenhouse, Editor

ARTICLES

Your Friendly Neighborhood Voderberg Tile

JADIE ADAMS
Westminster College
Salt Lake City, UT 84105
jadieraeadams@gmail.com

GABRIEL LOPEZ
California State University, San Bernardino
San Bernardino, CA 92407
gabriel.lopez-1@colorado.edu

CASEY MANN
University of Washington, Bothell
Bothell, WA 98011
cemann@uw.edu

NHI TRAN
University of Washington, Bothell
Bothell, WA 98011
vitamin933@gmail.com

In memory of Branko Grünbaum.

In 1934, K. Reinhardt asked if any tile exists such that two copies can completely enclose a third copy [4]. In 1936, Reinhardt's student Heinz Voderberg answered this question in the affirmative by inventing a special shape, now known as the *Voderberg tile*, shown in Figure 1. In fact, the two copies of the Voderberg tile can enclose* a third and a fourth copy (Figure 1)! More generally, a tile T has the *r-enclosure property* if two copies T_1 and T_2 of T can be arranged so that the complement of $T_1 \cup T_2$ has a bounded component, the closure of which is the union of r non-overlapping copies of T (as discussed, for example, by Grünbaum and Shepherd [3, 4]). In Figure 1, we see that the Voderberg tile has the 1- and 2-enclosure property.
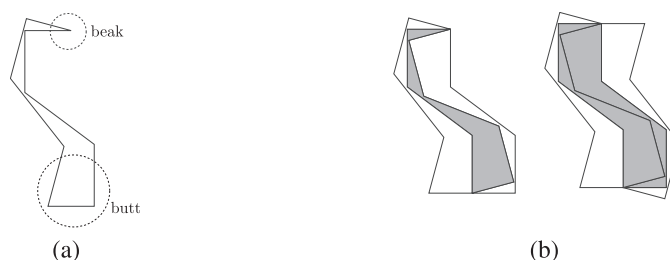

Figure 1 (a) The Voderberg tile. (b) The Voderberg tile has the 1-enclosure and the 2-enclosure property.

The enclosure property of the Voderberg tile is surprising and somewhat unintuitive, but this gangly tile's true beauty emerges in the tilings it generates, such as those shown in Figures 2(a) and 2(b). In this paper, we will explain how to construct a generalized

*Well, almost—see [5].

version of the tile in such a way that it can tessellate the plane, and we will go on to demonstrate some additional surprising properties that these tilings have.
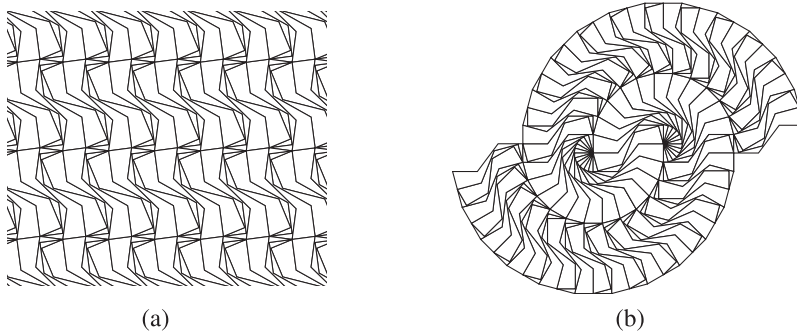

(a)   (b)

Figure 2   (a) A periodic tiling by the Voderberg tile. (b) A spiral tiling by the Voderberg tile.

Construction of a generalized Voderberg tile

The Voderberg tile was brought to a wide audience by Gardner [**1**], in 1977. Grünbaum and Shephard [**4**] describe a construction attributed to Goldberg [**2**] for a generalized family of tiles having members with the r-enclosure property for any $r \geq 1$. We will call these tiles *generalized Voderberg tiles* and denote them by V_n. For $n > 1$, these tiles are similar to the original Voderberg tile V_1, with the main differences being that they are thinner and that their "beaks" (see Figure 1) have n bends (or sections) instead of the single section of the original. Here we will reproduce and expound upon the construction in Grünbaum and Shephard [**4**] of the generalized Voderberg tile family.

Let n and k be positive integers with $k \gg 1$. Our goal is to describe the construction of V_n tiles with beak angle $\beta = \pi/k$. Having a beak angle that is commensurate with π allows for more kinds of spiral tilings to be formed from copies of the V_n tile, such as that in Figure 2. First we will give the construction of V_n that does not require β to be commensurate with π, then afterward we show how V_n with such specified β is obtained.

In the illustration of the construction in Figure 3 we use $n = 3$, but the construction described here is for general n. Start with 4 horizontal, parallel lines a_1, a_2, a_3, and a_4 spaced 1 unit apart, then construct the polygonal line $ABCD$ with $A \in a_3$, $B \in a_1$, $C \in a_4$, and $D \in a_2$ with right angles at B and C. Let α be the acute angle formed by AB and a_1. Let DE be the circular arc between a_2 and a_4 centered at A, and similarly let AF be the circular arc between a_1 and a_3 centered at D. Let EE' be $1/n$ of the arc DE, and divide the arc AF into n equal subarcs

$$AA_1, A_1A_2, \ldots, A_{n-1}F.$$

Let S be the polygonal line

$$AA_1A_2 \ldots A_{n-1}FBCE$$

and let S' be the result of rotating S about A until E coincides with E' (i.e., through an angle of $-\theta/n$). The tile bounded by S, S', and EE' is then a generalized Voderberg tile V_n. Note that V_n is a $(2n + 7)$-gon, and also observe that the beak angle β satisfies $\beta = \theta/n$ where $\theta = m(\angle DAE)$. In terms of the construction, the *beak* of the V_n tile is the thin part of the tile extending from F and F' to A, and the *butt* is the region near the segment EE'.
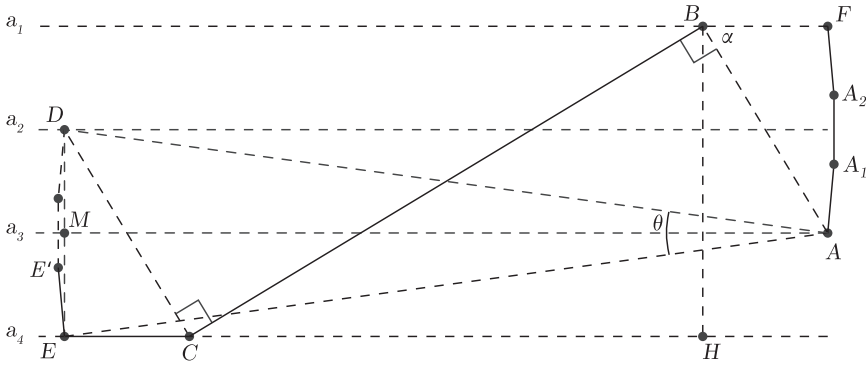
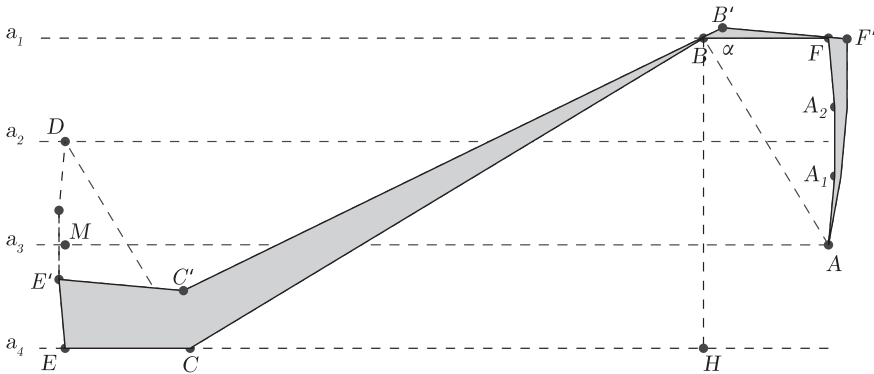**Figure 3** Construction of the generalized Voderberg tile.

Figure 4 Completed construction of a V_3 tile.

The angle $\alpha = \angle ABF$ in Figure 3 was chosen arbitrarily in the construction described above, but with some care we can specify α so that the beak angle β is a factor of π. To that end, let $M \in a_3$ be the midpoint of DE and let H be the point on the a_4 such that BH is perpendicular to BF. Notice that

$$\triangle AFB, \quad \triangle BHC, \quad \text{and} \quad \triangle CED$$

are all similar right triangles, since

$$\angle ABF = \angle CBH = \angle DCE = \alpha,$$

with $DE = AF = 2$ and $BH = 3$. Now

$$EC = 2\cot\alpha, \quad CH = 3\tan\alpha, \quad \text{and} \quad BF = 2\cot\alpha.$$

It follows that

$$AM = EC + CH + BF = 4\cot\alpha + 3\tan\alpha.$$

Now, in $\triangle AMD$ we see that $\angle DAM = \theta/2$, and so

$$\tan\theta/2 = \frac{1}{4\cot\alpha + 3\tan\alpha} = \frac{\tan\alpha}{4 + 3\tan^2\alpha}. \tag{1}$$

Set $T = \tan(\theta/2)$ and $x = \tan\alpha$ so that equation (1) gives

$$3Tx^2 - x + 4T = 0. \tag{2}$$

We solve equation (2) for x to get

$$\tan \alpha = \frac{1 \pm \sqrt{1 - 48 \tan^2(\theta/2)}}{6 \tan(\theta/2)}. \tag{3}$$

Now require that $\beta = \theta/n = \pi/k$, so that $\theta/2 = n\pi/2k$. Then we must have

$$1 - 48 \tan(\theta/2) = 1 - 48 \tan \frac{n\pi}{2k} \geq 0.$$

Solving this inequality for k gives

$$k \geq \frac{n\pi}{2 \tan^{-1}\left(\frac{1}{\sqrt{48}}\right)}. \tag{4}$$

Thus, based on n alone we can determine the range of allowable values of k.

For example, for a V_2 tile, we must have

$$k > \frac{\pi}{\tan^{-1}\left(\frac{1}{\sqrt{48}}\right)} \approx 21.9.$$

Let us choose $k = 25$. To construct a V_2 tile with beak angle $\beta = \pi/k = \pi/25$, we have

$$T = \tan \frac{n\pi}{2k} = \tan(\pi/25).$$

Thus, T in equation (2) is specified, and solving for x yields two roots x_1 and x_2, from which we determine two values of α:

$$\alpha = \tan^{-1} x_1 \approx 34.2611°,$$
$$\alpha = \tan^{-1} x_2 \approx 62.9389°.$$

Either of these values of α can be used to construct a V_2 with beak angle $\pi/25$, as in Figures 5 and 6. Notice that both shapes have the same length (from beak to butt). This is because in equation (1), the denominator of the first fraction is the length of AM, which depends only on θ. This was held constant in solving equation (2) for x.

In Figure 7, we see a V_2 tile exhibiting the 2-, 3-, and 4-enclosure property. Indeed, V_n tiles have the r- enclosure property for $r = 2n - 2, r = 2n - 1$, and $r = 2n$, which can be seen by noticing that that three enclosure values correspond to the manner in which the beaks meets the butts on the two enclosing tiles. Observe how the gray tiles meet in Figure 7. For general V_n, corresponding arrangements can be formed for the
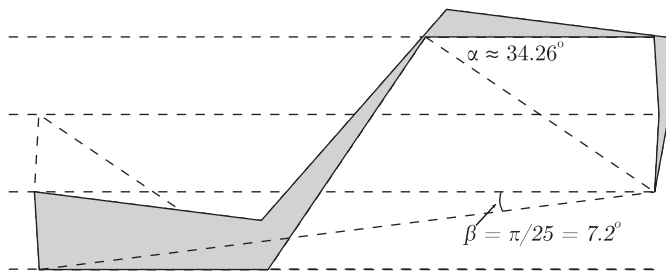
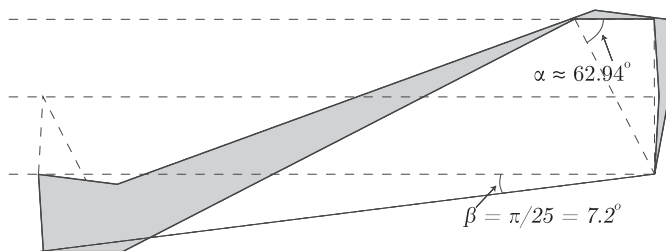**Figure 5**   A V_2 tile with $k = 25$ and beak angle $\beta = \pi/25$ and $\alpha \approx 34.2611°$

Figure 6   A V_2 tile with $k = 25$ and beak angle $\beta = \pi/25$ and $\alpha \approx 62.9389°$

enclosing (gray) tiles. Because of the number of beak segments and how the beak angle relates to the butt angle (at point E in the construction—that angle is exactly β more than a right angle), there will be room for exactly $2n - 2$, $2n - 1$, and $2n$ copies (respectively) in the middle, between the enclosing tiles.
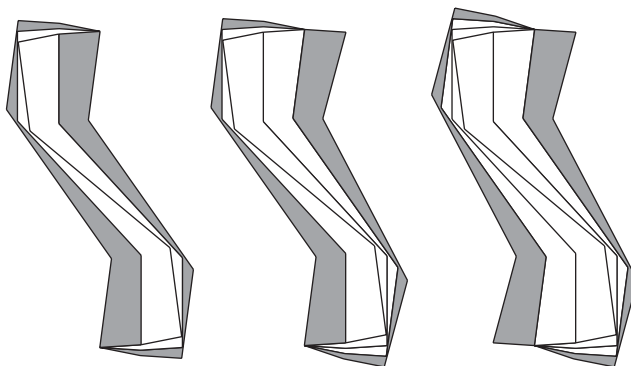

Figure 7   A V_2 tile exhibiting the 2-, 3-, and 4-enclosure property.

Problems from "G&S"

The contributions of Branko Grünbaum and G. C. Shephard to the development of a coherent and rigorous theory for tilings cannot be overstated, and much of their work is summarized in their magnum opus *Tilings and Patterns* [**3**]. In this work (and in others such as [**4**]) they were also generous in sharing open problems. In particular, these Johnny Appleseeds of tiling theory left two open problems for us which may be

answered using the V_n tiles. Before giving these problems and their solutions, we need to go over some terminology.

A *tile* is a topological disk in the Euclidean plane, and a *tiling* of the plane is a countable collection of tiles whose interiors are pairwise disjoint and whose union is the plane. A tiling in which all of the tiles are congruent to one another is a *monohedral* tiling. Two tiles of a tiling are *neighbors* if their intersection is nonempty, and the *neighborhood* $\mathcal{N}(T)$ of a tile T in a tiling is the collection of neighbors of T (including T itself). For example, the neighborhood of a square in a standard edge-to-edge tiling by squares is that square and the eight squares surrounding it. It is possible that the union of the tiles in $\mathcal{N}(T)$ can fail to be simply connected, though it is difficult to imagine this occurring in a monohedral tiling. The *patch* $\mathcal{A}(T)$ generated by a tile T in a tiling is $\mathcal{N}(T)$ together with the minimal collection of tiles necessary to form a simply connected union.

With this terminology, we are ready to state the problems.

1. Decide whether for some (or for each) $r \geq 3$ there exists a tile T having the r-enclosure property and which admits a tiling of the plane (see Grünbaum and Shephard [**3**, p. 129] and [**4**]).
2. Show that if $\mathcal{T}$ is a monohedral tiling and T is a tile of $\mathcal{T}$, then $\mathcal{N}(T) = \mathcal{A}(T)$ (see Grünbaum and Shephard [**4**, p. 26]).

Solution to Problem 1.   Problem 1 is readily solved by the V_n tiles. In particular, if the beak angle of a V_n tile is a factor of π, then such V_n tiles admit spiral tilings of the plane. As an added bonus, V_n can admit spiral tilings *while demonstrating the r-enclosure property*! The V_n tile also admits periodic tilings of the plane similar to that in Figure 2 while demonstrating the r-enclosure property.
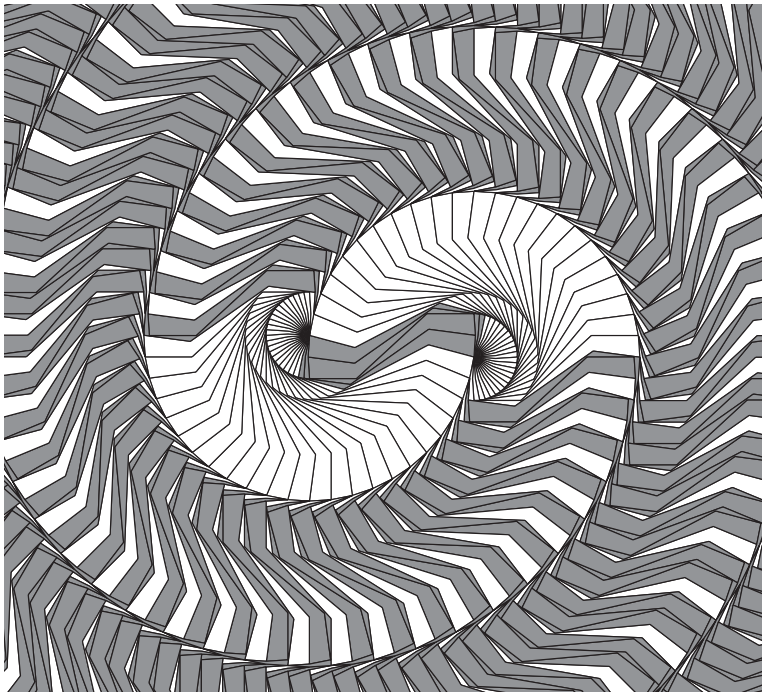

Figure 8   Spiral tiling admitted by the V_2 tile while displaying 4-enclosure property at the center.

Construction of the double spiral tiling by V_N

1. Begin at the center of the tiling by enclosing the desired amount of prototiles within two tiles.
2. To make the first layer of the spiral, fill in the two vertices where the enclosing tiles meet with beaks of tiles.
3. To make the second layer, move around the first layer adding a beak to butt pair of tiles followed by a single tile oriented so the beak is against the first layer.
4. To make the third layer, continue spiraling by adding two beak to butt pairs, then a single tile oriented so the beak is against the first layer.
5. The next layers are built in this same fashion. If you are adding layer k, first add $k - 1$ beak to butt pairs and then the single tile. One can continue tiling outward in this fashion infinitely so that the entire plane is covered.

Solution to Problem 2.   We provide a counterexample to show that the assertion of Problem 2 is false: In Figure 9, let T be the tile shaded in black. We see that $\mathcal{N}(T) \neq \mathcal{A}(T)$ because the tile shaded black is a member of $\mathcal{A}(T)$, but is not a member of $\mathcal{N}(T)$ (the tiles shaded gray are in both $\mathcal{N}(T)$ and $\mathcal{A}(T)$).
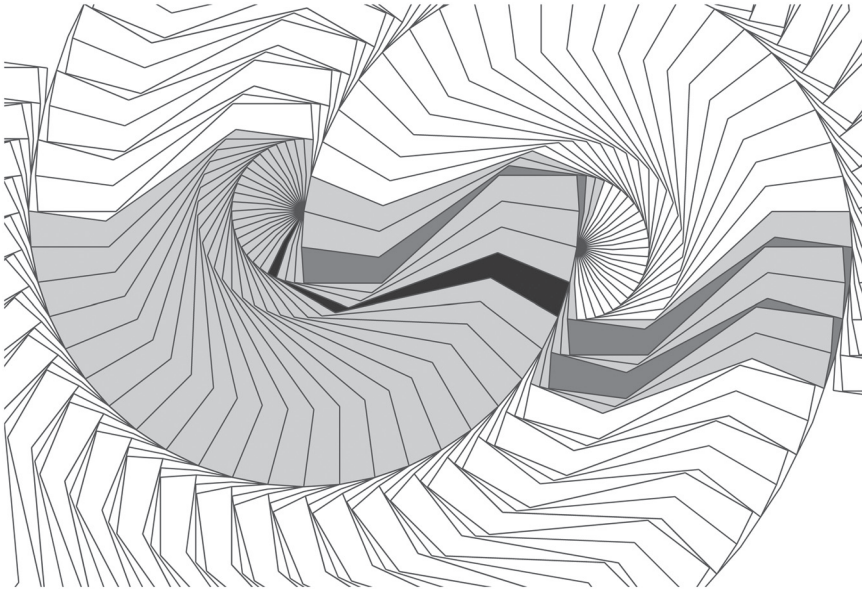

Figure 9 $\mathcal{A}(T) \neq \mathcal{N}(T)$. T is the black tile. The darker gray tiles are members of $\mathcal{A}(T)$ but not of $\mathcal{N}(T)$. The tiles shaded light gray are in both $\mathcal{N}(T)$ and $\mathcal{A}(T)$.

Discussion and open questions

We do not know what kind of witchcraft H. Voderberg wielded to invent his self-surrounding shape, but we do wonder if any tile with the r-enclosure property can be in any substantial way different from the Voderberg tile. Its construction allows for "self-spooning" (because of the rotated side), and we observe that this is critical to the behavior of the tiles.

We close with some open questions:

1. Is there a three-dimensional analog of the Voderberg tile? That is, is there a 3-D tile such that two copies can completely enclose some number of copies without gaps?

2. Is there an essentially different kind of 2-D tile that admits tilings of the plane while having the self-surrounding properties of the Voderberg tile?

3. Is there a 2-D tile such that two copies can *completely* surround (as an annulus) some number of copies while tiling the plane? In Mann [5], we see a modification of the Voderberg tile for which two copies can completely surround a number of copies, but that shape does not tile the plane.

4. In Grünbaum and Shephard [4, p. 129], we find a generalization of the r-enclosure property along with an open problem which we paraphrase here: A tile T has the (m, n)-enclosure property if two copies T_1 and T_2 of T can be arranged so that the complement of $T_1 \cup T_2$ has m bounded components, the closure of which is the union of n nonoverlapping copies of T. The open problem is: For each pair of positive integers m and n, find a tile with the (m, n)-enclosure property.

REFERENCES

[1] Gardner, M. (1977). Mathematical games. *Sci. Am.* 236(1): 110–121.
[2] Goldberg, M. (1955). Central tesselations. *Scr. Math.* 21: 253–260.
[3] Grünbaum, B., Shephard, G. C. (1987). *Tilings and Patterns*. New York: W. H. Freeman.
[4] Grünbaum, B., Shephard, G. C. (1998). Some problems on plane tilings. In: Klarner, D. A., ed. *Mathematical Recreations: A Collection in Honor of Martin Gardner*. Mineola: Dover, pp. 167–196.
[5] Mann, C. (2002). A tile with surround number 2. *Amer. Math. Monthly*. 109(4): 383–388.

Summary.   We present a generalization of the Voderberg tile, which, in addition to admitting periodic and nonperiodic spiral tilings of the plane, has the property that just two copies can surround 1 or 2 copies of the tile. We construct a generalization of this tile that admits periodic and nonperiodic spiral tilings of the plane while enjoying the property that any number of copies of the tile can be surrounded by just 2 copies. In doing so, we solve two open problems posed in the classic book *Tilings and Patterns* by Grünbaum and Shephard.

JADIE ADAMS earned her Bachelor of Science degree in math from Westminster College before going on to work in automatic speech recognition. She is currently working toward a Ph.D. in computing at the University of Utah, where she her research focuses on machine learning and statistical shape modeling for medical image analysis.

GABRIEL LOPEZ earned his Bachelor of Science in mathematics at California State University, San Bernardino. As of this publication, he is a graduate student in mathematics at the University of Colorado, Boulder. He hopes to use his training to not only go into higher education, but to work in outreach and create research opportunities for students who are interested in the mathematical sciences, in particular to those from historically underrepresented groups in the community.

CASEY MANN is a mathematics professor at the University of Washington Bothell. His research interests include tilings and knot theory. He thinks it is impactful to engage undergraduate students in the process of mathematical discovery, as exemplified in this article and the accomplishments of his coauthors!

NHI TRAN earned her Bachelor of Science in Mathematics from the University of Washington Bothell in 2017.

The Slippery Duck Theorem

JOCELYN R. BELL
Hobart and William Smith Colleges
Geneva, NY 14456
bell@hws.edu

FRANK WATTENBERG
United States Military Academy
West Point, NY 10996
frank.wattenberg@westpoint.edu

Imagine that a dog swims directly toward a swimming duck. The path taken by the dog during this chase is an example of a pursuit curve. Pursuit curves have long been a subject of mathematical inquiry.

An implicit pursuit curve was described in *The Man in the Moone*, a well-known story written by bishop Frances Godwin, published in 1638. In this story, a wedge of swans carries an astronaut from the Earth to the Moon. The swans always fly directly toward the Moon as the Moon orbits the Earth, so that the path of the astronaut is not a straight line. An analysis of this problem was given by Simoson [9].

In 1920, in the 27th volume of the *American Mathematical Monthly*, the mathematician A. S. Hathaway posed Problem 2801:

> A dog at the center of a circular pond makes straight for a duck which is swimming along the edge of the pond. If the rate of swimming of the dog is to the rate of swimming of the duck as n:1, determine the equation of the curve of pursuit and the distance the dog swims to catch the duck [4].

The next year, in the 28th volume of the *Monthly*, Hathaway published an analysis of his own problem [5], and an approximate solution for the equation of the pursuit curve was provided by F. V. Morley [6]. However, it turns out that finding the exact solution is impossible [7]. Today, computer algebra systems make short work of finding and plotting approximate solutions. In fact, experimenting with example pursuit scenarios is part of the fun.

This article offers some new and surprisingly general results pertaining to this classical pursuit problem. We will show that when a duck travels along a nice enough closed curve, if the dog is swimming more slowly than the duck then there is a unique limit cycle for the dog. In other words, the dog eventually settles into a periodic path; see Figure 1. Our result is an application of the Brouwer fixed point theorem.

Assumptions

Our dogs and ducks will be swimming in $\mathbb{R}^n$, where n is a positive integer, although all of our examples will be in two-dimensional space. We assume the duck is traveling with constant speed 1 along a closed curve, and the dog is pursuing with constant speed k, where $k < 1$. We assume that the duck's path is fixed in advance, and while being pursued the duck is not allowed to stray from its path. If the duck were allowed to evade the dog, a pursuit and evasion game would result, yielding a very different
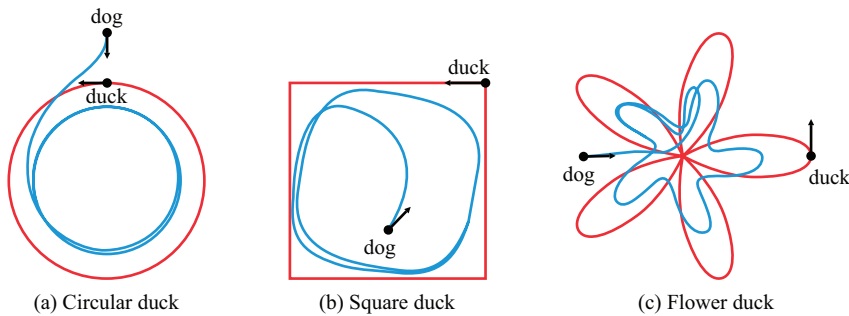
(a) Circular duck   (b) Square duck   (c) Flower duck

Figure 1   Dogs pursing ducks.

solution. The duck's path will be denoted by $x(t)$ and a dog's path by $y(t)$, where t is a parameter which we may think of as representing time.

The duck's path $x(t)$ is assumed to be continuous. We also assume that both the forward and backward derivative

$$\lim_{h \to 0^+} \frac{x(t + h) - x(t)}{h} \quad \text{and} \quad \lim_{h \to 0^-} \frac{x(t + h) - x(t)}{h}$$

exist for each t. However, we do not assume these two limits are equal. This allows us to consider, for example, a duck moving along a square path with sharp corners.

The path of the dog $y(t)$ is a solution to the initial value problem

$$y'(t) = \frac{k(x(t) - y(t))}{\|x(t) - y(t)\|} \quad \text{where} \quad y(0) = y_0. \tag{1}$$

In other words, the dog starts at the point y_0 and swims with speed k directly towards the duck. Dogs may begin pursuit anywhere: inside a duck's path, outside a duck's path, on a duck's path, or even on top of the duck—although this last case needs to be handled delicately.

If the dog catches the duck, that is, if $x(t) = y(t)$ for some time t, this will be called a *moment of capture*. Capture is indeed possible even though the dog is slower than the duck, since the duck does not evade the dog (see Figure 2). Immediately after capture, the duck, since it is swimming faster, pulls away from the dog and the dog resumes the chase. That is to say, we assume a slippery duck: in fact, the duck is so slippery that capture by a dog does not change the duck's path.

Examples

In Figure 1, three examples of dog and duck pursuit problems appear. At the far left, in Figure 1a, we begin with Hathaway's classical problem, with a twist. A duck with speed 1 starts swimming counterclockwise at the top of the circle, while the dog with speed 3/4 begins pursuit outside the duck's path, instead of at the center as specified in Hathaway's problem 2801. It is well-known that if k is the speed of the dog and $k < 1$ then with most starting positions the dog approaches a circular path of radius k [7]. This circular path is an example of a *cycle*: a path $y(t)$ of the dog which, after chasing the duck once along its closed path, returns to its starting position. That is, a cycle $y(t)$ is a solution to equation (1) satisfying $y(0) = y(p)$ where p is the period of the duck.

In Figure 1b, we see a dog chasing a duck whose path is a square. The dog starts inside the duck's path in this example, and eventually settles into a square-like path
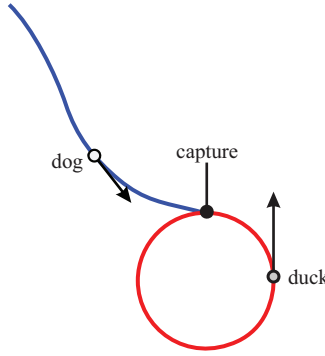
Figure 2   Capture!

with rounded corners. In Figure 1c, the duck is traveling along a flower shaped path and the dog eventually travels along a pinwheel-shaped curve. The presence of these three cycles in the three examples in Figure 1 is not exceptional: the goal of this article is to show that this phenomenon will always occur. In fact, we will show that these cycles are *limit cycles* in the usual sense (e.g., see chapter 16 of [**2**]).

When the dog is slower than the duck, the most common scenario is the one shown in Figure 1. That is, the dog eventually settles into a limit cycle without capturing the duck. It is, however, possible for the dog to catch the duck. If the slower dog is to catch the faster duck, then the dog must catch the duck from the front; see Figure 2. In effect, the duck swims into the dog.

The key lemma

Much of what is to come relies on the following observation: if two different dogs are pursuing the same duck at the same speed, and neither dog has yet caught the duck, then the two dogs are getting closer to each other unless all three are traveling in a straight line.

Lemma 1 (The key observation). *Suppose two dogs $y_1(t)$ and $y_2(t)$ are chasing the duck $x(t)$, and that at time t_0 neither dog captures the duck. Let $g(t) = \left\| y_2(t) - y_1(t) \right\|$. Then $g'(t_0) \le 0$ and $g'(t_0) = 0$ if and only if the two dogs and the duck lie on a straight line and the duck is not between the dogs.*

Proof. Suppose $y_1(t_0) \ne x(t_0)$ and $y_2(t_0) \ne x(t_0)$ and that $y_1(t)$ and $y_2(t)$ satisfy equation (1) on some neighborhood of t_0. We will show in the section on existence that y_1 and y_2 exist and are differentiable. Furthermore, suppose that $y_1(t_0) \ne y_2(t_0)$. That is, the two dogs are in different locations at time t_0.

Without loss of generality, we may assume that the second dog is closer to the duck:

$$\left\| y_1(t_0) - x(t_0) \right\| \ge \left\| y_2(t_0) - x(t_0) \right\|.$$

Consider a plane containing the three points $y_1(t_0)$, $x(t_0)$ and $y_2(t_0)$ and choose a coordinate system on this plane so that $y_1(t_0)$ is at the origin; see Figure 3a.

Let

$$h(t) = (g(t))^2 = \left\| y_2(t) - y_1(t) \right\|^2.$$
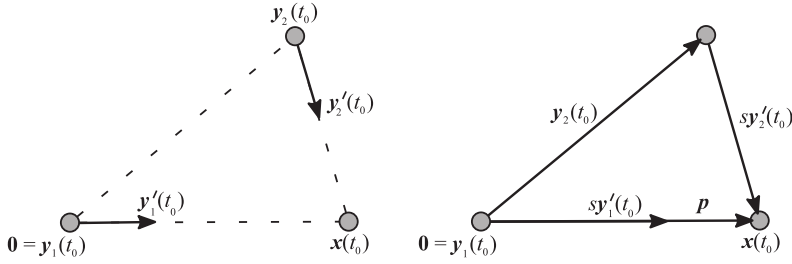
Then

$$h(t) = \left(y_2(t) - y_1(t) \right) \cdot \left(y_2(t) - y_1(t) \right),$$

Figure 3 A snapshot at time $t = t_0$.

$$h'(t) = 2\left(\boldsymbol{y}_2(t) - \boldsymbol{y}_1(t)\right) \cdot \left(\boldsymbol{y}_2'(t) - \boldsymbol{y}_1'(t)\right),$$

where $h'(t_0)$ is defined since $h(t)$ is differentiable at t_0. We will show that

$$\left(\boldsymbol{y}_2(t_0) - \boldsymbol{y}_1(t_0)\right) \cdot \left(\boldsymbol{y}_2'(t_0) - \boldsymbol{y}_1'(t_0)\right) \le 0.$$

Let s be the time required for the second dog to reach the duck if the duck were to remain fixed at $\boldsymbol{x}(t_0)$. That is,

$$s = \frac{\left\|\boldsymbol{x}(t_0) - \boldsymbol{y}_2(t_0)\right\|}{k} = \frac{\left\|\boldsymbol{x}(t_0) - \boldsymbol{y}_2(t_0)\right\|}{\left\|\boldsymbol{y}_2'(t_0)\right\|},$$

where k is the speed of the dog. Now, s is a scaling factor that scales $\boldsymbol{y}_2'(t_0)$ to a vector of length $\left\|\boldsymbol{x}(t_0) - \boldsymbol{y}_2(t_0)\right\|$. See Figure 3b. Then

$$\boldsymbol{y}_2(t_0) + s\boldsymbol{y}_2'(t_0) = \boldsymbol{x}(t_0).$$

Define

$$\boldsymbol{p} = \boldsymbol{x}(t_0) - s\boldsymbol{y}_1'(t_0),$$

as seen in Figure 3b. The three sides of the triangle in Figure 3b have lengths

$$\left\|\boldsymbol{y}_2(t_0)\right\|, \quad \left\|s\boldsymbol{y}_2'(t_0)\right\| = sk, \quad \text{and} \quad sk + \|\boldsymbol{p}\|.$$

Therefore, by the triangle inequality,

$$sk + \|\boldsymbol{p}\| \le \left\|\boldsymbol{y}_2(t_0)\right\| + sk$$

and therefore $\|\boldsymbol{p}\| \le \left\|\boldsymbol{y}_2(t_0)\right\|$. Now,

$$\boldsymbol{y}_2(t_0) + s\boldsymbol{y}_2'(t_0) = \boldsymbol{x}(t_0) = s\boldsymbol{y}_1'(t_0) + \boldsymbol{p},$$

so that

$$s\boldsymbol{y}_2'(t_0) - s\boldsymbol{y}_1'(t_0) = \boldsymbol{p} - \boldsymbol{y}_2(t_0).$$

By choice of coordinate system, $\boldsymbol{y}_1(t_0) = \boldsymbol{0}$ and therefore

$$\begin{aligned}
s\left(\boldsymbol{y}_2'(t_0) - \boldsymbol{y}_1'(t_0)\right) \cdot \left(\boldsymbol{y}_2(t_0) - \boldsymbol{y}_1(t_0)\right) &= \left(s\boldsymbol{y}_2'(t_0) - s\boldsymbol{y}_1'(t_0)\right) \cdot \boldsymbol{y}_2(t_0) \\
&= \left(\boldsymbol{p} - \boldsymbol{y}_2(t_0)\right) \cdot \boldsymbol{y}_2(t_0) \\
&= \boldsymbol{p} \cdot \boldsymbol{y}_2(t_0) - \boldsymbol{y}_2(t_0) \cdot \boldsymbol{y}_2(t_0) \\
&= \boldsymbol{p} \cdot \boldsymbol{y}_2(t_0) - \left\|\boldsymbol{y}_2(t_0)\right\|^2.
\end{aligned}$$
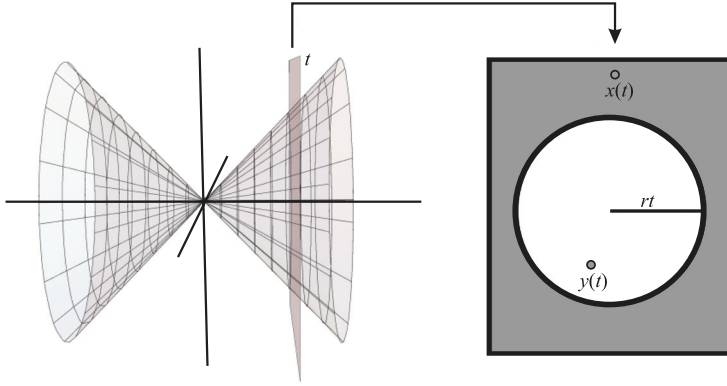
Figure 4 Capture at the origin: The dog is trapped inside the cone.

Now,

$$\boldsymbol{p} \cdot \boldsymbol{y}_2(t_0) = \|\boldsymbol{p}\| \, \|\boldsymbol{y}_2(t_0)\| \cos \theta,$$

where θ is the angle between $\boldsymbol{p}$ and $\boldsymbol{y}_2(t_0)$. Since $\|\boldsymbol{p}\| \leq \|\boldsymbol{y}_2(t_0)\|$,

$$\begin{aligned}
\boldsymbol{p} \cdot \boldsymbol{y}_2(t_0) &= \|\boldsymbol{p}\| \, \|\boldsymbol{y}_2(t_0)\| \cos \theta \\
&\leq \|\boldsymbol{y}_2(t_0)\| \, \|\boldsymbol{y}_2(t_0)\| \cos \theta \\
&\leq \|\boldsymbol{y}_2(t_0)\|^2,
\end{aligned}$$

and equality can only happen when $\cos \theta = 1$. That is, when the two dogs and the duck lie on a straight line and the duck is not between the two dogs. Otherwise, $\boldsymbol{p} \cdot \boldsymbol{y}_2(t_0) - \|\boldsymbol{y}_2(t_0)\|^2 < 0$ and so the proof is finished. ∎

In the event of capture, there is an interval of time during which no other captures by the same dog occur. This follows from the next lemma, which is an application of Taylor's theorem. In essence, in a neighborhood of capture, there is a cone which bounds the dog away from the duck. See Figure 4: the dog and duck are swimming in two dimensions, so the graphs of their position functions are in three dimensions. The dog is trapped inside the cone, while the duck is outside.

Lemma 2. *Suppose the dog $\boldsymbol{y}(t)$ captures the duck $\boldsymbol{x}(t)$ at the origin at time $t = 0$. Then there is a $\delta > 0$ and a constant $r < 1$ so that for all $t \neq 0$ with $-\delta \leq t \leq \delta$,*

$$\|\boldsymbol{x}(t)\| > r\,|t| > k\,|t| \geq \|\boldsymbol{y}(t)\|.$$

Proof. We apply Taylor's theorem. First, consider the case $t > 0$. Since by our assumptions $\lim_{h \to 0^+}(\boldsymbol{x}(h) - \boldsymbol{x}(0))/(h)$ exists, and since $\boldsymbol{x}(0) = \boldsymbol{0}$, we can define $\boldsymbol{q}(t)$ as follows:

$$\boldsymbol{q}(t) = \begin{cases} t \lim_{h \to 0^+} \frac{\boldsymbol{x}(h)}{h} & \text{if } t \leq 0, \\ \boldsymbol{x}(t) & \text{if } t > 0. \end{cases}$$

Then $\boldsymbol{q}(t)$ is differentiable at 0 with $\boldsymbol{q}'(0) = \lim_{h \to 0^+} \boldsymbol{x}(h)/h$, and $\boldsymbol{q}(t)$ is defined on an open interval containing 0. Note that $\|\boldsymbol{q}'(0)\| = 1$. By Taylor's theorem, for $t > 0$

$$\boldsymbol{x}(t) = \boldsymbol{x}(0) + \boldsymbol{q}'(0)t + \boldsymbol{\epsilon}(t)t = \boldsymbol{q}'(0)\,t + \boldsymbol{\epsilon}(t)\,t,$$

where $\boldsymbol{\epsilon}(t)$ is a vector-valued function which satisfies $\lim_{t \to 0} \boldsymbol{\epsilon}(t) = 0$.

Since $\lim_{t \to 0} \boldsymbol{\epsilon}(t) = 0$, we may choose δ so that if $0 < |t| \le \delta$ then $\|\boldsymbol{\epsilon}(t)\| < (1 - k)/2$, where k is the speed of the dog.

For the moment, assume that there is a solution $\boldsymbol{y}(t)$ to equation (1) if the dog begins his chase on top of the duck (that is, when the initial condition is $\boldsymbol{y}(0) = \boldsymbol{x}(0)$). We will show that this is the case in Theorem 2. The dog is moving at speed k, so for $t > 0$ we have $\|\boldsymbol{y}(t)\| \le kt$.

Now, for the duck, if $t > 0$, then we have

$$
\begin{aligned}
\|\boldsymbol{x}(t)\| &= \left\| \boldsymbol{q}'(0)t + \boldsymbol{\epsilon}(t)t \right\| \\
&= t \left\| \boldsymbol{q}'(0) + \boldsymbol{\epsilon}(t) \right\| \\
&\ge t \left(\left\| \boldsymbol{q}'(0) \right\| - \|\boldsymbol{\epsilon}(t)\| \right) \\
&= t \left(1 - \|\boldsymbol{\epsilon}(t)\| \right)
\end{aligned}
$$

so that if $0 < t \le \delta$ then

$$
\|\boldsymbol{x}(t)\| > \frac{(1 + k)t}{2}.
$$

Since $k < 1$, we have $k < (1 + k)/2$. Let $r = (1 + k)/2$. Then, $r < 1$ and if $0 < t \le \delta$, then

$$
\|\boldsymbol{x}(t)\| > \frac{(1 + k)t}{2} = rt > kt \ge \|\boldsymbol{y}(t)\|.
$$

If $t < 0$, define

$$
\boldsymbol{q}(t) = \begin{cases} t \lim_{h \to 0^-} \frac{\boldsymbol{x}(h)}{h} & \text{if } t \ge 0 \\ \boldsymbol{x}(t) & \text{if } t < 0 \end{cases}
$$

and proceed as in the previous paragraph.   ■

Existence of solutions

When faced with an initial value problem, the two most important questions are "Is there a solution?" and "If so, is the solution unique?" There are several so-called existence theorems available, possibly the most well-known of which is Peano's theorem [**8**], which appears in some form in nearly every introductory differential equations text. Peano's existence theorem states that if $\boldsymbol{g} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous on some open subset of its domain which includes $(t_0, \boldsymbol{y}_0)$, then the initial value problem

$$
\boldsymbol{y}'(t) = \boldsymbol{g}(t, \boldsymbol{y}(t)) \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0
$$

has a solution on an interval I of $\mathbb{R}$ containing t_0. However, this solution need not be unique. By a solution, we mean a differentiable vector-valued function $\boldsymbol{z} : I \to \mathbb{R}^n$ satisfying that, for all $t \in I$,

$$
\boldsymbol{z}'(t) = \boldsymbol{g}(t, \boldsymbol{z}(t)).
$$

In our dog and duck chase, when the dog is not on top of the duck, $\boldsymbol{f} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$
\boldsymbol{f}(t, \boldsymbol{u}) = \frac{k}{\|\boldsymbol{x}(t) - \boldsymbol{u}\|} (\boldsymbol{x}(t) - \boldsymbol{u}). \tag{2}
$$

(a) Neighborhood away from the duck   (b) Neighborhood surrounding the duck
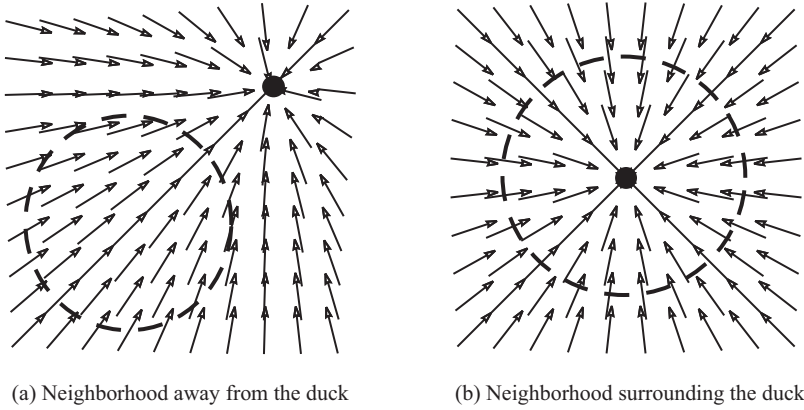
Figure 5 The vector field near the duck at a moment in time.

For each time t, if $\boldsymbol{u} \neq \boldsymbol{x}(t)$, $\boldsymbol{f}(t, \boldsymbol{u})$ can be pictured as a vector at $\boldsymbol{u}$ of length k pointed towards $\boldsymbol{x}(t)$. See Figure 5.

If $\boldsymbol{y}_0 \neq \boldsymbol{x}(t_0)$, then $\boldsymbol{f}$ is defined and continuous on some neighborhood of $(t_0, \boldsymbol{y}_0)$; see Figure 5a. By Peano's existence theorem, the initial value problem

$$\boldsymbol{y}'(t) = \boldsymbol{f}(t, \boldsymbol{y}(t)) \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0 \tag{3}$$

has a solution on an interval of the form $[t_0, t_0 + \delta)$, where no captures occur in this interval. It is easy to see that the key observation, Lemma 1, implies that such a solution is unique, for if there were two, they would agree at time t_0 and would thereafter be the same.

There is a small bump in the road on our way to the main theorem. The right hand side of equation (2) is not defined when $\boldsymbol{u} = \boldsymbol{x}(t)$, and there is no way to define $\boldsymbol{f}$ at these points which makes $\boldsymbol{f}$ continuous. This can be visualized by looking at Figure 5b. As soon as the duck moves an inch, some of the vectors will have to "flip" their direction.

We will address this difficulty by an application of Carathéodory's existence theorem. This theorem is slightly more advanced than Peano's existence theorem; for further reading see [2, 3]. We do not need the full power of this theorem; in fact, only the following stripped down version is necessary.

Theorem 1 (A version of Carathéodory's existence theorem)**.** *Suppose* $\boldsymbol{f}(t, \boldsymbol{u}) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ *satisfies that, for* $t_0 \leq t \leq t_0 + a$ *and* $\|\boldsymbol{u} - \boldsymbol{u}_0\| \leq b$, $\boldsymbol{f}(t, \boldsymbol{u})$ *is bounded, continuous in* $\boldsymbol{u}$ *for each fixed* t, *and continuous in* t *almost everywhere for every fixed* $\boldsymbol{u}$. *Then on a closed interval* $[t_0, t_0 + d]$, *where* $d > 0$, *there exists a solution of the initial value problem*

$$\boldsymbol{u}'(t) = \boldsymbol{f}(t, \boldsymbol{u}) \qquad \boldsymbol{u}(t_0) = \boldsymbol{u}_0$$

Our task is to show that a solution to the initial value problem (equation (3)) exists at capture. If the dog starts on top of the duck, the function $\boldsymbol{f}$ in equation (2) does *not* satisfy the hypotheses of the Carathéodory theorem on any neighborhood of the required type. Theorem 1 cannot be applied directly, and so we concoct a function which *does* satisfy the hypotheses of Theorem 1, and use this new function to find our solution. The cone from Lemma 2 will come into play.
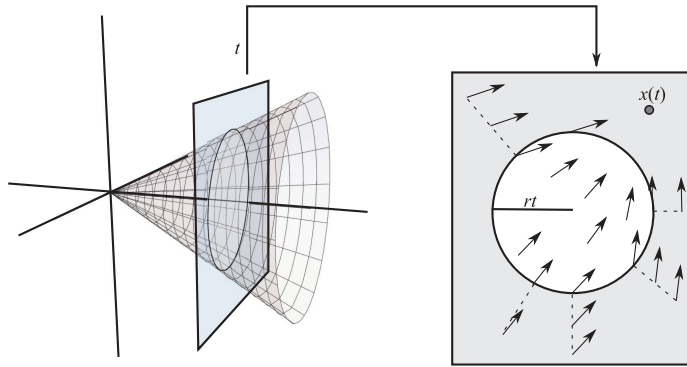
Figure 6　The vector field $v(t, u)$ at time t.

Theorem 2. *Suppose the dog $y(t)$ captures the duck at the origin at time $t = 0$. Then there is a solution to equation* (3) *with initial condition $y(0) = \mathbf{0}$ on an interval of the form* $[0, \alpha]$.

Proof. The idea of the proof is to define a new function $v(t, u)$ which agrees with

$$f(t, u) = \frac{k}{\|x(t) - u\|}(x(t) - u)$$

by the time the dog could possibly reach u. The function v will be bounded, continuous in t almost everywhere for each fixed u, and continuous in u for each fixed t.

The function $f(t, u)$ is continuous everywhere it is defined, as it is a rational expression composed of continuous functions. Since $\|f(t, u)\| = k$ whenever $f(t, u)$ is defined, $f(t, u)$ is bounded.

Choose δ and r as in the proof of Lemma 2. Let W be a closed ball centered at $x(0) = \mathbf{0}$ of radius

$$\|x(\delta) - x(0)\| = \|x(\delta)\|.$$

Then for $0 < t \leq \delta$, we have

$$\|x(t)\| > rt > kt,$$

where k is the speed of the dog. For $u \in W$, let $v(0, u) = \mathbf{0}$. Then $v(0, u)$ is continuous in u.

For $u \in W$ and $0 < t \leq \delta$, define the following vector-valued function:

$$v(t, u) = \begin{cases} f\left(t, \frac{(rt)u}{\|u\|}\right) & \text{if } rt < \|u\| \leq \|x(\delta)\|, \\ f(t, u) & \text{if } \|u\| \leq rt. \end{cases} \tag{4}$$

This function $v(t, u)$ is defined to be $f(t, u)$ inside the cone of Lemma 2 and is extended outside the cone in such a way that the extension agrees on the boundary of the cone; see Figure 6. If (t, u) is a point outside of the cone whose radial projection onto the cone is (t, w), then $v(t, u) = f(t, w)$.

Now, $v(t, u)$ is bounded since $f(t, u)$ is bounded. In fact, $\|v(t, u)\| = k$. We show that $v(t, u)$ is defined for all (t, u) with $0 < t \leq \delta$ and $u \in W$. Fix t with $0 < t \leq \delta$. If $\|u\| \leq rt$, then since $rt < \|x(t)\|$, we have that $\|u\| < \|x(t)\|$. Thus, $u \neq x(t)$ and $f(t, u)$ is defined.

If

$$rt < \|\boldsymbol{u}\| \leq \|\boldsymbol{x}(\delta)\|,$$

then

$$\left\| \frac{(rt)\boldsymbol{u}}{\|\boldsymbol{u}\|} \right\| = rt < \|\boldsymbol{x}(t)\|.$$

Thus, $(rt)\boldsymbol{u}/\|\boldsymbol{u}\| \neq \boldsymbol{x}(t)$ and so $\boldsymbol{f}(t, \boldsymbol{u})$ is again defined.

For each fixed t, with $0 < t \leq \delta$, the function $\boldsymbol{v}(t, \boldsymbol{u})$ is continuous in $\boldsymbol{u}$. This reason is that if $\|\boldsymbol{u}\| = rt$, then $(rt)\boldsymbol{u}/\|\boldsymbol{u}\| = \boldsymbol{u}$, so that

$$\boldsymbol{f}(t, \boldsymbol{u}) = \boldsymbol{f}\left(t, \frac{(rt)\boldsymbol{u}}{\|\boldsymbol{u}\|}\right).$$

For each fixed $\boldsymbol{u}$ with $\boldsymbol{u} \in W$, the function $\boldsymbol{v}(t, \boldsymbol{u})$ has only one possible discontinuity at $t = 0$, and so $\boldsymbol{v}(t, \boldsymbol{u})$ is continuous in t almost everywhere for every fixed $\boldsymbol{u}$.

Therefore, by Theorem 1, the differential equation

$$\boldsymbol{y}'(t) = \boldsymbol{v}(t, \boldsymbol{y}(t)) \qquad \boldsymbol{y}(0) = \boldsymbol{x}(0) = \boldsymbol{0}$$

has a solution on an interval of the form $[0, \alpha]$ where $0 < \alpha$. We can assume $\alpha \leq \delta$.

Now, $\|\boldsymbol{y}'(t)\| = k$ since $\|\boldsymbol{v}(t, \boldsymbol{u})\| = k$ for all $\boldsymbol{u}$ and all t. So at time t, the most $\|\boldsymbol{y}(t)\|$ can be is kt. Thus, for each t, with $0 < t \leq \delta$, we have that

$$\|\boldsymbol{y}(t)\| \leq kt < rt,$$

and therefore

$$\boldsymbol{v}(t, \boldsymbol{y}(t)) = \boldsymbol{f}(t, \boldsymbol{y}(t)).$$

Thus, any solution on an interval of the form $[0, \alpha]$ to the differential equation given by $\boldsymbol{y}'(t) = \boldsymbol{v}(t, \boldsymbol{y}(t))$ with initial condition $\boldsymbol{x}(0) = \boldsymbol{y}(0) = \boldsymbol{0}$ also satisfies the equation $\boldsymbol{y}'(t) = \boldsymbol{f}(t, \boldsymbol{y}(t))$ and the condition $\boldsymbol{x}(0) = \boldsymbol{y}(0) = \boldsymbol{0}$ on an interval of the form $[0, \alpha]$. ∎

We can demonstrate the uniqueness of this solution by applying the key observation and the mean value theorem.

Lemma 3. *The solution in Theorem 2 is unique.*

Proof. Suppose $\boldsymbol{y}_1(t)$ and $\boldsymbol{y}_2(t)$ are distinct solutions on an interval $[0, \alpha]$, where there are no other moments of capture during this interval. Since $\boldsymbol{y}_1 \neq \boldsymbol{y}_2$ there is some $t_i \in (0, \alpha]$ so that $\boldsymbol{y}_1(t_i) \neq \boldsymbol{y}_2(t_i)$. If at some time t_k with $0 < t_k < t_i$ we have $\boldsymbol{y}_1(t_k) = \boldsymbol{y}_2(t_k)$, then by Lemma 1, we have that $\boldsymbol{y}_1(t_i) = \boldsymbol{y}_2(t_i)$. So for all t with $0 < t < t_i$ it is the case that $\boldsymbol{y}_1(t) \neq \boldsymbol{y}_2(t)$. Then

$$g(t) = \|\boldsymbol{y}_1(t) - \boldsymbol{y}_2(t)\|$$

is continuous on the interval $[0, t_i]$. Since $\boldsymbol{y}_1(t)$ and $\boldsymbol{y}_2(t)$ are differentiable on $(0, t_i)$, and since the Euclidean norm is differentiable everywhere except the origin, we have that $g(t)$ is differentiable on the interval $(0, t_i)$.

By the mean value theorem, there is t_j with $0 < t_j < t_i$ satisfying

$$g'(t_j) = \frac{g(t_i) - g(0)}{t_i}.$$

Since $\mathbf{y}_1(0) = \mathbf{y}_2(0)$, we have $g(0) = 0$ and since $\mathbf{y}_1(t_i) \neq \mathbf{y}_2(t_i)$, $g(t_i) > 0$. Therefore, $g'(t_j) > 0$. This contradicts Lemma 1. ∎

From here it is a short argument to show that on any time interval and for any initial value, equation (1) has a unique, continuous solution. The idea is a natural one: we "paste" solutions together. It is clear that a solution can be uniquely extended up to, but not including, a point of capture, and if t is not a moment of capture, then there is an interval of the form $(t - \delta, t + \delta)$ during which no other captures occur. By this observation and Lemma 2, on a closed interval $[t_1, t_2]$, capture may occur at only finitely many moments. Furthermore, between moments of capture, the path of the dog is differentiable, as it is a solution to an initial value problem.

Lemma 4. *For any $t > 0$ and any $\mathbf{y}_0$ in $\mathbb{R}^n$, the initial value problem*

$$\mathbf{y}'(t) = \frac{k}{\|\mathbf{x}(t) - \mathbf{y}(t)\|}(\mathbf{x}(t) - \mathbf{y}(t)) \qquad \mathbf{y}(0) = \mathbf{y}_0$$

has a unique continuous solution on the interval $[0, t]$.

Proof. By an easy modification of Theorem 2, there is a unique solution $\mathbf{y}_1(t)$ to the initial value problem when $\mathbf{y}(0) = \mathbf{y}_0$ on an interval of the form $[0, t_1)$, where t_1 is a moment of capture. If $t_1 > t$, then we are done.

If not, there is a unique solution $\mathbf{y}_2(t)$ to the initial value problem when

$$\mathbf{y}(t_1) = \mathbf{y}_1(t_1) = \mathbf{x}(t_1)$$

on an interval of the form $[t_1, t_2)$ where t_2 is a moment of capture. Since $\mathbf{y}_1(t_1) = \mathbf{y}_2(t_1)$, the function

$$\mathbf{y} = \begin{cases} \mathbf{y}_1(t) & \text{if } 0 \leq t < t_1 \\ \mathbf{y}_2(t) & \text{if } t_1 \leq t < t_2 \end{cases}$$

is continuous on $[0, t_2)$. Continuing, this process stops at a finite stage, as there are only finitely many moments of capture in the interval $[0, t]$. ∎

The main result

We begin by observing that the key idea, dogs chasing the same duck cannot be moving away from each other, is still true even if one or both of the dogs captures the duck.

Lemma 5. *If two dogs are chasing the duck, then the distance between the two dogs is not increasing.*

Proof. Suppose the two dogs are $\mathbf{y}_1(t)$ and $\mathbf{y}_2(t)$ and that $t_1 < t_2$. On the interval $[t_1, t_2]$, capture occurs only finitely many times. By Lemma 1 the function

$$\|\mathbf{y}_1(t) - \mathbf{y}_2(t)\|$$

is non-increasing on each of the sub-intervals between points of capture on $[t_1, t_2]$. Since $\mathbf{y}_1(t)$ and $\mathbf{y}_2(t)$ are continuous by Lemma 4,

$$\|\mathbf{y}_1(t) - \mathbf{y}_2(t)\|$$

is continuous. It follows that $\|\mathbf{y}_1(t) - \mathbf{y}_2(t)\|$ is non-increasing on $[t_1, t_2]$. ∎
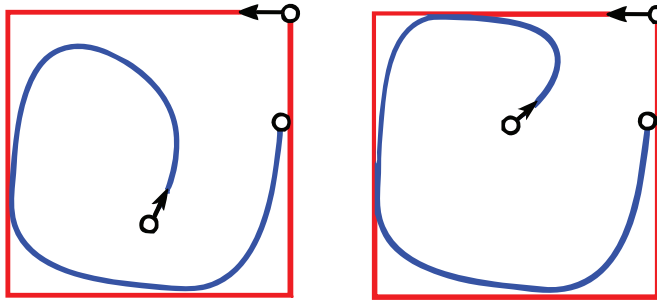
Figure 7 Two different dogs chasing a square duck.

Fixed point theorems play an important role in both theoretical and applied mathematics. Fixed point theorems are used, for example, to find equilibrium points. They are also used, as we will see, to argue that limit cycles must exist. Fixed points and fixed point theorems—for example, the search for a function f that is its own derivative—occur throughout mathematics.

The n-dimensional Brouwer fixed point theorem [1] states that any continuous map f from a closed ball B in $\mathbb{R}^n$ to itself must have a fixed point. That is, a point x in B satisfying $f(x) = x$. The one-dimensional Brouwer fixed point theorem has a straightforward proof, but even the two-dimensional Brouwer fixed point theorem is surprisingly difficult. The Brouwer fixed point theorem, as with many fixed point theorems, applies only to continuous functions defined on bounded, simply connected domains: that is, domains without "holes."

Theorem 3 (The slippery duck theorem)**.** *Suppose the duck's path is a continuous closed curve so that the forward and backward derivatives exist. If the dog is swimming slower than the duck then there is a cycle for the dog.*

Proof. Let $B \subseteq \mathbb{R}^n$ be a closed ball of large enough radius so that the curve of the duck's path is contained in the interior of B. For each $u \in B$, let $y_u(t)$ be the unique continuous solution to the initial value problem

$$y_u'(t) = \frac{k}{\|x(t) - y_u(t)\|}(x(t) - y_u(t)) \qquad y_u(0) = u$$

on the interval $[0, p]$ where p is the time it takes the duck to travel once along its path.

Since B is convex and $x(t)$ is in the interior of B, $y_u(t) \in B$ for each t.

Define a map $w : B \to B$ as follows: for each $u \in B$, let $w(u) = y_u(p)$. That is, $w(u)$ is defined by placing a dog at u at time 0 and seeing where the dog ends up after chasing the duck once around its path. See Figure 7. By Lemma 5

$$\|w(u) - w(v)\| \leq \|u - v\|$$

and so this map w is continuous.

By the Brouwer fixed point theorem, there is some $v \in B$ for which $w(v) = v$; in other words, $y_v(p) = y_v(0)$, and the dog which started at v has returned to v after chasing the duck once around its path. Then $y_v(t)$ is a cycle for the dog with period p. ∎

Next, we show that this cycle is unique and that it is a limit cycle: no matter where a dog begins the chase, that dog eventually becomes arbitrarily close to a dog traveling along this cycle (see Coddington [2]). Our slippery duck theorem is unusually general,
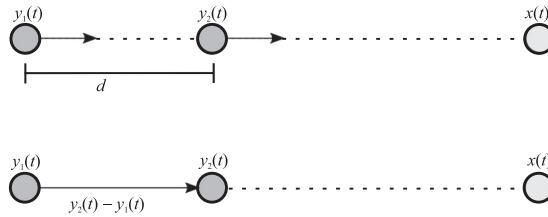
Figure 8 Two dogs and the duck in a line.

since it applies in any dimension, whereas most results on limit cycles apply only to motion in the plane.

Theorem 4. *The cycle of Theorem 3 is both unique and a limit cycle.*

Proof. Suppose there are two different cycles $y_1(t)$ and $y_2(t)$. By Lemma 5, $\left\| y_1(t) - y_2(t) \right\|$ is non-increasing and can only be constant when the dogs are following the duck in a straight line. Therefore, for all $t \in [0, p]$, where p is the period of the duck, if

$$d = \left\| y_1(0) - y_2(0) \right\|$$

we have

$$\left\| y_1(t) - y_2(t) \right\| = d.$$

That is, the dogs are always distance d apart. Furthermore, for each t, the points $x(t)$, $y_1(t)$, and $y_2(t)$ are collinear, and we can assume the second dog is closer to the duck than the first. Since the three points are always on a line and the duck's path is continuous, $y_2(t)$ is closer to $x(t)$ than $y_1(t)$ for each t. Then $y_1'(t) = y_2'(t)$, and

$$y_1'(t) = y_2'(t) = \frac{k\left(y_2(t) - y_1(t)\right)}{\left\| y_2(t) - y_1(t) \right\|} = \frac{k\left(y_2(t) - y_1(t)\right)}{d}$$

since the velocity vectors of the dogs point in the same direction as the line segment joining them; see Figure 8. If t is not a moment of capture for either dog, then $y_1'(t)$ and $y_2'(t)$ are differentiable. In this case

$$y_1''(t) = y_2''(t) = \frac{k}{d}\left(y_2'(t) - y_1'(t)\right) = 0.$$

Therefore $y_1'(t) = y_2'(t)$ is constant, and both dogs are moving in a straight line between moments of capture.

Suppose just before capture at time t the dogs are moving along line L_1. At capture,

$$y_2(t) = x(t) \quad \text{and} \quad y_1(t) \neq x(t)$$

since the second dog is closer to the duck. If L_2 is another line through $x(t) = y_2(t)$ with $L_1 \neq L_2$ then there is an $r > 0$ so that the ball of radius r about $y_1(t)$ does not intersect L_2. Then, since y_1 is continuous, there is $\epsilon > 0$ such that $y_1(t + \epsilon)$ is within distance r of $y_1(t)$. Then $y_1(t + \epsilon)$ is not on the line L_2; see Figure 9. Therefore, after capture the dogs cannot be traveling along line L_2.

Therefore, the dogs are always moving in the same straight line. Moreover, the dogs cannot change direction, as the path of the duck is continuous. This contradicts that $y_1(t)$ and $y_2(t)$ are cycles, and uniqueness follows.
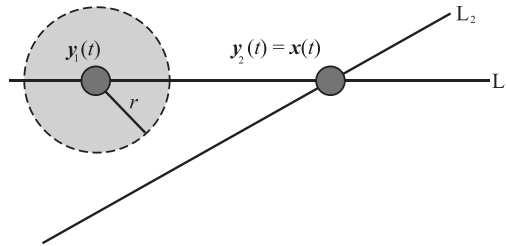
Figure 9 Two lines at capture.

The fact that the cycle is a limit cycle follows from the uniqueness of the cycle and Lemma 5. That is, if $v(t)$ is the unique cycle for the dog, then for any path $y(t)$ for the dog, we have

$$\lim_{t \to \infty} \| y(t) - v(t) \| = 0.$$

■

REFERENCES

[1] Brouwer, L. E. J. (1912). Über Abbildung von Mannigfaltigkeiten. *Math. Ann.* 71(1): 97–115. doi.org/10.1007/BF01456931

[2] Coddington, E. A., Levinson, N. (1955). *Theory of Ordinary Differential Equations*. New York: McGraw-Hill.

[3] Filippov, A. F. (2013). *Differential Equations with Discontinuous Righthand Sides: Control Systems*. Heidelberg: Springer.

[4] Hathaway, A. S. (1920). Problem 2801. *Amer. Math. Monthly*. 27(1): 31. doi.org/10.2307/2973244

[5] Hathaway, A. S. (1921). Solution to Problem 2801. *Amer. Math. Monthly*. 28(2): 93–97. doi.org/10.2307/2973052

[6] Morley, F. V. (1921). A curve of pursuit. *Amer. Math. Monthly*. 28(2): 54–61. doi.org/10.2307/2973034

[7] Nahin, P. J. (2012). *Chases and Escapes: The Mathematics of Pursuit and Evasion*. Princeton: Princeton University Press.

[8] Peano, G. (1890). Démonstration de l'intégrabilité des équations différentielles ordinaires. *Math. Ann.* 37(2): 182–228. doi.org/10.1007/BF01200235

[9] Simoson, A. J. (2007). Pursuit curves for the Man in the Moone. *College Math. J.* 38(5): 330–338. doi.org/10.1080/07468342.2007.11922257

Summary. Consider Hathaway's classical dog-and-duck problem: if a swimming dog heads directly toward a swimming duck, what is the path taken by the dog? Assuming a "slippery" duck is traveling along a smooth enough closed curve, using a combination of computer-based exploration and Brouwer's fixed point theorem we obtain a general result regarding the long term behavior of the dog. The proof of our slippery duck theorem relies mainly on familiar theorems from calculus. This topic lends itself well to undergraduate level student exploration.

JOCELYN R. BELL (MR Author ID: 1057578) received her Ph.D. in Mathematics in 2011 from the State University of New York, Buffalo. She has taught at the United States Military Academy, West Point and since 2016 has been at Hobart and William Smith Colleges.

FRANK WATTENBERG (MR Author ID: 237940) received his Ph.D. in Mathematics in 1968 from the University of Wisconsin, Madison. He has taught at a number of universities in the United States and Europe. Since 2001 he has been at the United States Military Academy, West Point and has been focusing on using mathematical modeling for public policy decisions and exploiting technology in undergraduate mathematics education.

Quadratics and the Floor Function

VINCENT J. MATSKO
Spring Hill, FL
www.vincematsko.com
www.cre8math.com
vince.matsko@gmail.com

My interest in the floor function blossomed while teaching functions and continuity in calculus. Textbooks tend to provide contrived examples of functions with jump discontinuities. Problems typically involve looking at graphs, but require no deeper analysis.

I began introducing the floor function in this context, as it allows for nontrivial examples which do require deeper analysis. For example, where is the expression

$$\frac{1}{\lfloor x \rfloor^2 - 3x + 2}$$

defined? Most students find the exceptions at $x = 1, 2$, but almost never find the exceptions at $x = 2/3, 11/3$. Moreover, I have not found a computer algebra system which is able to solve such problems in general.

We will examine quadratic equations where zero or more of the three occurrences of x are replaced by $\lfloor x \rfloor$. We look exclusively at real roots here, since defining the floor function uses an order relation on the reals which does not extend meaningfully to the complex numbers.

The results are intriguing. An equation may have no solutions at all. We will show by direct computation that if there are solutions and they are discrete, there may be arbitrarily many solutions. Moreover, if the solutions consist of a finite union of disjoint intervals (as illustrated in Figure 1), there may be arbitrarily many intervals, and further, the sum of their lengths is bounded.

In order to study these quadratic equations, we enumerate the following eight possibilities, where b and c are real numbers. We order them in this way to make the exposition easier to follow.

$$
\begin{array}{llll}
x^2 + bx + c = 0 & (1) & \lfloor x^2 \rfloor + bx + c = 0 & (5) \\
\lfloor x \rfloor^2 + b\lfloor x \rfloor + c = 0 & (2) & x\lfloor x \rfloor + b\lfloor x \rfloor + c = 0 & (6) \\
\lfloor x \rfloor^2 + bx + c = 0 & (3) & x^2 + b\lfloor x \rfloor + c = 0 & (7) \\
x\lfloor x \rfloor + bx + c = 0 & (4) & \lfloor x^2 \rfloor + b\lfloor x \rfloor + c = 0 & (8)
\end{array}
$$

We assume that $b \neq 0$, since these cases are easy to analyze. The two simplest equations to consider are (1) and (2). Of course (1) is the usual quadratic equation. Equation (2) has solutions if and only if the quadratic $y^2 + by + c = 0$ has an integer solution. If y is such a solution, then clearly $[y, y + 1)$ contains solutions to (2).

Equation (3)

To develop an intuition for studying these equations, we begin with (3), and recalling that $b \neq 0$, rewrite it as

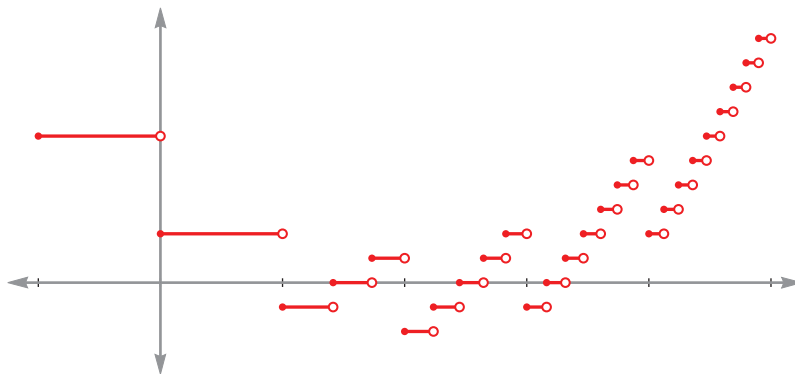$$-\frac{\lfloor x \rfloor^2 + c}{b} = x.$$

Figure 1   Graph of $y = \lfloor x^2 \rfloor - 4\lfloor x \rfloor + 2$, for $x \in [-1, 5]$.

Putting the equation in this form allows for a more straightforward graphical analysis.

To show that the number of solutions is unbounded in general, we seek parameterizations of b and c with a positive integer, n. There are many ways to accomplish this. One method involves the observation that the line $y = x$ is tangent to the graph of $y = (x^2 + n^2)/(2n)$ at (n, n), illustrated in Figure 2 with $n = 3$. As a result, solutions to $(\lfloor x \rfloor^2 + n^2)/(2n) = x$ will cluster around (n, n).
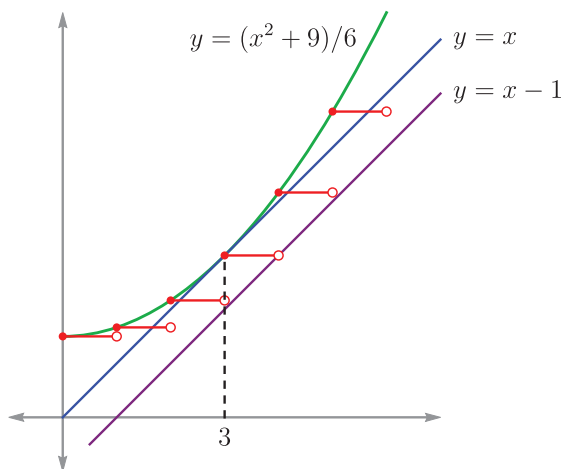
Figure 2   Equation (3) with $n = 3$: graphs of $y = (\lfloor x \rfloor^2 + 9)/6$ (horizontal segments), $y = (x^2 + 9)/6$, $y = x$, and $y = x - 1$.

This suggests using $b = -2n$ and $c = n^2$, so that we seek to find solutions to

$$\lfloor x \rfloor^2 - 2nx + n^2 = 0. \tag{9}$$

Clearly, $x = n$ is a solution. As we are clustering around $x = n$, we seek solutions of the form $n + k + \varepsilon$, where $k \in \mathbb{Z}$ and $\varepsilon \in [0, 1)$. Writing solutions in this form will make it easier both to evaluate the floor function and to count solutions.

Substituting $x = n + k + \varepsilon$ into (9) and rearranging terms gives

$$k^2 = 2n\varepsilon.$$

Solutions for $k \in \mathbb{Z}$ and $\varepsilon \in [0, 1)$ exist when $|k| < \left\lfloor \sqrt{2n} \right\rfloor$. Note that the inequality is strict, otherwise we would have $\varepsilon = 1$ when $n = 2m^2$ and m is a positive integer.

When $n \geq 1$, we have an odd number of solutions. By varying n, we may obtain any number of odd solutions greater than one. To see that any number of solutions $n \in \mathbb{N}$ is possible, we refer again to Figure 2.

As $x = 3$ is a solution, we see that $(3, 3)$ lies both on $y = (\lfloor x \rfloor^2 + 9)/6$ and $y = x$. Thus, we may shift the graph of $y = (\lfloor x \rfloor^2 + 9)/6$ down by a small amount (algebraically, by slightly decreasing 9) to generate an even number of solutions. Shifting this graph up 1 gives no solutions—although with a near miss at $(4, 3)$. Shifting the graph down slightly will then generate a single solution. Thus, any number of solutions is possible.

Equation (4)

We consider (4) next as the solution technique is similar to that used to solve (3). Again using $b = -2n$ and $c = n^2$, we have

$$x \lfloor x \rfloor - 2nx + n^2 = 0. \tag{10}$$

We substitute $x = n + k + \varepsilon$ as before (with $\varepsilon \in [0, 1)$), giving

$$\varepsilon = \frac{k^2}{n - k}. \tag{11}$$

Now $k < n$ gives $\varepsilon \geq 0$. To have $\varepsilon < 1$, we rewrite (11) as

$$k^2 + k - n < 0. \tag{12}$$

This requires

$$\frac{-1 - \sqrt{1 + 4n}}{2} < k < \frac{-1 + \sqrt{1 + 4n}}{2}.$$

A brief table of admissible values for k given a specific n is given in Table 1.

TABLE 1: Admissible values for k.

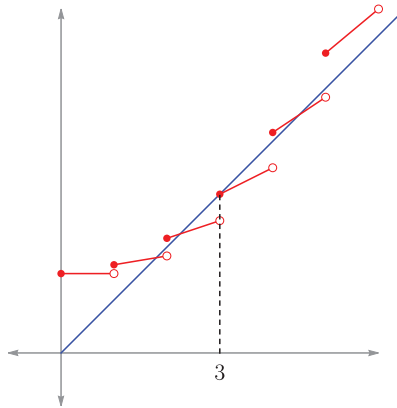| Range for n	Minimum k	Maximum k
$1, \dots, 2$	-1	0
$3, \dots, 6$	-2	1
$7, \dots, 12$	-3	2
$13, \dots, 20$	-4	3

Figure 3 Equation (4) with $n = 3$: solutions to $(x \lfloor x \rfloor + 9)/6 = x$.

It should be clear from (12) that jumps occur in the table when n is of the form $m(m + 1)$. Indeed, we see that there are $2(m + 1)$ admissible values for k when

$$m(m + 1) < n \le (m + 1)(m + 2).$$

Thus, any even number of solutions is possible. The case when $n = 3$ is illustrated in Figure 3, where (10) is rewritten (as will be the case from this point on) in the form

$$\frac{x \lfloor x \rfloor + n^2}{2n} = x.$$

Again, because both graphs pass through (n, n), it is clear we may obtain any odd number of solutions by shifting the graph of $(x \lfloor x \rfloor + n^2)/(2n)$ up by slightly increasing n^2, or no solutions at all by sufficiently shifting the graph up.

Equation (5)

Here, we take a slightly different approach, although we still look at (5) with $b = -2n$ and $c = n^2$ in the form

$$\frac{\lfloor x^2 \rfloor + n^2}{2n} = x. \tag{13}$$

Since $\lfloor x^2 \rfloor$ jumps when x^2 is an integer and $x = n$ is a solution to (13), we look for x in intervals such as

$$\left[\sqrt{n^2 + k}, \sqrt{n^2 + k + 1} \right).$$

In this case,

$$\lfloor x^2 \rfloor = n^2 + k,$$

and so we seek those $k \in \mathbb{Z}$ for which

$$\sqrt{n^2 + k} \le n + \frac{k}{2n} < \sqrt{n^2 + k + 1}.$$

This is easily seen to be equivalent to

$$0 \leq \frac{k^2}{4n^2} < 1,$$

giving $|k| < 2n$. This allows the number of solutions to be an arbitrary odd number congruent to 3 mod 4. The case $n = 2$ is illustrated in Figure 4.
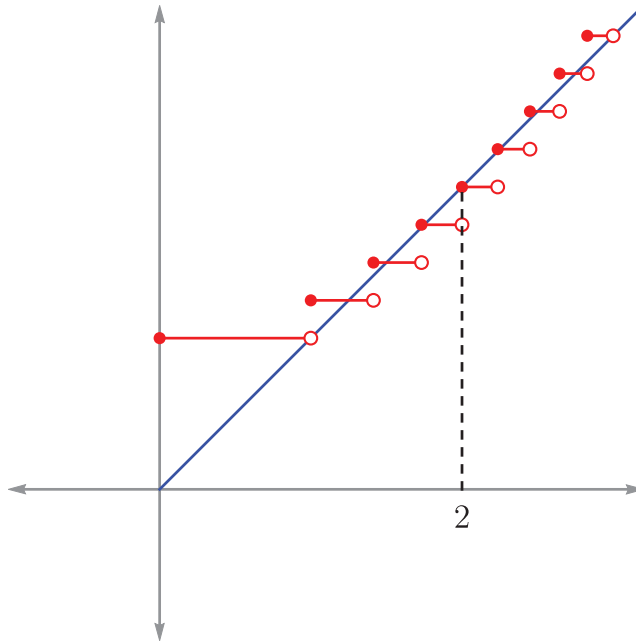

Figure 4   Equation (5) when $n = 2$: solutions to $(\lfloor x^2 \rfloor + 4)/4 = x$.

To see that an arbitrary number of solutions is possible is not difficult, but does require comment. A typical "step" in Figure 4, given k, is illustrated in Figure 5.
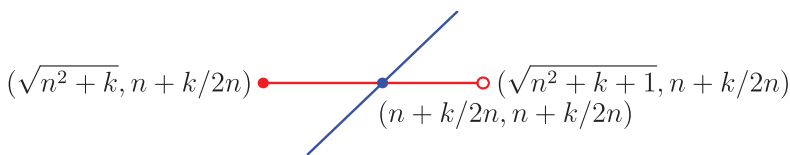

$(\sqrt{n^2 + k}, n + k/2n)$ •——————○ $(\sqrt{n^2 + k + 1}, n + k/2n)$
$(n + k/2n, n + k/2n)$

Figure 5   A typical step in Figure 4.

Thus, moving the line down by

$$\delta_k = \sqrt{n^2 + k + 1} - (n + k/2n)$$

will cause it to pass through the open circle, thereby missing an intersection. If two such differences were equal (say, for j and k), we would have

$$\sqrt{n^2 + j + 1} - \sqrt{n^2 + k + 1} \in \mathbb{Q}.$$

Now, it is not hard to see that $\sqrt{p} - \sqrt{q}$ is rational for integers p, q only if p and q are both perfect squares. But $n^2 + k + 1$ is a perfect square only when $k = -1$ since $|k| < 2n$. Thus, two such differences δ_j and δ_k cannot be equal. Therefore, we can suitably shift the line down so that an additional 1, 2, or 3 intersections are avoided, thereby obtaining any positive number of solutions.
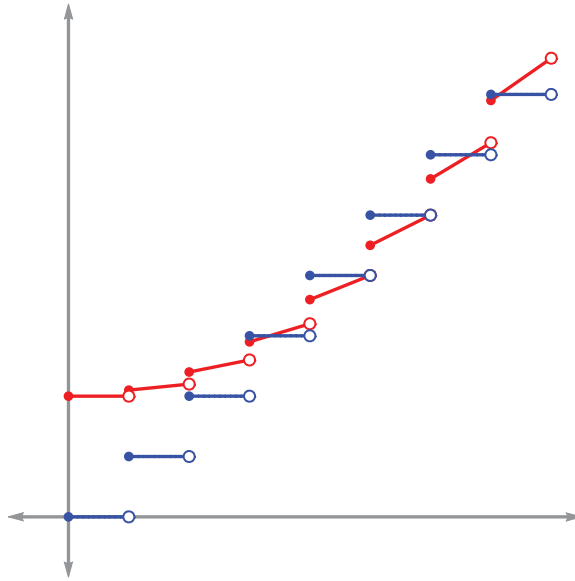
Figure 6 Equation (6) with $n = 5$: solutions to $(x \lfloor x \rfloor + 20)/10 = \lfloor x \rfloor$.

Equation (6)

Here, we put $b = -2n$ and $c = n(n-1)$, so that (6) becomes

$$x \lfloor x \rfloor - 2n \lfloor x \rfloor + n(n-1) = 0, \qquad (14)$$

illustrated in Figure 6 when $n = 5$.

As before, put $x = n + k + \varepsilon$, where $n, k \in \mathbb{Z}$ and $\varepsilon \in [0, 1)$. Then (14) becomes

$$k^2 + \varepsilon n + \varepsilon k - n = 0,$$

which, when $n \neq -k$, may be rewritten

$$\varepsilon = \frac{n - k^2}{n + k}.$$

It is not hard to show that $\varepsilon \in [0, 1)$ precisely when

$$-\sqrt{n} < k < -1 \quad \text{or} \quad 0 < k < \sqrt{n}$$

when n is not a perfect square, giving $2 \lfloor \sqrt{n} \rfloor - 1$ possible solutions. Note that when m is a positive integer, putting $n = m^2 + 1$ gives

$$2 \lfloor \sqrt{n} \rfloor - 1 = 2m - 1.$$

Thus, any number of odd solutions is possible in this case.

It is not difficult to show that all the ε's are different, since when $k_1 \neq k_2$, some simple algebra yields

$$\frac{n - k_1^2}{n + k_1} = \frac{n - k_2^2}{n + k_2} \implies (k_1 + k_2 + 1)n = -k_1 k_2,$$

which is impossible since $0 < |k_1 k_2| < n$ when k_1 and k_2 are in the ranges given above (again, assuming n is not a perfect square). Thus, shifting the quadratic to the left by slightly more than the smallest ε will result in an even number of solutions.

Equation (7)

In this case, we use the parameterization $b = -2n$ and $c = n(n-1)$, and therefore look at the equation

$$x^2 - 2n \lfloor x \rfloor + n(n-1) = 0.$$

As we did with equations (3) and (4), we look at

$$x = n + k + \varepsilon,$$

where $n, k \in \mathbb{Z}$ and $\varepsilon \in [0, 1)$. This results in

$$2\varepsilon(n + k) + \varepsilon^2 = n - k^2. \tag{15}$$

It is not difficult to show that this quadratic in ε has no real solutions when $k < -n$. When $k \geq -n$, the left-hand side of (15) is increasing and has range

$$[0, 2n + 2k + 1)$$

as ε varies over $[0, 1)$. Thus, because

$$0 \leq n - k^2 < 2n + 2k + 1$$

when $|k| \leq \lfloor \sqrt{n} \rfloor$, we may always find exactly one ε for such k. When $n = 0$, this gives one solution, and for a positive integer m, putting $n = m^2 + 1$ results in $2m + 1$ solutions. Thus, an arbitrary number of odd solutions is possible. One case with three solutions when $m = 1$ is shown in Figure 7.
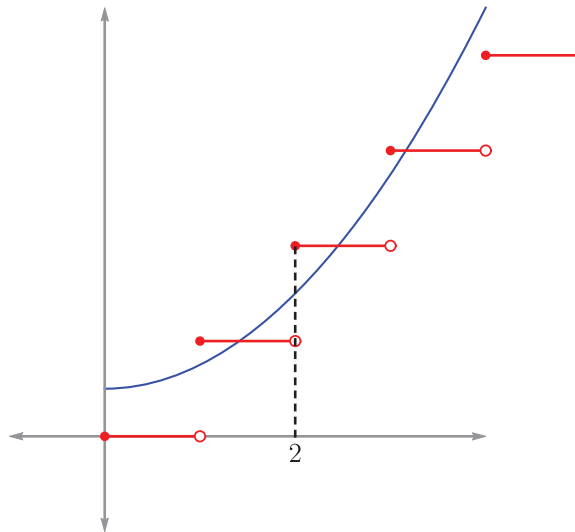
Figure 7   Equation (7) with $n = 2$: solutions to $(x^2 + 2)/4 = \lfloor x \rfloor$.

To see that any even number of solutions is possible, we proceed with the equation

$$x^2 - (2n + 1) \lfloor x \rfloor + n(n-1) = 0.$$

In looking for k and ε with

$$2\varepsilon(n + k) + \varepsilon^2 = 2n - k^2 + k,$$

it is not hard to show that this quadratic in ε has no real solutions when $k < -n$. When $k \geq -n$, the left-hand side must be nonnegative for $\varepsilon \in [0, 1)$, and so the inequality

$$k^2 - k - 2n \leq 0 \tag{16}$$

is generated. An analysis exactly like that following (12) may then be carried out, since replacing k with $-k$ in (16) is quite similar to (12). We do need to allow $n = 0$, however, to obtain two solutions.

We remark that it is not difficult to show that a shift of the quadratic to the left (as with Equation (6)) will also result in an even number of solutions. But we wish to use the results here to analyze (8)—and the shift argument would be unwieldy in that case, as will be remarked upon later.

Equation (8)

Again, we choose $b = -2n$ and $c = n(n - 1)$, giving

$$\lfloor x^2 \rfloor - 2n \lfloor x \rfloor + n(n - 1) = 0. \tag{17}$$

Most of the work in studying this equation is already done as a result of studying (7). Observe that moving from x^2 to $\lfloor x^2 \rfloor$ will still generate the same number of intersections with $\lfloor x \rfloor$. However, in this case, the intersections will be intervals, as shown in Figure 8 (where $\lfloor x \rfloor$ is graphed as a dashed horizontal line). Thus, the same argument given for (7) applies here, so that (8) may have a solution containing an arbitrary number of intervals. We remark about the difficulty of shifting the case with an odd number of solutions to one with an even number: the size of the intervals of the odd-numbered cases would be important as well, which would involve substantially more calculation.

In any case, it turns out that the sum of the lengths of these intervals is bounded. It is not difficult to show this by case analysis; we illustrate with one case here.
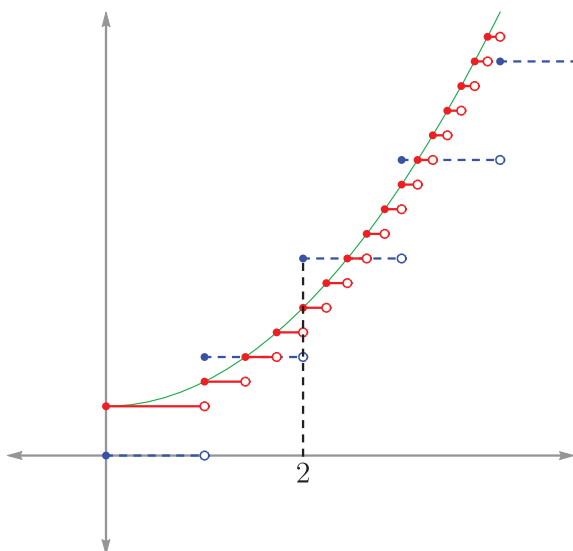

Figure 8   Equation (8) with $n = 2$: solutions to $\left(\lfloor x^2 \rfloor + 2 \right) / 4 = \lfloor x \rfloor$. The quadratic $y = (x^2 + 2)/4$ is included for reference.

As we want to show boundedness generally, we must consider the arbitrary quadratic

$$y_1 = \lfloor x^2 \rfloor + b \lfloor x \rfloor + c.$$

We look at the case $b < 0$, and consider the quadratic

$$y_2 = x^2 + bx + c - 1.$$

We note, that since $b < 0$, we have

$$x^2 - 1 < \lfloor x^2 \rfloor, \qquad bx \leq b \lfloor x \rfloor.$$

This implies that

$$y_2 = x^2 + bx + c - 1 < \lfloor x^2 \rfloor + b \lfloor x \rfloor + c = y_1.$$

Since the graph of y_1 is above the graph of y_2, any solutions to $y_1 = 0$ must lie between the solutions to $y_2 = 0$, which is a finite interval. (Of course, if $y_2 = 0$ has no solutions, then neither does $y_1 = 0$.) The case $b > 0$ is handled similarly.

Concluding remarks

By replacing "x" with "$\lfloor x \rfloor$" in the usual quadratic equation, a remarkable range of behaviors can be observed. The more interesting cases, equations (3)–(8), allow arbitrarily many solutions. Moreover, equation (8) admits arbitrarily many *intervals* as solutions. What remains elusive is a general boundedness result. Based on numerous graphical examples, it seems that the sum of the lengths of the solution intervals is never more than 2—which is the sum of the lengths of solution intervals of simple equations, like (2). A simple proof of this conjecture—or a counterexample—would be most welcome.

Finally, a similar analysis may be applied to replacing occurrences of x with the ceiling function, $\lceil x \rceil$, by using the relationship $\lceil x \rceil = -\lfloor -x \rfloor$.

Summary.   What happens when the floor function meets the quadratic equation? Replacing one or more occurrences of the variable in a quadratic equation with its floor results in diverse and interesting behavior. The new quadratic may have arbitrarily many solutions—even arbitrarily many intervals as solutions. In this case, however, the sum of the lengths of the intervals is bounded. We examine all possible cases of such replacement, and show how it is possible to find quadratics with any desired number of solutions or solution intervals.

VINCENT J. MATSKO (MR Author ID: 692090) is a former college professor turned consultant. He now has time to work on unfinished projects, such as his book on applying spherical trigonometry to the study of uniform polyhedra, as well as to enjoy traveling, cooking, and yoga.

An Optimal Oval

IVAR HENNING SKAU
University of South-Eastern Norway
3800 Bø, Telemark
ivar.skau@usn.no

KAI FORSBERG KRISTENSEN
University of South-Eastern Norway
3918 Porsgrunn, Telemark
kai.f.kristensen@usn.no

The Danish writer and inventor Piet Hein had aesthetic optimality in mind when he sought to find the simplest and most pleasing curve between an ellipse and a rectangle. He considered superellipses, of the form $(|x|/a)^n + (|y|/b)^n = 1$, where $a > b > 0$ and $n > 2$. For some reason, he concluded that the choice $n = 2.5$ (see Figure 1) would give the most beautiful curves (see [1]). In 1968, production of superelliptic tables was initiated with $a/b = 3/2$, and the landmark known as "Sergel's square" in Stockholm ($a/b = 6/5$) shows that superellipses have been used in urban architectural design.

Figure 1 A Piet Hein superellipse.

We will mainly consider race tracks, and since the participants are exposed to centripetal forces proportional to curvature, optimality in this connection has to be defined by the curvature properties of an oval.

Definition 1. An oval is a closed, plane curve which is symmetric with respect to two perpendicular axes, and which forms the boundary of a convex region.

Traditionally, "centripetal shocks" along tracks have been avoided using clothoids (also called Euler spirals) to smooth the transition between straight lines and circular segments. This is also a well-known technique for the construction of roads, railways, and roller coasters. However, here we will suggest a track with a smoother curvature behavior, expressed by a single formula. We call the track optimal since the optimization criteria we shall apply might well agree with how participants would perceive track quality.

Minimum requirements for oval race tracks

Ordinary, and in many ways evident, minimum requirements for an oval representing a horse race track or a velodrome, may be formulated in the following way:

(O1) The oval must be sufficiently smooth that the curvature can be calculated.

(O2) When inscribed in a rectangle with vertices $(\pm a, \pm b)$, where $a > b > 0$, the oval is required to have zero curvature at the midpoints $(0, \pm b)$ of the long sides.

(O3) Moving from $(0, \pm b)$ the curvature is a non-decreasing function until it reaches its maximum at the points $(\pm a, 0)$, called the horizontal extreme points.

A track consisting of two parallel lines and two semi circles, as shown upper left in Figure 2, will satisfy these requirements but will be unfavorable, especially as a horse race track, because of the sudden changes in curvature. This is justified by taking a quick look at the discontinuous curve at the bottom left of Figure 2.
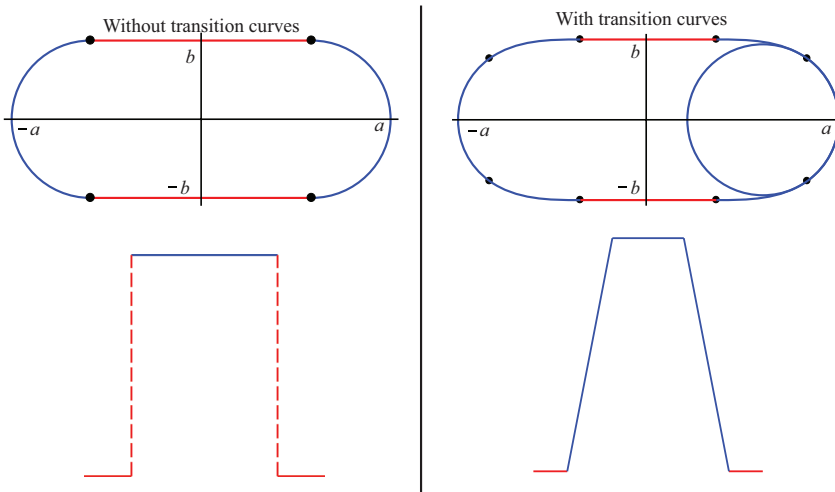
Figure 2   The graphs in the bottom row show the curvature between the two midpoints $(0, \pm b)$ as a function of arc length.

The curvature graph at the bottom right of Figure 2 demonstrates that the use of clothoids gives a radical improvement. This is no surprise, since a clothoid is defined to be a curve where the curvature changes linearly with its curve length. At the transition points, the clothoids are connected to the circles and straight lines with the same slope *and curvature*. Raph Levien gives a compact historical treatment of those curves, based on his own PhD thesis. His presentation provides a number of interesting applications [**2**].

However, a problem connected to the use of clothoid transition curves arises because of banking. Moving along the clothoid from B to D in Figure 3, the banked track outside the clothoid curve must be made gradually steeper because of the increasing curvature. So when the banking is "correctly" performed, an unfortunate edge CD will be generated across the track since the slope suddenly stops increasing when circular motion (DF and CE) takes over. A minor, but similar, phenomenon occurs in AB, where the slope suddenly starts increasing. With a unified formula this problem is eliminated and no smoothing will be needed.

Figure 3   When clothoids are used, the banked track has an edge CD across the track in the turn, through the transition point D.

Optimality criteria

To avoid edges due to transition points and to obtain a smoother curvature behavior, it is natural to restrict our interest to curves expressible by the formula

$$\left(\frac{|x|}{a}\right)^m + \left(\frac{|y|}{b}\right)^n = 1 \quad (a > b > 0). \tag{1}$$

Since the curvature κ is defined to be the rate of change $d\theta/ds$ of the rotation angle θ along the curve, we will obtain the result 2π when we perform the integral $\int_C \kappa \, ds$, traversing the curve C counterclockwise. Our goal is to determine m and n so that the shape of the oval distributes the curvature in an optimal way, given by the following definition:

Definition 2. An optimal oval is the curve, satisfying (1) and (O1)–(O3), that for fixed a and b ($a > b > 0$) has the least maximum curvature at the horizontal extreme points $(\pm a, 0)$.

Definition 2 makes sense, especially considering horse race tracks. Smooth curvature behavior (supplied by (1) and (O1)–(O3)) will reduce the risk of galloping (at CD), while minimizing the maximum curvature will be beneficial since the corresponding centripetal force will be minimized as well.

In Figure 4, an optimal oval (Oval 1) is sketched, together with an oval based on clothoids (Oval 2). Their curvature graphs are also displayed. The two ovals have the same arc length.
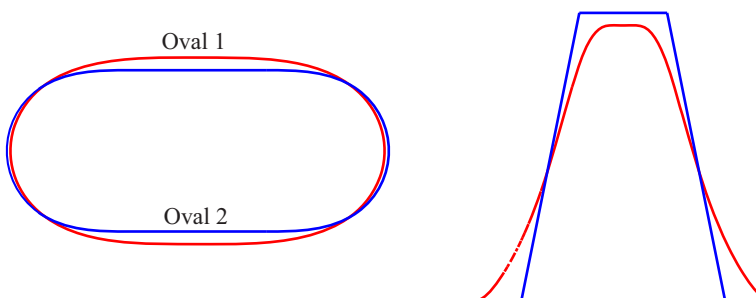

Figure 4   Oval 1 is an optimal oval, while clothoids have been used to generate Oval 2. To the right, we see their curvature graphs between the two points $(0, \pm b)$. Oval 1 represents the smoothest one.

The optimal oval

We shall now see how criteria (O1)–(O3) rule out several possible choices for m and n. We note that cases where $0 < m < 1$ or $0 < n < 1$ generate curves bounding non-convex regions, and possess at least two cusps. Figure 5 indicates that we must require $m > 2$ and $n = 2$ to fulfill these criteria.
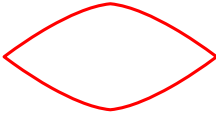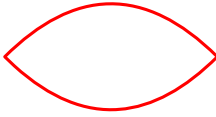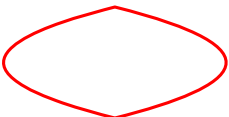
	$m < 2$	$m = 2$	$m > 2$
$n < 2$	$|X|^{1.5}/2^{1.5} + |Y| = 1$	$X^2/2^2 + |Y| = 1$	$|X|^3/2^3 + |Y| = 1$
$n = 2$	$|X|/2 + Y^2 = 1$	$X^2/2^2 + Y^2 = 1$	$|X|^3/2^3 + Y^2 = 1$
$n > 2$	$|X|^{1.5}/2^{1.5} + |Y|^3 = 1$	$X^2/2^2 + |Y|^3 = 1$	$|X|^3/2^3 + |Y|^3 = 1$

Figure 5   The oval types depend on the exponents m and n.

However, to see theoretically how m and n will affect curvature, we need to take a closer look at the curvature formula

$$\kappa(x, y) = \frac{|x''(y)|}{(1 + (x'(y))^2)^{3/2}}.$$

We first simplify the situation by substituting $x = bX$ and $y = bY$, a similarity transformation that will not change the shape of the curve. Letting K denote the curvature of the transformed curve in the (X, Y) plane, it is straightforward to see the relationship

$$K(X, Y) = b \cdot \kappa(x, y). \tag{2}$$

For convenience, we proceed by considering the first quadrant only. Putting $\alpha \equiv a/b$, we may now rewrite equation (1) to get

$$X = \alpha(1 - Y^n)^{1/m}, \quad 0 \le X \le \alpha, \ 0 \le Y \le 1.$$

In the remaining part of this section we find it suitable to emphasize K's dependence on Y and $k = 1/m$ by writing $K = K(Y, k)$.

When the necessary differentiations are done, we obtain

$$K(Y, k) = \frac{|X''(Y)|}{(1 + (X'(Y))^2)^{3/2}}$$

$$= \frac{nk\alpha\{(n-1)(1-Y^n)^{k-1}Y^{n-2} - n(k-1)Y^{2n-2}(1-Y^n)^{k-2}\}}{(1 + n^2\alpha^2 k^2 Y^{2n-2}(1-Y^n)^{2k-2})^{3/2}}.$$

We recall that $Y = 0$ corresponds to the horizontal extreme points $(\pm a, 0)$ and that requirement (O3) will exclude 0 as a possible value of K here. When $n > 2$ and $Y = 0$, both terms in the numerator are 0 while the denominator equals 1. Further on, when $1 < n < 2$, we realize that the factor Y^{n-2}, in the numerator's first term, will dominate and make $K \to \infty$ when $Y \to 0$. Hence we have obtained

$$K(0, k) = \begin{cases} 0 & \text{if } n < 2 \\ \infty & \text{if } 1 < n < 2, \end{cases}$$

and we conclude that $n = 2$. One could ask if a slightly smaller n value would cause trouble in a real horse race. The answer is no, since horses in that case will be exposed to just a tiny change in the slope of the tangent when passing through the horizontal extreme points. Furthermore, the curvature collapse at $Y = 0$ is just a one point collapse, unable to affect a running horse.

We can now write $K(Y, k) = G(Y^2, k)$, where

$$G(t, k) = \frac{2k\alpha\{(1-t)^{2-2k} - 2(k-1)t(1-t)^{1-2k}\}}{((1-t)^{2-2k} + 4\alpha^2 k^2 t)^{3/2}}.$$

The problem is that different values of k will give different curvature behaviors expressed by $G(t, k)$, some of which are more favorable than others with respect to optimality in Definition 2. So from now on we are looking for an optimal k value. Even though $G(t, k)$ may look rather complicated, the calculations, allowing us to find k, are surprisingly simple. In fact, we end up solving a quadratic equation.

First, let us show how $k < 1/2$, which is equivalent to $m > 2$, can be deduced from requirement (O2). This requirement states that the curvature has to be zero at the midpoints $(0, \pm b)$ of the track, corresponding to $(0, \pm 1)$ in (X, Y) coordinates. Now $Y \to 1$ also implies $t = Y^2 \to 1$, so we will look into how $\lim_{t \to 1} G(t, k)$ depends on k. We observe that the denominator of G will never give us any trouble, but the numerator is infinite at $t = 1$ when $k > 1/2$. Hence we must have $k \le 1/2$, from which we can exclude $k = 1/2$ because $\lim_{t \to 1} G(t, k) > 0$ in the elliptic case (where also the horizontal extreme points are too acute). When $k < 1/2$, i.e., $m > 2$, we have $\lim_{Y \to 1} K(Y, k) = 0$, fulfilling (O2).

So far we know that $n = 2$ and $m > 2$, but there is still a ways to go. For example, the oval given by $|X|^3/2^3 + Y^2 = 1$ in Figure 5 satisfies these requirements, but it is not optimal. The following calculations will show why the exponent $m = 3$ in this case must be replaced by $m = (1 + \sqrt{33})/2 \approx 3.37$ to achieve optimality. In general, it remains to determine m such that the curvature increases along the curve from $(0, \pm 1)$ to $(\pm\alpha, 0)$ (requirement (O3)) *and* such that the maximum curvature at $(\pm\alpha, 0)$, i.e.,

$$K(0, k) = G(0, k) = 2\alpha k = 2\alpha/m, \tag{3}$$

becomes as small as possible (Definition 2). In other words, we must find the greatest value of m (or the smallest value of k) that is compatible with the first of these two claims.

Figure 6 displays the curvature surface, given by $z = K(Y, k)$ when $\alpha = 2$ and $0.2 \le k \le 0.4$. The curve $z = K(0, k)$ represents the linear curvature of the horizontal extreme points. If we fix a sufficiently small k, we notice that the Y-curve $z = K(Y, k)$ will have a minimum when $Y = 0$. As k increases, we will obtain the particular value $k = k_\alpha$ where $K(0, k)$ changes from being a minimum to being a maximum. A formula

of k_α may therefore be found by the requirement $K_{YY}(0, k) = 0$. The second curve marked in Figure 6 is given by $z = K(Y, k_\alpha)$ when $\alpha = 2$. Since
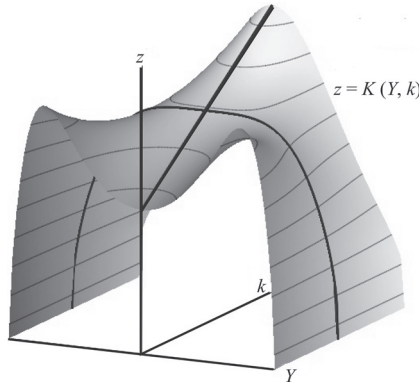

Figure 6   Here the curvature is represented by a surface $z = K(Y, k)$ when $\alpha = 2$.

$$K(Y, k) = G(Y^2, k) = G(t, k),$$

we have

$$K_{YY}(0, k) = 2G_t(0, k) = -12k\alpha(2\alpha^2 k^2 + k - 1),$$

resulting in the positive solution

$$k = k_\alpha = \frac{-1 + \sqrt{1 + 8\alpha^2}}{4\alpha^2},$$

of the equation $K_{YY}(0, k) = 0$. We note that $K_{YY}(0, k) > 0$ when $k < k_\alpha$ and $K_{YY}(0, k) < 0$ when $k > k_\alpha$, matching what we observe in Figure 6. The corresponding exponent $m_\alpha = 1/k_\alpha$ becomes

$$m_\alpha = \frac{1}{2} + \frac{1}{2}\sqrt{1 + 8\alpha^2}.$$

Before we can conclude that $m = m_\alpha$ will give the optimal solution to our problem, it remains to prove that $K(Y, k)$ has the desired shape when $k = k_\alpha$. To be more specific, we must show that the curvature decreases when $Y^2 \in \langle 0, 1 \rangle$, which is the same as proving that $G(t, k)$ decreases when $0 < t < 1$. Now we substitute $u = 1 - t$ and $\beta = 2 - 2k$ in G to obtain

$$g(u) = \frac{(2 - \beta)\alpha(u^\beta + \beta(1 - u)u^{\beta-1})}{(u^\beta + \beta(1 - u))^{3/2}},$$

where we have used the relation $2\alpha^2 k^2 + k - 1 = 0$ to see that $4\alpha^2 k^2 = \beta$. Since $1 < \beta < 2$ and $0 < u < 1$, differentiation shows that the numerator increases and the denominator decreases, giving that $g(u)$ is an increasing function. Hence we can conclude that all the requirements of an optimal oval are fulfilled, giving the formula

$$\left(\frac{|X|}{\alpha}\right)^{\frac{1}{2}\left(1+\sqrt{1+8\alpha^2}\right)} + Y^2 = 1.$$

By resubstituting

$$X = x/b, \quad Y = y/b, \quad \text{and} \quad \alpha = a/b,$$

we get, finally

$$\left(\frac{|x|}{a}\right)^{\frac{1}{2}\left(1+\sqrt{1+8\left(\frac{a}{b}\right)^2}\right)} + \left(\frac{y}{b}\right)^2 = 1. \tag{4}$$

$a/b = 2.5$ $a/b = 2$ $a/b = (1 + \sqrt{5})/2 \approx 1.62$
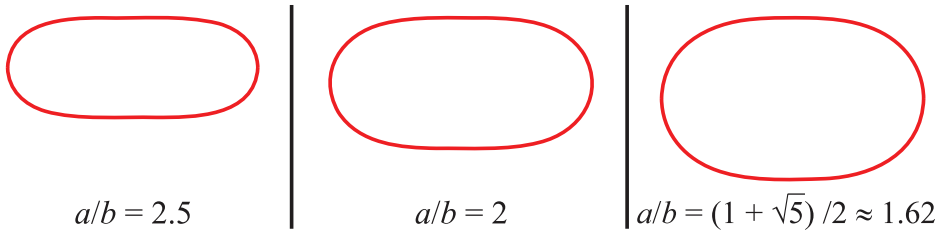
Figure 7   Optimal ovals.

According to (2) and (3), the curvature of (1) in $(\pm a, 0)$ becomes

$$\kappa = \frac{2ka}{b^2} = \frac{2a}{mb^2}.$$

One notices that the turn of an optimal oval in Figure 7 is partly an approximate circle segment, in accordance with Figures 4 and 6. Note also that the limit curve as $a \to b$ in (4), is a circle.

When designing real horse race tracks or velodromes, one has to decide what will be a favorable ratio a/b, and a and b must be scaled to fulfill the requirements of track length. In addition, one must adjust the banking according to how fast horses or bicycles are expected to move.

Comparing ovals of equal length

In Definition 2, a and b are required to be fixed. This means that we have compared curves within a fixed rectangle to find the one that is optimal. However, all relevant curves (satisfying (1) and (O1)–(O3)) have different lengths, and hence one might object that the comparison is unfair. In this section, we show that the optimal oval found in the preceding section will also be optimal when we compare curves of equal length within rectangles with the same length/width ratio.

Let us fix $a > 1$ and consider curves satisfying

$$(|x|/a)^m + y^2 = 1.$$

Now we perform the transformation $x = S(a, m)X$ and $y = S(a, m)Y$, where

$$S(a, m) = 4 \int_0^a \sqrt{1 + (dy/dx)^2} \, dx$$

is the curve length. Substituting $b = S(a, m)$ in (2), we get

$$K(X, Y) = S(a, m) \cdot \kappa(x, y).$$

This transformation will generate curves of length 1 in the (X, Y) plane because $dY/dX = dy/dx$ and $dX = dx/S(a, m)$. Since a is fixed, the rectangles inscribing curves of length 1 will be similar, with half axes given by $a/S(a, m)$ and $1/S(a, m)$. We notice that the calibrated curvature

$$S(a, m)\kappa(x(y), y)$$

varies as a function of y solely through κ. This is why we can inherit much of the reasoning that helped us find the optimal oval in the fixed frame case. Thus it suffices to prove that

$$h(a, m) \equiv S(a, m)\kappa(a, 0)$$

decreases with m (increases with $k = 1/m$). Since $\kappa(a, 0) = 2a/m$, we have

$$h(a, m) = \frac{2a}{m} \cdot 4 \int_0^a \sqrt{1 + (dy/dx)^2} \, dx.$$

We substitute $u = x/a$ to obtain

$$h(a, m) = 4a \int_0^1 \sqrt{\frac{4a^2}{m^2} + \frac{u^{2m-2}}{1 - u^m}} \, du.$$

Since $0 < u < 1$ and $m > 2$, we observe that both terms in the square root decreases with increasing m, and the conclusion is that the optimal oval is given by (4), also when comparing ovals of equal length within similar rectangles.

REFERENCES

[1] Gardner, M. (1990). *Mathematical Carnival*. New York: Penguin Books, pp. 240–254.
[2] Levien, R. (2008). The Euler spiral, a mathematical history. PhD dissertation. University of California at Berkeley. levien.com/phd/euler_hist.pdf

Summary. In designing horse race tracks and velodromes, clothoids, also called Euler spirals, have been used to achieve a smooth transition between straight and circular track segments. This is also a well-known technique for the construction of roads, railways, and roller coasters. Clothoids provide a steady increase in curvature, preventing participants from experiencing a "centripetal shock." We suggest a track with a smoother curvature behavior, expressed by a single formula. We call the track optimal since the optimization criteria we apply might well agree with how participants would perceive track quality.

IVAR HENNING SKAU (MR Author ID: 358725) received his cand. real. degree in mathematics from the University of Oslo in 1967. He has had a long career teaching mathematics and statistics to various types of students, mainly at the University of South-Eastern Norway. His research interests are primarily in number theory and combinatorics.

KAI FORSBERG KRISTENSEN (MR Author ID: 600751) received his cand. scient. degree in mathematics from the University of Oslo in 1985. For most of his professional career he has been teaching mathematics and statistics to students of engineering and technology, but he also has experience with educating teachers and candidates of business administration. He is happy whenever he can help illuminate classical themes in new ways.

Exploring the Evolutes of Cubic Polynomials

RUSSELL A. GORDON
Whitman College
Walla Walla, WA 99362
gordon@whitman.edu

The derivative of a function f at a given point tells us the slope of the line that best approximates the curve $y = f(x)$ at that point. While not as familiar, we can also determine the circle that best approximates the curve at a given point. The center of this circle is referred to as the center of curvature. These two concepts are illustrated in the left and middle graphs found in Figure 1 for the function $f(x) = x^3$ at the point $(0.7, 0.343)$.
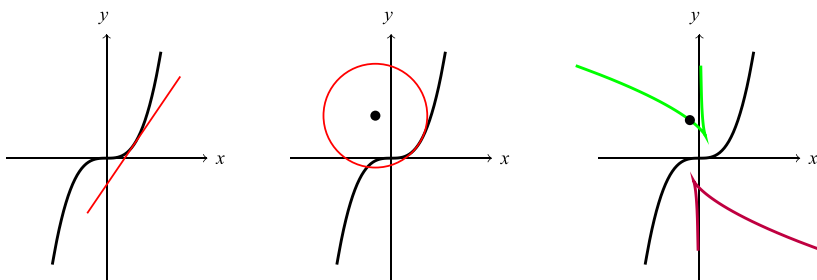

Figure 1   The line of best fit, the circle of best fit, and the evolute.

Just as the graph of $y = f'(x)$ reveals information about the function f, the curve traced out by all the centers of curvature also provides insight into the properties of f. This curve (illustrated by the graph on the right in Figure 1) is known as the *evolute* of the given curve $y = f(x)$, and examples of such curves have been studied for centuries (see Coolidge [3] and Eves [4]). For instance, the evolute of an ellipse is an astroid (see Hartmann [6] or Lawrence [7]) and the evolute of a parabola is a semicubical parabola (see Bains [1] or Lawrence [7]). In general, the evolute of a polynomial becomes dramatically more complicated as the degree of the polynomial increases. We will confine our attention to a careful study of the evolutes of cubic polynomials.

Given a twice differentiable function f defined on $\mathbb{R}$, the center of curvature $(h(x), k(x))$ at a point x is given by

$$h(x) = x - \frac{f'(x)}{f''(x)}\big(1 + (f'(x))^2\big)$$

$$k(x) = f(x) + \frac{1}{f''(x)}\big(1 + (f'(x))^2\big).$$

(see Burgette [2], Giv [5], Lawrence [7], or Leithold [8]; the first two references derive the formulas without reference to arc length). At those values of x for which $f''(x) = 0$, the curve $y = f(x)$ is essentially linear, and the circle of best fit has infinite radius. The evolute of $y = f(x)$ is the curve defined parametrically by the points $(h(x), k(x))$ as x ranges over the set of real numbers. For a cubic polynomial, we expect the evolute to have two branches since there is no point defined on the evolute when $f''(x) = 0$.

In fact, if $f''(x_0) = 0$, then the normal line at the point $(x_0, f(x_0))$ is a slant asymptote for the evolute. Since pictures speak much more clearly than words, we present (see Figure 2) the graphs of the evolutes of four different cubic equations so that the reader can visualize how these curves look and change when different coefficients are used for the cubic equations.
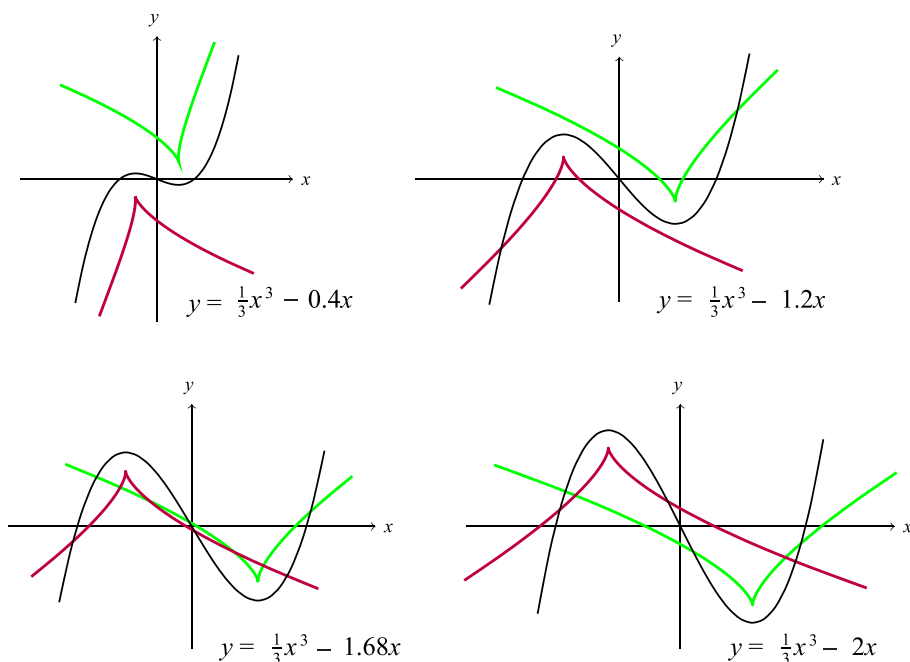

$$y = \tfrac{1}{3}x^3 - 0.4x$$

$$y = \tfrac{1}{3}x^3 - 1.2x$$

$$y = \tfrac{1}{3}x^3 - 1.68x$$

$$y = \tfrac{1}{3}x^3 - 2x$$

Figure 2   The evolutes of several cubic graphs.

It is evident that all of these evolute curves have two cusps. As we will soon see, the structure of these cusps is more intricate than these graphs indicate.

Based on these pictures, we pose and solve the following problems for curves of the form $y = \tfrac{1}{3}x^3 - \lambda x$.

1. Find the value of λ where the cusp of one branch of the evolute touches the other branch.
2. Find the value of λ when the region between the two branches of the evolute first becomes a connected region.
3. Express the area of the connected region in analytic terms as a function of λ.
4. Determine the asymptotic behavior of this area as $\lambda \to \infty$.

In addition to these questions, for each point (s, t) in the plane, let $N_\lambda(s, t)$ denote the number of points on the curve $y = \tfrac{1}{3}x^3 - \lambda x$ for which the normal line at that point passes through (s, t). It turns out that the evolute divides the plane into regions where this number is constant (see Sanchez [9]). In fact, it is this connection that allows us to answer the above questions.

Before proceeding, we need to establish that there really is only one parameter to consider for cubics. Given the cubic polynomial $C(x) = ax^3 + bx^2 + cx + d$, we find that

$$T(x) \equiv C\left(x - \frac{b}{3a}\right) - \frac{2b^3 - 9abc + 27a^2d}{27a^2} = ax^3 + \frac{3ac - b^2}{3a}x.$$

This shows that a simple translation of the axes reduces the number of parameters to two. Without loss of generality (the other option is just a reflection about the y-axis), we may assume that $a > 0$. By choosing appropriate constants and relabeling, we can write $y = T(x)$ as

$$y = \frac{a^2}{3} x^3 \pm b^2 x \quad \Leftrightarrow \quad ay = \frac{1}{3}(ax)^3 \pm b^2(ax).$$

This last equation is just a rescaling of the equation $y = \frac{1}{3}x^3 \pm b^2 x$. Hence, it is suffi-cient to answer the above questions for cubics of this form.

General properties of the cubic evolute

To answer the questions concerning properties of the cusp of the evolute, we first need to obtain a clear picture of the shape of the cusp. Since the parameter λ is posi-tive when the cusps have the various properties described in our four questions posed above, let $f(x) = \frac{1}{3}x^3 - b^2 x$, where b is some positive constant. For such functions, the graphs have two relative extrema. From the graphs and first impressions, it seems as though the tips of the cusps of the evolutes correspond to the relative extreme points on the cubic graph. For example, it is easy to believe that the smallest radius for an approximating circle occurs at a relative extremum. However, since the evolutes of cubic graphs without relative extrema also have cusps, it becomes clear that other factors must be involved. To see what the cusps look like up close, we consider the curve $y = \frac{1}{3}x^3 - 2x$ and focus on the cusp that appears in the fourth quadrant (refer to Figure 2). Omitting the algebraic details behind finding the equations for the center of curvature (these will be discussed shortly), a magnified view of this cusp can be obtained by looking at the parametric equations

$$h(x) = 10000 \left(-\frac{1}{2} x^5 + 3x^3 - \frac{11}{2} x + \frac{5}{x} \right) - 14140;$$

$$k(x) = 1000 \left(\frac{5}{6} x^3 - 4x + \frac{5}{2x} \right) + 1536,$$

as x ranges from 1.40 to 1.47. The resulting image (which, as these equations show, has been scaled differently for the $h(x)$ and $k(x)$ values) is shown in Figure 3.
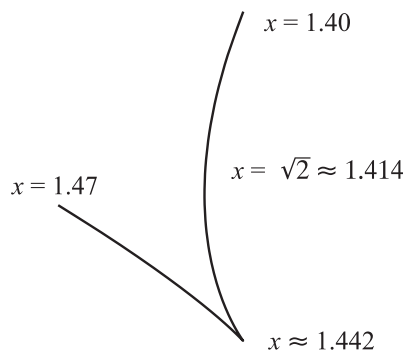

Figure 3 A magnified version of the cusp of an evolute.

From this picture, we expect $h'(x)$ to have two zeros in this range of x values while $k'(x)$ will only have one zero. Recalling that the parametric equations

$$h(x) = x - \frac{f'(x)}{f''(x)} \left(1 + (f'(x))^2\right)$$

$$k(x) = f(x) + \frac{1}{f''(x)} \left(1 + (f'(x))^2\right)$$

define the evolute, we find that

$$h(x) = x - f'(x) \left((k(x) - f(x))\right)$$

$$h'(x) = 1 - f'(x) \left(k'(x) - f'(x)\right) - f''(x) \left(k(x) - f(x)\right) = -f'(x)k'(x).$$

If we consider the curve $x = h(t)$, $y = k(t)$ defined parametrically, then both x and y (most likely even in the case of higher degree polynomials) change directions at values of t for which $k'(t) = 0$, while just x changes directions at values of t for which $f'(t) = 0$ (typically the relative extreme values of the graph of the polynomial). Note that if f' has no zeros, then the cusps have a simpler (more "vee-like") structure.

Applying the general formulas for the evolute to the function $f(x) = \frac{1}{3}x^3 - b^2 x$, we find that

$$h(x) = x - \frac{x^2 - b^2}{2x} \left(1 + (x^2 - b^2)^2\right)$$

$$= -\frac{1}{2} x^5 + \frac{3}{2} b^2 x^3 + \frac{1}{2} (1 - 3b^4)x + \frac{b^2(b^4 + 1)}{2x},$$

$$k(x) = \frac{1}{3} x^3 - b^2 x + \frac{1}{2x} \left(1 + (x^2 - b^2)^2\right)$$

$$= \frac{5}{6} x^3 - 2b^2 x + \frac{b^4 + 1}{2x},$$

$$k'(x) = \frac{5}{2} x^2 - 2b^2 - \frac{b^4 + 1}{2x^2}.$$

It is easy to verify that $k'(x) = 0$ when

$$x = \pm\sqrt{\frac{1}{5} \left(2b^2 + \sqrt{9b^4 + 5}\right)}.$$

As indicated in Figure 3, the x value corresponding to the plus sign is larger than b. Hence, the smooth bend in the magnified cusp graph occurs when $x = b$ and the cusp itself occurs when x has the positive value recorded above. As b increases, these two x values become very close (in fact, they are rather close even for small b values) and they are equal in the limit, but each of the cusps on the evolute of a cubic with relative extrema has the general shape illustrated in Figure 3.

As an aside, we make a couple of observations on the general shape of the evolute graph for this one parameter family of cubic functions. For values of x near 0, we see that

$$h(x) \approx b^2(b^4 + 1)/(2x) \quad \text{and} \quad k(x) \approx (b^4 + 1)/(2x).$$

Thus, the graph of the evolute resembles the line $y = x/b^2$, the normal line to the cubic at the origin. For large values of x, the functions h and k behave like $-\frac{1}{2}x^5$ and $\frac{5}{6}x^3$, respectively. Hence, the other ends of the evolute graph approach the curve $y = -\frac{5}{6}(2x)^{3/5}$ asymptotically as $|x|$ goes to infinity. These two extreme behaviors are illustrated in Figure 4 for the evolute of the curve $y = \frac{1}{3}x^3 - x$. The dashed line represents the normal to the cubic at the origin and one side of each branch of the evolute approaches this line as x goes to 0. The dashed curve in the figure is the graph of $y = -\frac{5}{6}(2x)^{3/5}$ and it indicates the shape of the evolute graph as $|x|$ goes to infinity. It is interesting to note that the asymptotic behavior of the evolute of the curve $y = \frac{1}{3}x^3 \pm b^2x$ as $|x|$ goes to infinity is independent of b. This fact is somewhat evident in the collection of graphs presented in Figure 2.
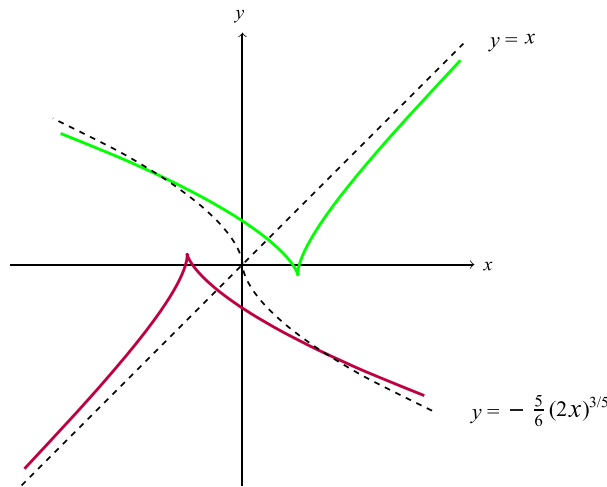

Figure 4   Asymptotic behavior of the graph of the evolute.

Values for λ when the cusps have certain properties

To answer the first two questions posed earlier, we need to examine more closely the connection between the evolute and the number of normals. A point (s, t) lies on a normal line of the curve $y = f(x)$ when

$$\frac{f(x) - t}{x - s} = -\frac{1}{f'(x)} \quad \Leftrightarrow \quad f(x)f'(x) - tf'(x) + x - s = 0.$$

If f is a polynomial of degree n, this last equation generates a polynomial of degree $2n - 1$. As shown in Sanchez [9], this polynomial has a repeated root whenever the point (s, t) lies on the evolute. If, further, the point (s, t) is at the tip of the cusp, then the polynomial has a root of multiplicity three. Figure 5 helps illustrate these ideas for the cubic $y = \frac{1}{3}x^3 - 2x$. For each graph, the star represents a point in the plane and the bullets indicate the points on the curve where the normal line passes through the star. By looking at these graphs, you can notice how the points on the cubic coalesce as the star approaches a cusp on the evolute, thus leading to a root of multiplicity three.

We now apply these ideas to our cubic function f defined by $f(x) = \frac{1}{3}x^3 - b^2x$. Referring to the equation
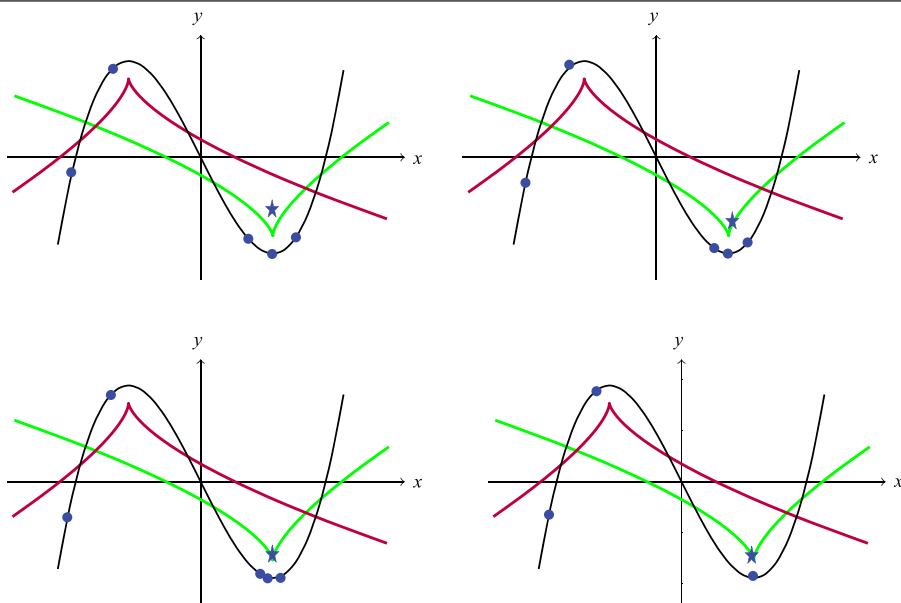
$$f(x)f'(x) - tf'(x) + x - s = 0,$$

Figure 5   The polynomial has a root of multiplicity three at a cusp.

let P be the polynomial given by

$$P(x) = 3\big(f(x)f'(x) - tf'(x) + x - s\big)$$
$$= (x^3 - 3b^2x)(x^2 - b^2) - 3t(x^2 - b^2) + 3x - 3s$$
$$= x^5 - 4b^2x^3 - 3tx^2 + 3(b^4 + 1)x + 3(b^2t - s).$$

In order for the cusp of one branch of the evolute to touch the other branch of the evolute, we need P to have two real roots, one of multiplicity 3 and the other of multiplicity 2. The root w of multiplicity 3 must have one of the values determined earlier for the location of the cusps. For definiteness, let

$$w = \sqrt{\frac{1}{5}\left(2b^2 + \sqrt{9b^4 + 5}\right)}.$$

Since the coefficient of x^4 is 0, we must choose a value for b so that the roots of P are $w, w, w, -\frac{3}{2}w$, and $-\frac{3}{2}w$. It follows that

$$P(x) = (x - w)^3 \left(x + \frac{3}{2}w\right)^2$$
$$= x^5 - \frac{15}{4}w^2x^3 + \frac{5}{4}w^3x^2 + \frac{15}{4}w^4x - \frac{9}{4}w^5.$$

Equating the x^3 coefficients, we find that

$$15w^2 = 16b^2 \;\;\Rightarrow\;\; 3\left(2b^2 + \sqrt{9b^4 + 5}\right) = 16b^2 \;\;\Rightarrow\;\; 9(9b^4 + 5) = 100b^4$$

and thus

$$b^2 = \sqrt{45/19} \approx 1.539.$$

Although it is a bit tedious, equating the other coefficients yields the same value for b^2. Note that this value for $\lambda = b^2$ is consistent with the set of graphs shown in Figure 2.

The second cusp question involves seeking a value for b^2 so that the region between the two branches of the evolute first becomes a connected region. Referring to the graphs in Figure 2 as well as the discussion concerning the number of normals, we see that the interior region is first connected when there are five points on the cubic curve for which the normal line passes through the origin. With $s = 0 = t$, the polynomial P defined earlier becomes

$$
\begin{aligned}
P(x) &= x^5 - 4b^2x^3 + 3(b^4 + 1)x \\
&= x\left((x^2 - 2b^2)^2 + (3 - b^4)\right).
\end{aligned}
$$

Assuming that $b^2 > \sqrt{3}$, both of the numbers $2b^2 \pm \sqrt{b^4 - 3}$ are positive. It follows that the polynomial P has five distinct real roots when $b^2 > \sqrt{3}$. Hence, the five normal line region is connected when the parameter λ is greater than $\sqrt{3}$. (The reader may find it interesting to determine the value of λ when the cusps lie on the x-axis.)

Area of the five normal line region for $\lambda > \sqrt{3}$

Consider once again the curve corresponding to the function $f(x) = \frac{1}{3}x^3 - b^2x$ and assume that $b^2 > \sqrt{3}$. We want to find the area of the region that consists of all of the points that lie on five normal lines of the curve. To be specific, we want to find the area of the shaded region in Figure 6. The quantities shown at the edges of the branches of the evolute curve indicate the values of x that generate the points on the evolute.
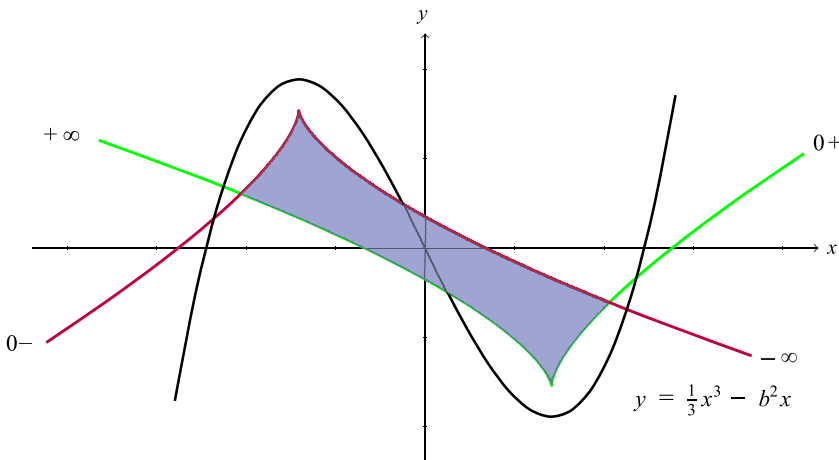
Figure 6   Each point in the shaded region lies on five normal lines.

The first step is to find the corners of the region. We have already determined the location of the cusps but, as we shall see, these are not necessary for the area computation. Let (s, t) be the leftmost point of intersection. By symmetry, the other intersection point is $(-s, -t)$. Since (s, t) lies on both branches of the evolute curve, the polynomial

$$
f(x)f'(x) - tf'(x) + x - s
$$

has two repeated roots, one negative and one positive. As we have done previously, multiplying this polynomial by 3 and calling the resulting polynomial $P(x)$, we find that

$$P(x) = x^5 - 4b^2x^3 - 3tx^2 + 3(b^4 + 1)x + 3(b^2t - s).$$

Let $u < 0 < v$ be the two repeated roots of the polynomial P. Since the sum of the roots of P is 0, we know that

$$P(x) = (x - u)^2(x - v)^2(x + 2u + 2v).$$

In expanded form, this becomes

$$x^5 - (3u^2 + 4uv + 3v^2)x^3 + \left(2(u^3 + v^3) + 8uv(u + v)\right)x^2$$
$$- \left(7u^2v^2 + 4uv(u^2 + v^2)\right)x + 2u^2v^2(u + v).$$

Equating the x^3 and x coefficients, we find that

$$3u^2 + 4uv + 3v^2 = 4b^2$$

and

$$7u^2v^2 + 4uv(u^2 + v^2) = -3(b^4 + 1).$$

Let $\alpha = u^2 + v^2$ and $\beta = uv$. Then $\beta < 0 < \alpha$, $u + v = \sqrt{\alpha + 2\beta}$ (the negative square root corresponds to the other corner point), and the equations become

$$3\alpha + 4\beta = 4b^2$$

and

$$7\beta^2 + 4\alpha\beta = -3(b^4 + 1).$$

The first equation tells us that $\alpha = \frac{4}{3}(b^2 - \beta)$. Since

$$\alpha + 2\beta = (u + v)^2 \geq 0,$$

we find that $\beta \geq -2b^2$. Substituting the expression for α into the other equation yields

$$7\beta^2 + \frac{16}{3}\beta(b^2 - \beta) + 3(b^4 + 1) = 0 \quad \Leftrightarrow \quad 5\beta^2 + 16b^2\beta + 9(b^4 + 1) = 0.$$

We thus find that

$$\beta = \frac{1}{10}\left(-16b^2 \pm \sqrt{256b^4 - 180(b^4 + 1)}\right)$$
$$= \frac{1}{5}\left(-8b^2 \pm \sqrt{19b^4 - 45}\right).$$

Since $b^2 > \sqrt{3}$, the square root generates real numbers. It is easy to verify that both of these roots are negative and that the root corresponding to the plus sign yields a value for β that satisfies the inequality $-2b^2 \leq \beta < 0$. (For values of b that satisfy $\sqrt{45/19} < b^2 < \sqrt{3}$, there is a five normal line region, but it consists of two symmetric parts. In this case, the value for β involving the minus sign comes into play. We leave the determination of the area of the five normal line region for this range of b^2 values as a challenging exercise for the reader.)

Using the plus sign to determine a value for β, we thus have

$$\alpha = \frac{4}{3}(b^2 - \beta) = \frac{4}{15}\left(13b^2 - \sqrt{19b^4 - 45}\right)$$

and

$$\beta = \frac{1}{5}\left(\sqrt{19b^4 - 45} - 8b^2\right).$$

Since

$$v + u = \sqrt{\alpha + 2\beta} \quad \text{and} \quad v - u = \sqrt{\alpha - 2\beta},$$

it follows that

$$u = \frac{1}{2}\left(\sqrt{\alpha + 2\beta} - \sqrt{\alpha - 2\beta}\right)$$

and

$$v = \frac{1}{2}\left(\sqrt{\alpha + 2\beta} + \sqrt{\alpha - 2\beta}\right).$$

The normals to the cubic at $x = u$ and $x = v$ pass through the leftmost corner of the region while the normals to the cubic at $x = -u$ and $x = -v$ pass through the rightmost corner of the region. It is possible to express the coordinates of the corner point (s, t) in terms of α and β (equate the x^2 and constant coefficients in the two polynomials), but these values are rather messy to write out and provide little useful information.

We are now in a position to use Green's theorem to compute the area of the five normal line region. This theorem is a standard result from vector calculus that allows us to compute the area of a region with a line integral. The region of interest for us is traversed in a counterclockwise direction on the boundary curve C in the following way (refer to Figure 6). Start at the leftmost corner and use the positive branch of the evolute for x values from v to $-u$, then switch to the negative branch of the evolute using x values from $-v$ to u. Since $hk' - kh'$ is an odd function, the area A of the region is

$$A = \frac{1}{2}\int_C (x\,dy - y\,dx)$$

$$= \frac{1}{2}\left(\int_v^{-u}\left(h(x)k'(x) - k(x)h'(x)\right)\,dx + \int_{-v}^u\left(h(x)k'(x) - k(x)h'(x)\right)\,dx\right)$$

$$= \frac{1}{2}\left(\int_v^{-u}\left(h(x)k'(x) - k(x)h'(x)\right)\,dx + \int_v^{-u}\left(h(x)k'(x) - k(x)h'(x)\right)\,dx\right)$$

$$= \int_v^{-u}\left(h(x)k'(x) - k(x)h'(x)\right)\,dx.$$

Using the expressions for h, h', k, and k' recorded earlier and omitting the details, we find that

$$h(x)k'(x) - k(x)h'(x) = \frac{5}{6}x^7 - 4b^2x^5$$

$$+ \left(5b^4 + \frac{7}{3}\right)x^3 - \frac{4}{3}\left(b^6 + b^2\right)x - \frac{(b^4 + 1)^2}{2x}.$$

It then follows that

$$A = -\frac{5}{48}(v^8 - u^8) + \frac{2}{3}b^2(v^6 - u^6) - \left(\frac{5}{4}b^4 + \frac{7}{12}\right)(v^4 - u^4)$$

$$+ \frac{2}{3}\left(b^6 + b^2\right)(v^2 - u^2) + \frac{(b^4 + 1)^2}{2}\ln(v/|u|),$$

an admittedly complicated looking expression. Since u and v depend on the parameter $\lambda = b^2$, the area A is a function of λ. It is thus possible to express this value in terms of λ, but it is not all that enlightening to write out this expression.

The graph in Figure 6 corresponds to the curve $y = \frac{1}{3}x^3 - 2x$. The area of the five normal line region in this case (where $b^2 = 2$) is about 2.94, a value that seems reasonable based on the graph. Suppose we want to find the area of the five normal line region for a curve of the form $y = x^3 - cx$. This equation can be written as

$$\sqrt{3}y = \frac{1}{3}\left(\sqrt{3}x\right)^3 - \left(\sqrt{c}\right)^2\left(\sqrt{3}x\right),$$

indicating that the scale factor (discussed earlier in the paper) is $\sqrt{3}$. The desired value for the area is thus $A(c)/3$. For example, the area of the five normal line region for the curve $y = x^3 - 2x$ is about $2.94/3 = 0.98$.

For the record, the exact value of the area is reasonably simple for the cases in which $\lambda = \sqrt{3}$ and $\lambda = \sqrt{15}$. It seems that most other values for λ generate complicated looking expressions for the exact value of $A(\lambda)$, but with some care, it is possible to find other values of λ that generate reasonably nice exact answers. Of course, using a computer to integrate the function $hk' - kh'$ over an appropriate interval can quickly give a decimal approximation for the value of the area.

As a final observation, we consider the asymptotic behavior of $A(b^2)$ by computing $\lim_{b \to \infty} A(b^2)/(b^2)^4$. We begin by noting that

$$\alpha_* \equiv \lim_{b \to \infty} \frac{\alpha}{b^2} = \frac{4}{15}\left(13 - \sqrt{19}\right)$$

$$\beta_* \equiv \lim_{b \to \infty} \frac{\beta}{b^2} = \frac{1}{5}\left(\sqrt{19} - 8\right)$$

$$u_* \equiv \lim_{b \to \infty} \frac{u}{b} = \frac{1}{2}\left(\sqrt{\alpha_* + 2\beta_*} - \sqrt{\alpha_* - 2\beta_*}\right)$$

$$v_* \equiv \lim_{b \to \infty} \frac{v}{b} = \frac{1}{2}\left(\sqrt{\alpha_* + 2\beta_*} + \sqrt{\alpha_* - 2\beta_*}\right).$$

Using these limits, it follows that

$$\lim_{b \to \infty} \frac{A(b^2)}{b^8} = -\frac{5}{48}(v_*^8 - u_*^8) + \frac{2}{3}(v_*^6 - u_*^6)$$

$$-\frac{5}{4}(v_*^4 - u_*^4) + \frac{2}{3}(v_*^2 - u_*^2) + \frac{1}{2}\ln(v_*/|u_*|)$$

$$= \frac{1}{1215}\left(49 - 13\sqrt{19}\right)\sqrt{5 + 40\sqrt{19}}$$

$$+ \frac{1}{4}\ln\left(\frac{26 - 2\sqrt{19} + \sqrt{5 + 40\sqrt{19}}}{26 - 2\sqrt{19} - \sqrt{5 + 40\sqrt{19}}}\right)$$

$$\approx 0.431773.$$

We have thus shown that it is possible to express the area of the five normal line region, as well as the limiting value of this area, in fairly simple analytic terms. Although these analytic forms may not provide any additional insight, it is helpful to know that such expressions are possible.

We have carefully examined many of the properties of the evolutes of cubic polynomials. For polynomials of higher degree, the number of branches of the evolutes increases and the regions representing portions of the plane with a given number of normal lines become much more complicated (see Sanchez [**9**] for some graphs). It is also necessary to introduce more parameters in order to represent all of the possible forms. Determining parameter values for certain behaviors of the evolute curves and finding the areas of the various bounded regions in these cases is much more difficult and, most likely, it is not possible to represent these areas in analytic form.

REFERENCES

[1] Bains, M., Thoo. J. (2007). The normals to a parabola and the real roots of a cubic. *College Math. J.* 38(4): 272–277. doi.org/10.1080/07468342.2007.11922248

[2] Burgette, L., Gordon, R. (2004). On determining the non-circularity of a plane curve. *College Math. J.* 35(2): 74–83. doi.org/10.1080/07468342.2004.11922055

[3] Coolidge, J. (1952). The unsatisfactory story of curvature. *Amer. Math. Monthly.* 59(6): 375–379. doi.org/10.2307/2306807

[4] Eves, H. (1976). *An Introduction to the History of Mathematics*, 4th ed. New York: Holt, Rinehart, and Winston.

[5] Giv, H. (2017). The osculating circle without the unit normal vector. *Math. Mag.* 90(5): 347–352. doi.org/10.4169/math.mag.90.5.347

[6] Hartmann, F., Jantzen, R. (2010). Apollonius's ellipse and evolute revisited—The alternative approach to the evolute. *Convergence* (online). maa.org/press/periodicals/convergence/apolloniuss-ellipse-and-evolute-revisited.

[7] Lawrence, J. (1972). *A Catalog of Special Plane Curves*. New York: Dover.

[8] Leithold, L. (1981). *The Calculus with Analytic Geometry*, 4th ed. New York: Harper and Row.

[9] Sanchez, D., Smith, K. (2000). Normal lines and the evolute curve. *College Math. J.* 31(5): 397–403. doi.org/10.2307/2687456

Summary. We consider the evolutes of cubic polynomials $y = \frac{1}{3}x^3 - \lambda x$, where λ is a positive constant, and explore some of their properties. In particular, we determine values for the parameter λ for which the evolute has certain characteristics.

RUSSELL A. GORDON (MR Author ID: 75470) received his Ph.D. from the University of Illinois in 1987, writing his dissertation under the supervision of Jerry Uhl. He has been teaching mathematics at Whitman College since then and is becoming increasingly aware that his current students believe that 1987 was a long time ago. When not pursuing various mathematical ideas, he enjoys eating his spouse's wonderful vegetarian cooking (for which doing the dishes is a small price to pay), watching movies with his family, and hiking in the local mountains.

A Curious Possible Prime Pattern

BENJAMIN M. BOLKER
McMaster University
Hamilton, Ontario, Canada
bolker@mcmaster.ca

ELEANOR J. B. BOLKER
Phillips Exeter Academy
Exeter, NH 03833
ejbbolker@comcast.net

ETHAN D. BOLKER
University of Massachusetts, Boston
Boston, MA 02125
ebolker@gmail.com

The second author is particularly fond of the prime 29 (it's her birthday). When she was 12 she observed that

$$29_{29} = (2 \times 29) + 9 = 67$$

is also prime. So she continued:

$$67_{67} = (6 \times 67) + 7 = 409$$
$$409_{409} = 669133$$

are prime too. The next round of arithmetic was too much for her, but Python and Mathematica say that, sadly,

$$669133_{669133} = 804848507724119045807275881508$$
$$= 2^2 \times 10143450113 \times 19836655643739326729$$

is composite.[*]

Patterns that depend on representing integers in a particular base are often anomalies.[**] To connect her observations to some mainstream ideas in number theory we consider her construction for arbitrary bases. For positive integers b and $n > b$ let $L_b(n)$ be the base n expansion of the base b string representation of n. When both n and $L_b(n)$ are prime we call $(n, L_b(n))$ a *base b Eleanor prime pair*. Her first observation was that $(29, 67)$ is such a pair for base 10. $(23, 29)$ is a base 17 Eleanor prime pair because $23 = 16_{17}$ and

$$L_{17}(23) = 1 \times 16_{17} + 6 = 23 + 6 = 29$$

is prime.

We wrote a Python program to look for Eleanor prime pairs. Here are all the chains of length at least four starting with primes $p \leq 1069$ for all bases up to 30:

---

[*]In retrospect we could have known that number is even because 669133 has an even number of odd digits.

[**]One reviewer called them "the kind of thing that gives number theory a bad name."

```
base 10:   29,    67,    409,      669133
base 11:   17,    23,     47,         191
base 14:   29,    59,    239,       57839
base 15:   17,    19,     23,          31
base 17:   23,    29,     41,          89,   449
base 19;   59,   179,   1619,   10499219
base 25;   67,   151,    907,      832633
base 26:   53,   107,    431,        6911
base 28:   61,   127,    523,        9433
```

The base 15 line in this table with gaps 2 : 2, 4, 8 led us to

Lemma 1. $b < n < 2b$ *implies* $L_b(n) = 2n - b$ *and thus* $L_b(n) - n = n - b$.

Proof. The b's place digit in the two digit base b expansion of n is 1 so n_b is the string $1(n - b)$. Then

$$L_b(n) = n + (n - b) = 2n - b. \qquad \blacksquare$$

Corollary 2. *If p and q are primes with $p < q < 3p/2$ then (p, q) is a base $2p - q$ Eleanor prime pair.*

Proof. Since $2(2p - q) = 4p - 2q > 4p - 3p = p$, Lemma 1 applies:

$$L_{2p-q}(p) = 2p - (2p - q) = q. \qquad \blacksquare$$

This explains (in a sense) the base 17 Eleanor prime pairs (23, 29) and (29, 41) but not the rest of that row in the table.

In particular, whenever $(p, p + 2)$ is a pair of twin primes it's a base $p - 2$ Eleanor prime pair. We do not know whether there are infinitely many twin primes, but we do know enough to prove

Theorem 3. *Every sufficiently large prime is the first prime in an Eleanor prime pair for some base.*

Proof. Hardy and Wright [2] used the prime number theorem to prove that for any $\varepsilon > 0$ there is an $N > 0$ such that for all $n > N$ there is a prime q such that $n < q < (1 + \varepsilon)n$. Let $\varepsilon = 1/2$ and suppose $p > N$ is prime. Then there is a prime $q < 3p/2$. Corollary 2 implies (p, q) is a base $2p - q$ Eleanor prime pair. $\qquad \blacksquare$

What about chains of Eleanor prime pairs? Write $L_b^2(n)$ for the composition $L_b(L_b(n))$. Then

Lemma 4. $b < n < L_b(n) < 2b$ *implies*

$$L_b^2(n) - L_b(n) = 2(L_b(n) - n).$$

Proof.

$$
\begin{aligned}
L_b^2(n) - L_b(n) &= L_b(L_b(n)) - L_b(n) \\
&= 2L_b(n) - b - L_b(n) \\
&= L_b(n) - b \\
&= L_b(n) - (2n - L_b(n)) = 2(L_b(n) - n).
\end{aligned}
$$

$\blacksquare$

As long as $b < L_b^j(n) < 2b$, applying L doubles the gap. That explains the base 15 Eleanor prime pair chain 17, 19, 23, 31 with gaps 2, 4, 8. Since 31 is more than twice the base, applying L again no longer doubles the gap. But if we just continue to double the gap, the primes continue as far as

$$17, 19, 23, 31, 47, 79.$$

Primality fails at the next entry: $79 + 64 = 143 = 11 \times 13$.

In 1904 Leonard Dickson conjectured that any finite integer sequence defined by linear functions represents only primes infinitely often unless it obviously does not [**1**]. Dickson's conjecture is quite powerful. It implies some known results like the infinitude of the primes and Dirichlet's theorem on primes in arithmetic progressions and many deeper unproved ones like the twin prime conjecture.

The sequence $n, n + 2, n + 4$ obviously fails to produce primes after $(3, 5, 7)$ since it always contains a multiple of 3, but there's no problem with $n, n + 2, n + 6$, or, equivalently, $n + 2, n + 4, n + 8$.

Theorem 5. *Dickson's conjecture implies that for each k there are infinitely many bases for which there is a chain of k Eleanor prime pairs beginning with a pair of twin primes.*

Proof. Consider the sequence

$$n + 2, n + 4, \ldots, n + 2^{k+1}$$

in which the gap starts at 2 and doubles at each step.

The obvious obstruction to applying Dickson's conjecture is the existence of an odd prime p such that for every n at least one number in the sequence is a multiple of p. There is no such p here: consider what happens when p divides n. Then each of the numbers in the sequence is congruent (mod p) to a power of 2 and hence not 0 (mod p).

Then Dickson's conjecture implies that for each k there are infinitely many such sequences all of whose entries are prime. Each one with $n > 2^{k+1}$ is a length k chain of base n Eleanor prime pairs starting with a pair of twin primes.   ∎

Our original proof of Theorem 3 relied on Zhang's bounded gap theorem [**3**]: if a difference $q - p$ between primes occurs infinitely often then eventually $q < 2(2p - q)$ and we have Eleanor prime pairs. If there were an analogue of Zhang's theorem bounding Dickson-like sequences the infinitude of long chains of Eleanor prime pairs would follow without the need for Dickson.

All our results so far turn the problem around, searching for bases that have Eleanor prime pairs. We find them because Lemma 1 tells us how $L_b(n)$ behaves for two digit base b integers n with leading digit 1. What we'd really like are Eleanor prime pair existence theorems for arbitrary bases. That would require understanding $L_b(n)$ when $n \gg b$. Then n has about $\log_b(n)$ base b digits and

$$\log_b(L_b(n)) \approx (\log_b(n))^2$$

so $L_b(n)$ is much larger than n. The gaps in a potential chain of base b Eleanor prime pairs starting at n grow very rapidly as a function of n. Our proof of Theorem 5 using Dickson will not work.

We leave these questions to the reader (in increasing order of difficulty):

- Are there Eleanor prime pairs for every base?
- Are there infinitely many Eleanor prime pairs for every base?

- Are there infinitely many arbitrarily long chains of Eleanor prime pairs for every base?

We conjecture "yes" for all three, with decreasing confidence. Until you find a base 10 chain of length 5, Eleanor's birthday holds the everyday record.

REFERENCES

[1] Dickson, L. E. (1904). A new extension of Dirichlet's theorem on prime numbers. *Messenger Math.* 33: 155–161.
[2] Hardy, G. H., Wright, E. M. (2008). *An Introduction to the Theory of Numbers*, 6th ed. Oxford: Oxford University Press, p. 494.
[3] Zhang, Y. (2014). Bounded gaps between primes. *Ann. Math.* 179(3): 1121–1174. doi.org/10.4007/annals.2014.179.3.7.

Summary.   29 and $29_{29} = 67$ are both prime. We explore that curiosity and find connections to deep questions on the distribution of the primes: the prime number theorem, Dickson's conjecture, and Zhang's bounded prime gap theorem.

BENJAMIN M. BOLKER, Eleanor's uncle, is a professor at McMaster University. When he is not doing research and teaching in ecology, epidemiology, or statistics, or writing R code, he folk dances and plays the fiddle.

ELEANOR J. B. BOLKER is a student at Phillips Exeter Academy. She is an avid NPR listener and reader of magical realism. Her math classmates sometimes complain that whenever she asks a question they get more homework.

ETHAN D. BOLKER is Professor of Mathematics, Emeritus, at UMass Boston. His favorite research topics come from elementary mathematics and ideas he first encountered in Hugo Steinhaus' *Mathematical Snapshots* when he was about his granddaughter Eleanor's age.

Sierpiński's Topological Characterization of $\mathbb{Q}$

KRZYSZTOF CHRIS CIESIELSKI
West Virginia University
Morgantown, WV 26506-6310
KCies@math.wvu.edu

In a 1920 paper [5], Wacław Sierpiński proved the following result characterizing the space $\mathbb{Q}$ of rational numbers considered with the standard topology:

Theorem 1. *Any countable metric space $\langle X, d \rangle$ without isolated points is homeomorphic to $\mathbb{Q}$.*

Its simple and natural form seems to indicate that its proof could be included in an undergraduate topology course as soon as the notion of "homeomorphism" is introduced. This result can help to illuminate the difference between the standard topologies on $\mathbb{R}$ and on $\mathbb{Q}$. According to Theorem 1, $\mathbb{Q}$ is homeomorphic to $\mathbb{Q}^2$ and to $\mathbb{Q}_\ell$ (that is, $\mathbb{Q}$ with the "Sorgenfrey topology," generated by all left closed intervals $[p, q)$). In contrast, their real counterparts $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}_\ell$—obtained by replacing $\mathbb{Q}$ with $\mathbb{R}$ in their respective definitions—are pairwise topologically different. Also, Theorem 1 implies that the Furstenberg topology on the integers $\mathbb{Z}$, which has been used to prove the infinitude of primes (see Aigner [1] or Furstenberg [4]), is actually homeomorphic to the standard topology on $\mathbb{Q}$.

Nevertheless, so far Theorem 1 could not have been included early in topological education for a simple reason—all of the published proofs are too complicated for such a purpose. The proofs presented in [3] and [2] are relatively simple, but they are both considerably longer than our proof, and they are not self contained, since they both rely on Cantor's characterization of the linear structure of $\mathbb{Q}$: *Any linearly ordered dense set with neither smallest nor greatest element is order-isomorphic to $(\mathbb{Q}, \leq)$.*

We now prove Theorem 1.

Proof. Let S be the set of all infinite sequences $s = \langle s_1, s_2, \ldots \rangle$ of natural numbers that are eventually zero, that is, such that $0 = s_n = s_{n+1} = \cdots$ for some $n \in \mathbb{N} := \{1, 2, 3, \ldots\}$. Notice that S is countable, and so is the set $\mathbb{N}^{<\omega}$ of all finite sequences of natural numbers. Equip S with a topology τ generated by a basis formed by all sets $[t]$, with $t \in \mathbb{N}^{<\omega}$, defined as

$$[t] := \{s \in S : t \subset s\},$$

where "$t \subset s$" means "the sequence s extends t."[*] To finish the proof it is enough to show that there is a homeomorphism $h \colon X \to S$, since then there also exists a homeomorphism $H \colon \mathbb{Q} \to S$, implying that $H^{-1} \circ h \colon X \to \mathbb{Q}$ is a homeomorphism and proving Theorem 1.

Let $\{x_n : n < \omega\}$ be an enumeration, with no repetitions, of the set X and let

$$D := \{d(x, y) : x, y \in X\}$$

be the set of all distances between the elements in X. Notice that D is countable. Also, for any $r > 0$ with $r \notin D$, the open ball

$$B_d(c, r) := \{x \in X : d(c, x) < r\}$$

[*]Of course, $\langle S, \tau \rangle$ is actually a subspace of $\mathbb{N}^{\mathbb{N}}$ considered with the product topology.

centered in $c \in X$ and with radius r is also a closed set in X, since it is equal to

$$B_d[c, r] := \{x \in X : d(c, x) \le r\}.$$

All open balls in X considered below will be with radii *not* in D. So, they will be clopen sets.

In the construction of h we will repeatedly use the following simple fact:

(*) For every $k \in \mathbb{N}$ and nonempty open subset U of X there exists a sequence

$$S_k(U) := \langle B_i : i \in \mathbb{N} \rangle$$

of pairwise disjoint clopen balls contained in U, each of radius $\le 2^{-k}$, such that $U = \bigcup_{i \in \mathbb{N}} B_i$. Moreover, we assume that B_0 contains $x_{n(U)}$, where

$$n(U) := \min\{i < \omega : x_i \in U\}.$$

The balls are chosen by induction on $i \in \mathbb{N}$: each B_i is centered at the point $x_{n(U_i)}$, where

$$U_i := U \setminus \bigcup_{j < i} B_j,$$

and has radius $r_i \in (0, 2^{-k}) \setminus D$ small enough so that

$$B_i \subsetneq U \setminus \bigcup_{j < i} B_j.$$

More specifically, each U_i is open, as a difference of open U and a finite union of balls, each of which is a closed set according to our choice of balls of radii not in D. Each nonempty U_{i-1}, including $U_0 = U$, has more than one point, since X has no isolated points. This allows us to choose each r_i small enough so that $U_i = U_{i-1} \setminus B_i$ is nonempty, as long as $U_{i-1} \ne \emptyset$. Finally, each $x_k \in U$ belongs to $\bigcup_{j \le i} B_j$, according to our rule of choosing the center of each B_i with the smallest possible available index.

Next, we construct the family

$$\{B_s : s \in \mathbb{N}^{<\omega}\}$$

of nonempty clopen sets in X. The construction is by induction on the length of sequences s. Thus, for the sequence $\emptyset$ of length 0 we put $B_\emptyset := X$. Also, if $\mathbb{N}^k$ is the set of all sequences in $\mathbb{N}^{<\omega}$ of length k (possibly 0) and for every $s \in \mathbb{N}^k$ and $i \in \mathbb{N}$ the symbol $s\hat{\ }i$ denotes the sequence s extended by one more term with value i, then we define

$$\langle B_{s\hat{\ }i} : i \in \mathbb{N} \rangle := S_k(B_s).$$

Notice that

for every $x \in X$ and $k \in \mathbb{N}$ there exists a unique $s \in \mathbb{N}^k$ so that $x \in B_s$.   (1)

This is justified by an easy inductive argument. For $k = 1$ this holds, since the sets $\{B_{\emptyset\hat{\ }i} : i \in \mathbb{N}\}$ form a partition of $B_\emptyset = X$. Also, if $x \in B_s$ for some $s \in \mathbb{N}^k$, then there is a unique $t \in \mathbb{N}^{k+1}$, which must be of the form $t = s\hat{\ }i$, for which $x \in B_{s\hat{\ }i}$, as the sets $\{B_{s\hat{\ }i} : i \in \mathbb{N}\}$ form a partition of B_s.

Notice that, by (1), for every $x \in X$ there is a unique sequence $s = s_x \in \mathbb{N}^{\mathbb{N}}$ such that

$$x \in \bigcap_{k \in \mathbb{N}} B_{s \restriction k},$$

where $s \restriction k$ is the restriction of s to its first k elements. Define $h \colon X \to \mathbb{N}^{\mathbb{N}}$ by letting $h(x) := s_x$ for every $x \in X$. We claim that this is the desired homeomorphism.

Clearly h is one-to-one. To see that h is onto S, first notice that $h[X] \subset S$. Indeed, by $(*)$, for every $x_j \in X$ we have $h(x_j) \restriction k = 0$ for every $k > n$, where n is such that

$$d(x_i, x_j) > 2^{-n},$$

for every $i < j$. Thus, $h(x_j) \in S$, since $h(x_j)$ is eventually 0. Also, h is onto S, since for every $s \in S$ there exists a $k \in \mathbb{N}$ such that $s_m = 0$ for all $m \geq k$. Let $j = n(B_{s \restriction k})$. Then, by $(*)$, $x_j \in B_{s \restriction m}$ for all $m \geq k$ and so $h(x_j) = s$.

Finally, we need to show that both h and h^{-1} are continuous. Clearly, h is continuous, since for every basic open set $[s]$ in S, $s \in \mathbb{N}^{<\omega}$, we have

$$h^{-1}([s]) = B_s$$

is open in X. Also, h^{-1} is continuous, since

$$\{B_s : s \in \mathbb{N}^{<\omega}\}$$

is a basis for X and, for every $s \in \mathbb{N}^{<\omega}$,

$$(h^{-1})^{-1}(B_s) = h(B_s) = [s]$$

is open in S.

∎

REFERENCES

[1] Aigner, M., Ziegler, G. (2001). *Proofs from THE BOOK*. Heidelberg: Springer-Verlag.
[2] Błaszczyk, A. (forthcoming). A simple proof of Sierpiński's theorem. To appear in the *Amer. Math. Monthly*.
[3] Eberhart, C. (1977). Some remarks on the irrational and rational numbers. *Amer. Math. Monthly*. 84(1): 32–35. doi.org/10.2307/2318303
[4] Furstenberg, H. (1955). On the infinitude of primes. *Amer. Math. Monthly*. 62(5): 353. doi.org/10.2307/2307043
[5] Sierpiński, W. (1920). Sur une propriété topologique des ensembles dénombrables denses en soi. *Fundam. Math.* 1: 11–16.

Summary.   In a 1920 paper, Sierpiński proved the following theorem characterizing the space $\mathbb{Q}$ of rational numbers considered with the standard topology: *Any countable metric space $\langle X, d \rangle$ without isolated points is homeomorphic to $\mathbb{Q}$*. In this note, we provide a simple proof of this result, that requires only basic topological background. As such, it can be incorporated into an undergraduate topology curriculum.

KRZYSZTOF CHRIS CIESIELSKI (MR Author ID: 49415) received his Masters and Ph.D. degrees in Pure Mathematics from Warsaw University, Poland, in 1981 and 1985, respectively. He has worked at West Virginia University since 1989. Since 2006, he has been an adjunct professor in the Department of Radiology at the University Pennsylvania. He is the author of three books and over 140 journal research articles. He is an editor of *Real Analysis Exchange*, *Journal of Applied Analysis*, and *Journal of Mathematical Imaging and Vision*.

How Fundamental is the Fundamental Theorem of Algebra?

STEVEN KRANTZ
Washington University in St. Louis
St. Louis, MO 63130
sk@math.wustl.edu

Mathematicians spent several centuries studying and finding the roots of polynomials. It was Carl Friedrich Gauss who finally proved the fundamental theorem of algebra—that any polynomial of degree $k \geq 1$ has k complex roots (counting multiplicities). Over the course of his career, he produced a total of five proofs of this decisive result.

Today there are a great many proofs of the fundamental theorem of algebra. Here we present several rather unusual proofs. These come from many different parts of mathematics, including topology, complex analysis, and algebra. Useful and entertaining references for these matters include the books by Fine and Rosenberger [1], Greene and Krantz [2], Narasimhan and Nievergelt [3], and Saff and Snider [6], as well as the online articles at PlanetMath [5] and Wikipedia [7].

We concentrate here on proving that any polynomial of degree $k \geq 1$ has *one* root. The full result then follows from the Euclidean algorithm.

It is a pleasure to thank the referees and the editor for a number of perceptive remarks that helped to clarify the exposition.

Topology

Proof. Let p be a polynomial of degree $k \geq 1$. Write

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k.$$

If $a_0 = 0$, then obviously $z = 0$ is a root of the polynomial. So we may assume that $a_0 \neq 0$.

Let $\epsilon > 0$ be very small (much smaller than $a_0 \neq 0$). Then the image of the circle

$$C(0, \epsilon) = \{z \in \mathbb{C} : |z - 0| = \epsilon\}$$

under p is a small perturbation of the constant function $q(z) \equiv a_0$. Specifically, if $z = \epsilon e^{i\theta}, 0 < \theta \leq 2\pi$, then

$$p(z) = a_0 + a_1 \epsilon e^{i\theta} + a_2 \epsilon^2 e^{2i\theta} + \cdots + a_k \epsilon^k e^{ki\theta} = a_0 + \epsilon \cdot (\mathcal{R}(z)),$$

where $\mathcal{R}$ is a remainder term that is bounded. So the image will be a closed curve that *does not* surround the origin (because $a_0 \neq 0$ and $\epsilon \ll |a_0|$).

Now let $R > 0$ be very large. Then the image of the circle

$$C(0, R) = \{z \in \mathcal{C} : |z - 0| = R\}$$

under p will be approximately the image of $C(0, R)$ under $q(z) = a_k z^k$. (Specifically, we can write

$$p(z) = R^k \cdot \left(\frac{a_0}{R^k} + \frac{a_1 z}{R^k} + \cdot + \frac{a_k z^k}{R^k} \right).$$

If R is very large, then the fraction z^j/R^k is small for $|z| = R$ and $0 \leq j < k$. Thus, the dominant term inside the parentheses is $a_k z^k/R^k$.) That will, of course, be a circle centered at the origin with radius $|a_k|R^k$. In particular, this image *will* encircle the origin.

Now we reason that, as the radius of a circle varies continuously from ϵ to R, the images vary continuously (use the implicit function theorem for a rigorous verification of this statement). Since, for radius ϵ, the image does *not* encircle the origin, while for radius R the image *does* encircle the origin, we may conclude that for some intermediate radius the image *passes through* the origin. Thus, p takes the value 0. ∎

Hurwitz's theorem

Proof. Let p be a polynomial of degree $k \geq 1$. Write

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k .$$

Now consider the sequence of functions

$$f_j(z) \equiv j^{-k} p(jz) .$$

It is clear that the sequence f_j converges uniformly on compact sets to the function $f_0 = a_k z^k$. If p is nonvanishing, then each f_j is nonvanishing. Hurwitz's theorem (see Greene and Krantz [2]) then tells us that the limit is either nonvanishing or identically 0. But f_0 satisfies neither of these properties, which is a contradiction. So p must have a zero. ∎

Rouché's theorem

Proof. Let p be a polynomial of degree $k \geq 1$. Write

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k .$$

Of course, we assume that $a_k \neq 0$.

Define $g(z) = a_k z^k$. Then, for R large, it is clear that

$$|p(z) - g(z)| < |g(z)|$$

on $\partial D(0, R)$. In particular, if $z = Re^{i\theta}, 0 < \theta \leq 2\pi$, then

$$|p(z) - g(z)| = \left| a_0 + a_1 Re^{i\theta} + a_2 R^2 e^{2i\theta} + \cdots + a_{k-1} R^{k-1} e^{(k-1)i\theta} \right|$$

$$= R^{k-1} \cdot \left(\frac{a_0}{R^{k-1}} + \frac{a_1 e^{i\theta}}{R^{k-2}} + \frac{a_2 e^{2i\theta}}{R^{k-3}} + \cdots + a_{k-1} e^{(k-1)i\theta} \right) .$$

When R is large, all the terms in the parentheses on the right are small, with the exception of the last term. So the expression on the right is essentially of size $R^{k-1}|a_{k-1}|$, and that is smaller than $|g(z)| = |a_k|R^k$ if R is large enough.

By Rouché's theorem, p and g have the same number of zeros in $D(0, R)$. Hence p has k zeros. ∎

Counting zeros

Proof. Let p be a polynomial of degree $k \geq 1$. Write

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k .$$

If f is a function holomorphic on a neighborhood of a closed disc

$$\overline{D}(0, r) = \{z \in \mathcal{C} : |z - 0| \leq r\}$$

and nonvanishing on the boundary, then it is a well-known result (see Greene and Krantz [2]) that the number of zeros of f in the interior of the disc is given by

$$\mathcal{N}(f, r) = \frac{1}{2\pi i} \oint_{\partial D(0,r)} \frac{f'(z)}{f(z)} \, dz.$$

This formula follows from the Cauchy theory—first verify it for monomials and then approximate.

If p is zero-free, then $\mathcal{N}(p, r) = 0$ for every $r > 0$. Now let $r \to \infty$ Then we may calculate that

$$\mathcal{N}(p, r) = \frac{1}{2\pi i} \oint_{\partial D(0,r)} \frac{p'(z)}{p(z)} \, dz$$

is asymptotically equal to

$$\frac{1}{2\pi i} \oint_{\partial D(0,r)} \frac{k a_k z^{k-1}}{a_k z^k} \, dz = \frac{1}{2\pi i} \oint_{\partial D(0,r)} \frac{k}{z} \, dz = k \neq 0 .$$

This is a contradiction.   ∎

More topology

Proof. Let p be a polynomial of degree $k \geq 1$. Write

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k .$$

We shall prove that the image of p is both open and closed. Since it is clearly nonempty, this will imply it must be all of $\mathcal{C}$.

Open: Since the function p is clearly holomorphic, it is an open mapping. Therefore the image of p is open.

Closed: Let q_j be a Cauchy sequence of elements in the image of p. Then $\{q_j\}$ is certainly a bounded set. If r_j is a preimage of q_j for each j, then $\{r_j\}$ must be a bounded set (for certainly $p(z)$ blows up as z tends to infinity). Thus, there is a convergent subsequence $\{r_{j_k}\}$ with limit r_0. But then, by the continuity of the function p, the point $q_0 = p(r_0)$ is the limit of the sequence $\{q_{j_k}\}$. Since the sequence $\{q_j\}$ is Cauchy, we see that the image point q_0 is the limit of the full sequence $\{q_j\}$. So the image of p is closed.

Since the image of p is open, closed and nonempty (after all, $a_0 = p(0)$ is in the image of p), we may conclude that the image of p is all of $\mathcal{C}$.   ∎

Concluding remarks

Peter Roth, in his book *Arithmetica Philosophica*, published in 1608, formulated a version of the fundamental theorem of algebra. Although polynomial equations had been studied for centuries, this was one of the first formal enunciations of the great theorem. Gauss validated the theorem, and the use of the complex numbers, in his 1798 Ph.D. thesis. Since that time there have been many studies of the fundamental theorem and many new proofs. The book by Krantz and Parks [**4**] even contains a new proof using an unusual version of the implicit function theorem.

One of the strengths of mathematics is that even ancient ideas are still valid today. Another is that the ideas keep developing. It is a pleasure to be a part of that process.

REFERENCES

[1] Fine, B., Rosenberger, G. (1997). *The Fundamental Theorem of Algebra*. New York: Springer.
[2] Greene, R. E., Krantz, S. G. (2006). *Function Theory of One Complex Variable*, 3rd ed. Providence, RI: American Mathematical Society.
[3] Narasimhan, R., Nievergelt, Y. (2000). *Complex Analysis in One Variable*, 2nd ed. Boston, MA: Birkhäuser.
[4] Krantz, S. G., Parks, H. R. (2002). *The Implicit Function Theorem*. Boston, MA: Birkhäuser.
[5] PlanetMath. (2013). Proof of fundamental theorem of algebra. planetmath.org/ProofOfFundamentalTheoremOfAlgebraRouchesTheorem.
[6] Saff, E. B., Snider, A. D. (1993). *Fundamentals of Complex Analysis for Mathematics, Science, and Engineering*, 2nd ed. New York: Prentice Hall.
[7] Wikipedia. (2019). Fundamental theorem of algebra. wikipedia.org/wiki/Fundamental_theorem_of_algebra.

Summary.   We give several non-obvious proofs of the fundamental theorem of algebra, coming from different parts of mathematics. These include the idea of homotopy, Hurwitz's theorem, Rouché's theorem, the open mapping principle, and the argument principle.

STEVEN KRANTZ is a professor of mathematics at Washington University in St. Louis. Born in San Francisco, he graduated *summa cum laude* from the University of California at Santa Cruz in 1971 and earned his Ph.D. in 1974 from Princeton University under the direction of E. M. Stein.

Krantz has held professorships at the University of California in Los Angeles, Penn State University, and more recently at Washington University in St. Louis (where he has been for nearly 34 years). Krantz has directed 20 Ph.D. students and 9 Masters students. He has written 250 scholarly papers and 135 books. He is the holder of the Chauvenet Prize and the Beckenbach Book Award.

Silos, Snow Cones, Railroad Spikes, and Spherical Coordinates

JOHN ENGBERS
Marquette University
Milwaukee, WI 53233
john.engbers@marquette.edu

ADAM HAMMETT
Cedarville University
Cedarville, OH 45314
ahammett@cedarville.edu

In the typical multivariate calculus development, the spherical coordinate transformation

$$x = \rho \sin\phi \cos\theta, \quad y = \rho \sin\phi \sin\theta, \quad z = \rho \cos\phi,$$

is introduced to simplify computations involving certain foundational solids, especially portions of spheres and cones. Of primary concern is to reduce the complexity of certain triple integrals, and ultimately we arrive at the formula

$$\iiint_{\mathcal{D}} f(x, y, z)\, dx\, dy\, dz$$

$$= \iiint_{\mathcal{D}} f(\rho \sin\phi \cos\theta,\ \rho \sin\phi \sin\theta,\ \rho \cos\phi)\, \rho^2 \sin\phi\, d\rho\, d\theta\, d\phi. \quad (1)$$

Although the factor $\rho^2 \sin\phi$ may be found using the Jacobian (see, e.g., Stewart [2, pp. 1046–1047]), most calculus texts cover integration using spherical coordinates *before* they cover the Jacobian. Thus, in the derivation of this formula much attention is given to explaining this factor. The common approach essentially amounts to computing the volume of a "spherical box" in these new coordinates (see the last solid in Figure 1; we will revisit this factor later). More precisely, given fixed values
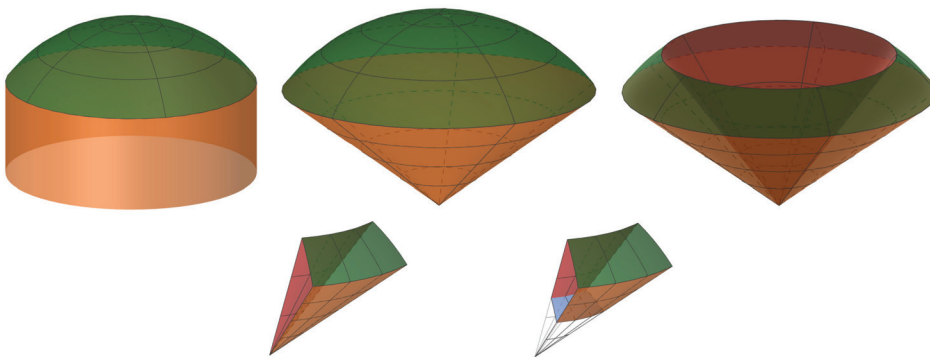

Figure 1 Progressing from "silo," to "snow cone," to "snow cone cup," to "railroad spike," and finally to "spherical box."

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

$$0 \leq \rho_1 \leq \rho_2, \quad 0 \leq \theta_1 \leq \theta_2 \leq 2\pi, \quad \text{and} \quad 0 \leq \phi_1 \leq \phi_2 \leq \pi,$$

let

$$\mathcal{B} = \{(\rho, \theta, \phi) \mid \rho_1 \leq \rho \leq \rho_2, \, \theta_1 \leq \theta \leq \theta_2, \, \phi_1 \leq \phi \leq \phi_2\}$$

denote a "box" in spherical coordinates, and let $V_{\mathcal{B}}$ denote the volume of $\mathcal{B}$. Of course, by *utilizing* the integral formula (1) we have

$$V_{\mathcal{B}} = \frac{1}{3} \left(\rho_2^3 - \rho_1^3 \right) (\theta_2 - \theta_1) \left(\cos \phi_1 - \cos \phi_2 \right), \tag{2}$$

but is there a relatively simple way to arrive at this formula for $V_{\mathcal{B}}$ without appealing to (1)? (Note that formula (2) is the result of computing

$$\int_{\phi_1}^{\phi_2} \int_{\theta_1}^{\theta_2} \int_{\rho_1}^{\rho_2} \rho^2 \sin \phi \, d\rho \, d\theta \, d\phi.)$$

The elegant form of (2) suggests that there ought to be. We remark that there is a basically canonical presentation (see, e.g., Edwards and Larson [**1**] or Stewart [**2**]) whereby an *approximation* of $V_{\mathcal{B}}$ is computed, and this is enough to explain (1) via Riemann sums. We seek an explanation of (2) (independent of (1)) that does *not* rely on estimates, something that is noticeably absent in these standard treatments.

We shall outline a proof of (2) that we have observed to be quite illuminating in the classroom setting, and which we have not seen in any textbook. It has the added benefit of revisiting double integrals in polar coordinates and some elementary facts from solid geometry. Moreover, we shall see that the precise expression for $V_{\mathcal{B}}$, once we pass to differentials, gives rise to an interesting alternative explanation for the appended factor $\rho^2 \sin \phi$ in (1) coming from this triple integral transformation.

We organize our approach in several steps, each of which could serve as a short interactive classroom exercise. Our progression of ideas is illustrated in Figure 1, and in addition to $V_{\mathcal{B}}$ our approach will yield exact volume formulas for *all* the solids there!

Step 1: Volume of a "silo"

Informally, a "silo" $\mathcal{S}$ is a right circular cylinder with base in the xy-plane, centered about the z-axis, and with a spherical top; see the first solid in Figure 1. Formally, we let the top of $\mathcal{S}$ be defined by the spherical equation $\rho = \rho_1$, and we let $\phi_1 \in [0, \pi/2]$ denote the angle from the positive z-axis at which a ray emanating from the origin intersects the circle where the spherical top and cylinder meet. Under these assumptions, the height of just the cylindrical portion within the silo is given by $\rho_1 \cos \phi_1$ and its radius by $\rho_1 \sin \phi_1$. Notice also that by the Pythagorean theorem, the height of the silo over a point at distance $r \leq \rho_1 \sin \phi_1$ from the origin in the xy-plane is given by $\sqrt{\rho_1^2 - r^2}$. Thus, integrating over the base of the cylinder in polar coordinates, the volume of the silo is

$$V_{\mathcal{S}} = \int_0^{2\pi} \int_0^{\rho_1 \sin \phi_1} \sqrt{\rho_1^2 - r^2} \, r \, dr \, d\theta = \frac{2\pi}{3} \rho_1^3 \left(1 - \cos^3 \phi_1 \right). \tag{3}$$

Step 2: Volume of a "snow cone"

A "snow cone" $\mathcal{SC}$ is a right circular cone with tip at the origin and a spherical top added to the base of the cone; see the second solid in Figure 1. We imagine the snow

cone being obtained by shaving off parts of the cylinder within the silo from step 1. More precisely, in the last step we saw that this cylinder has height $\rho_1 \cos \phi_1$ and radius $\rho_1 \sin \phi_1$, and hence it has volume

$$\pi (\rho_1 \sin \phi_1)^2 \rho_1 \cos \phi_1.$$

The conic bottom of the snow cone is obtained by keeping precisely 1/3 the volume of this cylinder. This means that to calculate the volume of a snow cone, we simply need to *subtract* 2/3 of the volume of the cylinder included in (3):

$$V_{SC} = \frac{2\pi}{3} \rho_1^3 \left(1 - \cos^3 \phi_1\right) - \frac{2}{3}\pi (\rho_1 \sin \phi_1)^2 \rho_1 \cos \phi_1$$

$$= \frac{2\pi}{3} \rho_1^3 \left(1 - \cos \phi_1\right). \tag{4}$$

Step 3: Volume of a "snow cone cup"

A "snow cone cup" $\mathcal{SCC}$ is obtained by fixing ρ_1 and taking the volume of a snow cone corresponding to ϕ_1 away from a snow cone corresponding to $\phi_2 \geq \phi_1$; see the third solid in Figure 1. Therefore, invoking (4),

$$V_{SCC} = \frac{2\pi}{3} \rho_1^3 (1 - \cos \phi_2) - \frac{2\pi}{3} \rho_1^3 \left(1 - \cos \phi_1\right)$$

$$= \frac{2\pi}{3} \rho_1^3 \left(\cos \phi_1 - \cos \phi_2\right). \tag{5}$$

Step 4: Volume of a "railroad spike"

A "railroad spike" $\mathcal{RS}$ is obtained by slicing a snow cone cup with two vertical half-planes hinged at the z-axis, keeping one connected portion of the solid subtended by these planes; see the fourth solid in Figure 1. Assuming the planes correspond to θ_1 and $\theta_2 \geq \theta_1$, under (5) we obtain

$$V_{\mathcal{RS}} = \left(\frac{\theta_2 - \theta_1}{2\pi}\right) \frac{2\pi}{3} \rho_1^3 \left(\cos \phi_1 - \cos \phi_2\right)$$

$$= \frac{1}{3} \rho_1^3 \left(\theta_2 - \theta_1\right) \left(\cos \phi_1 - \cos \phi_2\right). \tag{6}$$

Step 5: Volume of a "spherical box"

A "spherical box" is obtained by cutting a railroad spike with radius ρ_1 out of a railroad spike with radius $\rho_2 \geq \rho_1$; see the fifth solid in Figure 1. By (6) we obtain

$$V_{\mathcal{B}} = \frac{1}{3} \rho_2^3 \left(\theta_2 - \theta_1\right) \left(\cos \phi_1 - \cos \phi_2\right) - \frac{1}{3} \rho_1^3 \left(\theta_2 - \theta_1\right) \left(\cos \phi_1 - \cos \phi_2\right)$$

$$= \frac{1}{3} \left(\rho_2^3 - \rho_1^3\right) \left(\theta_2 - \theta_1\right) \left(\cos \phi_1 - \cos \phi_2\right), \tag{7}$$

which is our (2) above!

Explaining the spherical coordinates integral formula

With the spherical box volume formula in hand, we can now explain the appended factor $\rho^2 \sin\phi$ appearing in (1). If we cover the solid $\mathcal{D}$ with n disjoint spherical boxes $\mathcal{B}_1, \ldots, \mathcal{B}_n$ and choose $(x_i, y_i, z_i) \in \mathcal{B}_i$ and $(\rho_i, \theta_i, \phi_i)$ so that

$$x_i = \rho_i \sin\phi_i \cos\theta_i, \quad y_i = \rho_i \sin\phi_i \sin\theta_i, \quad z_i = \rho_i \cos\phi_i, \quad 1 \le i \le n,$$

then, given certain suitable conditions on the function $f(x, y, z)$ over $\mathcal{D}$, the triple integral of f over $\mathcal{D}$ is approximated by the Riemann sum

$$\sum_{1 \le i \le n} f(\rho_i \sin\phi_i \cos\theta_i, \ \rho_i \sin\phi_i \sin\theta_i, \ \rho_i \cos\phi_i) \, V_{\mathcal{B}_i}. \tag{8}$$

It is now evident why volumes of spherical boxes are of preeminent concern when one endeavors to explain (1). Moreover, this Riemann sum approximation becomes precise if we let $n \to \infty$ with $\max_{1 \le i \le n} V_{\mathcal{B}_i} \to 0$. Thus, our task becomes determining how the volumes of these spherical boxes behave as they get small, and so we explore this now.

First, we outline a computation that does not rely on deep understanding of differentials, as some students have only limited exposure to this topic. We insert

$$\rho_1 = \rho, \quad \rho_2 = \rho + \Delta\rho, \quad \theta_1 = \theta,$$
$$\theta_2 = \theta + \Delta\theta, \quad \phi_1 = \phi, \quad \phi_2 = \phi + \Delta\phi$$

into (7), and using

$$x^3 - y^3 = (x - y)(x^2 + xy + y^2)$$
$$\cos(\alpha + \beta) = \cos\alpha \cos\beta - \sin\alpha \sin\beta,$$

we have

$$V_{\mathcal{B}} = \frac{1}{3} \left((\rho + \Delta\rho)^3 - \rho^3 \right) \left((\theta + \Delta\theta) - \theta \right) (\cos\phi - \cos(\phi + \Delta\phi))$$

$$= \frac{1}{3} \left(\Delta\rho(3\rho^2 + 3\rho\Delta\rho + (\Delta\rho)^2) \right) (\Delta\theta)$$

$$\cdot (\cos\phi(1 - \cos(\Delta\phi)) + \sin\phi \sin(\Delta\phi)). \tag{9}$$

As

$$\Delta\rho \to d\rho, \quad \Delta\theta \to d\theta, \quad \text{and} \quad \Delta\phi \to d\phi,$$

note that $\cos(\Delta\phi) \to 1$ and $\sin(\Delta\phi) \approx \Delta\phi \to d\phi$. This gives

$$V_{\mathcal{B}} = \frac{1}{3} \left(\Delta\rho \left(3\rho^2 + 3\rho\Delta\rho + (\Delta\rho)^2 \right) \right) (\Delta\theta)$$

$$\cdot (\cos\phi \, (1 - \cos(\Delta\phi)) + \sin\phi \sin(\Delta\phi))$$

$$\to \frac{1}{3}(3\rho^2)(\sin\phi) \, d\rho \, d\theta \, d\phi = \rho^2 \sin\phi \, d\rho \, d\theta \, d\phi.$$

This, together with (8), proves the spherical coordinates integral formula (1)!

We conclude by presenting a second approach through differentials in tandem with (7) to explain (1). Recall that if $g(x)$ is differentiable at x, we have the differential identity

$$\frac{d\,(g(x))}{dx}\,dx = g(x+dx) - g(x)$$

(see, e.g., Stewart [2, pp. 253–254]). Thus, passing to differentials we let

$$\Delta\rho \to d\rho, \quad \Delta\theta \to d\theta, \quad \text{and} \quad \Delta\phi \to d\phi$$

in (9) to obtain

$$V_{\mathcal{B}} = \frac{1}{3}\left((\rho+d\rho)^3 - \rho^3\right)\left((\theta+d\theta)-\theta\right)\left(\cos\phi - \cos(\phi+d\phi)\right)$$

$$= \frac{1}{3}\left(\frac{d\left(\rho^3\right)}{d\rho}\,d\rho\right)\left(\frac{d\,(\theta)}{d\theta}\,d\theta\right)\left(-\frac{d\,(\cos\phi)}{d\phi}\,d\phi\right)$$

$$= \rho^2 \sin\phi\,d\rho\,d\theta\,d\phi.$$

Remark. As a final remark, we note that the equation for the volume of a sphere of radius $a > 0$, $\frac{4}{3}\pi a^3$, is a simple consequence of any one of the formulas (3)–(7).

REFERENCES

[1] Edwards, B., Larson, R. (2014). *Calculus: Early Transcendental Functions*, 6th ed. Boston: Brooks Cole.
[2] Stewart, J. (2015). *Calculus: Early Transcendentals*, 8th ed. Boston: Brooks Cole.

Summary.   We present a fun alternative geometric approach for computing the volume of a "box" in spherical coordinates, which we then use to explain the formula for spherical triple integrals. Each step of our approach works well as an interactive classroom exercise.

JOHN ENGBERS (MR Author ID: 973477) earned his Ph.D. in graph theory from the University of Notre Dame in 2013. Currently, he is an assistant professor at Marquette University. When not immersed in mathematics, he loves spending time with his wife Ruth and kids Sam, Luke, and Joel, riding his bicycle, and enjoying nature.

ADAM HAMMETT (MR Author ID: 840691) earned his Ph.D. in combinatorial probability from The Ohio State University in 2007. Ever since, he has taught mathematics at the college level, and has served as an associate professor at Cedarville University since 2015. He enjoys overseeing undergraduate research projects, reading, and spending time outdoors with his wife Rachael and their four children Isabelle, Madison, Daniel, and Esther.
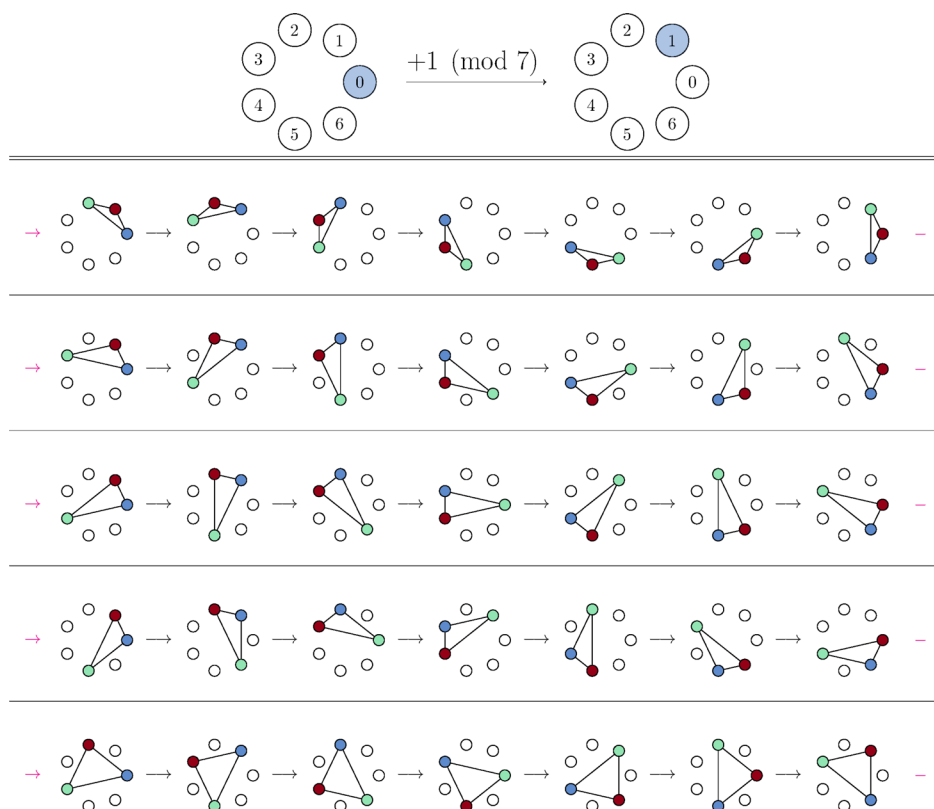
PROOF WITHOUT WORDS

Binomial Coefficients Modulo Primes

TOM EDGAR
Pacific Lutheran University
Tacoma, WA 98447
edgartj@plu.edu

Theorem. *If p is prime and $1 \leq k \leq p - 1$, then $\binom{p}{k} \equiv 0 \pmod{p}$.*

Proof. We illustrate the proof for the case $n = 7$ and $k = 3$, noting that the general idea of this diagram extends to any prime and collection of k-subsets except $k = 1$ and $k = p$.

Anderson, Benjamin, and Rouse [1] provided an argument that inspired this visual proof. The illustration can also serve as a verification of the stronger result that $\binom{n}{k} \equiv 0 \pmod{n}$ whenever $\gcd(n, k) = 1$. However, in the general case, it is not clear from the diagram that every orbit will have exactly n representatives.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.
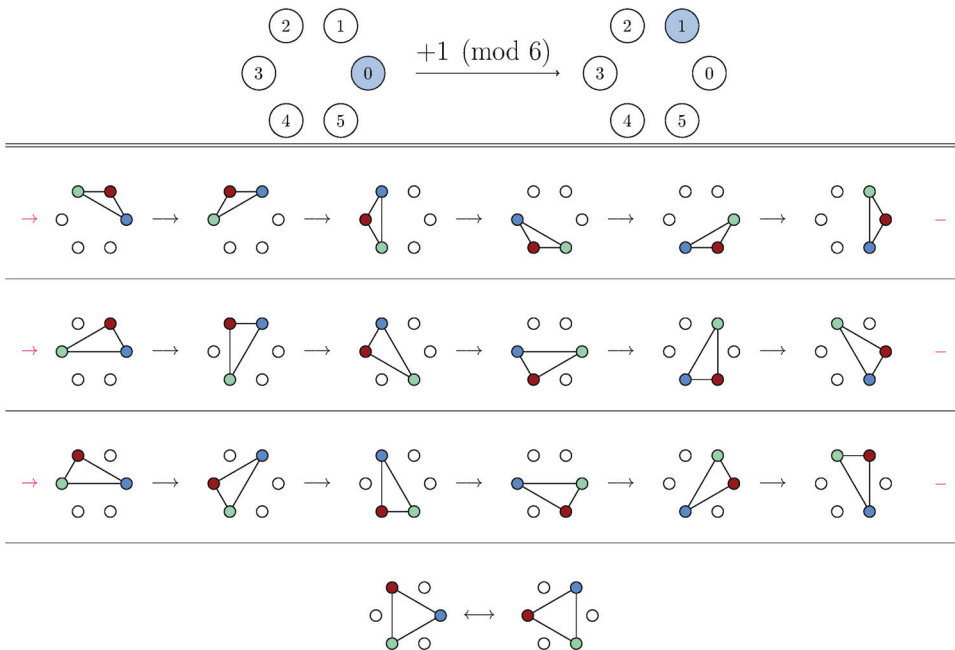
Figure 1 The binomial coefficient $\binom{6}{3}$ is not congruent to 0 modulo 6 because there is a single orbit that has only two subsets while the other orbits all have six. This orbit arises from the set containing the multiples of $2 = 6/3$ (modulo 6).

In Figure 1, we demonstrate that the binomial coefficient $\binom{n}{k}$ might not be zero modulo n when n and k are not relatively prime: the orbit of a k-subset need not contain a full collection of n subsets as required in the visual proof. In general, the orbit of the set

$$\{x \cdot \gcd(n, k) \mid 0 \leq x < n/\gcd(n, k)\}$$

(pictured in the last row of Figure 1 when $n = 6$ and $k = 3$) will have $n/\gcd(n, k)$ subsets, which is less than n when $\gcd(n, k) > 1$.

Nonetheless, we can find pairs of integers, n and k, with a common divisor larger than one where $\binom{n}{k} \equiv 0 \pmod{n}$ (e.g., $n = 10$ and $k = 4$). Robbins [3, 4] provided characterizations for the set of squarefree integers n with the property that $\binom{n}{k} \equiv 0 \pmod{n}$, if and only if $\gcd(n, k) = 1$. Robbins' work was inspired by Harborth's proof [2] that almost every binomial coefficient $\binom{n}{k}$ is divisible by n.

REFERENCES

[1] Anderson, P. G., Benjamin, A. T., Rouse, J. A. (2005). Combinatorial proofs of Fermat's, Lucas's, and Wilson's theorems. *Amer. Math. Monthly*. 112(3): 266–268. doi.org/10.1080/00029890.2005.11920193

[2] Harborth, H. (1977). Divisibility of binomial coefficients by their row number. *Amer. Math. Monthly*. 84(1): 35–37. doi.org/10.2307/2318304

[3] Robbins, N. (1982). On the number of binomial coefficients which are divisible by their row number. *Can. Math. Bull.* 25(3): 363–365. doi.org/10.4153/CMB-1982-052-3

[4] Robbins, N. (1985). On the number of binomial coefficients which are divisible by their row number: II. *Can. Math. Bull.* 28(4): 481–486. doi.org/10.4153/CMB-1985-059-0

Summary. We wordlessly demonstrate that the binomial coefficient $\binom{p}{k}$ is congruent to 0 modulo p for positive k less than p.

TOM EDGAR (MR Author ID: 821633) is an associate professor at Pacific Lutheran University and the editor of *Math Horizons*.

PROBLEMS

LES REID, *Editor*
Missouri State University

EUGEN J. IONAŞCU, *Proposals Editor*
Columbus State University

RICHARD BELSHOFF, Missouri State University; *EYVINDUR ARI PALSSON*, Virginia Tech; *CODY PATTERSON*, Texas State University; *ROGELIO VALDEZ*, Centro de Investigación en Ciencias, UAEM, Mexico; *Assistant Editors*

Proposals

To be considered for publication, solutions should be received by September 1, 2020.

2091. *Proposed by Marian Tetiva, National College "Gheorghe Roşca Codreanu," Bârlad, Romania.*

Let *ABC* be a triangle with sides of lengths a, b, c, altitudes h_a, h_b, h_c, inradius r, and circumradius R. Prove that the following inequality holds:

$$h_a + h_b + h_c \geq 9r + \frac{a^2 + b^2 + c^2 - ab - ac - bc}{4R},$$

with equality if and only if $\triangle ABC$ is equilateral.

2092. *Proposed by Seán M. Stewart, Bomaderry, Australia.*

Let n be a non-negative integer. Evaluate

$$\int_0^\infty \frac{1}{x^{2n+3}} \left(\sin x - \sum_{k=0}^n \frac{(-1)^k x^{2k+1}}{(2k+1)!} \right) dx.$$

2093. *Proposed by Jacob Siehler, Gustavus Adolphus College, Saint Peter, MN.*

Suppose π is a permutation of $\{1, 2, \ldots, 2m\}$, where m is a positive integer. Consider the (possibly empty) subsequence of $\pi(m+1), \pi(m+2), \ldots, \pi(2m)$ consisting of only those values which exceed $\max\{\pi(1), \ldots, \pi(m)\}$. Let $P(m)$ denote the probability that this subsequence never decreases (note that the empty sequence has this property), when π is a randomly chosen permutation of $\{1, \ldots, 2m\}$. Evaluate $\lim_{m \to \infty} P(m)$.

2094. *Proposed by George Stoica, Saint John, New Brunswick, Canada.*

We invite readers to submit original problems appealing to students and teachers of advanced undergraduate mathematics. Proposals must always be accompanied by a solution and any relevant bibliographical information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution. Submitted problems should not be under consideration for publication elsewhere.

Proposals and solutions should be written in a style appropriate for this MAGAZINE.

Authors of proposals and solutions should send their contributions using the Magazine's submissions system hosted at http://mathematicsmagazine.submittable.com. *More detailed instructions are available there. We encourage submissions in PDF format, ideally accompanied by LᴬTEX source. General inquiries to the editors should be sent to* mathmagproblems@maa.org.

Find the smallest number $f(n)$ such that for any set of unit vectors $x_1, \ldots, x_n$ in $\mathbb{R}^n$, there is a choice of $a_i \in \{-1, 1\}$ such that $|a_1 x_1 + \cdots + a_n x_n| \leq f(n)$.

2095. *Proposed by Mircea Merca, University of Craiova, Romania.*

Show that

$$\sum_{k=1}^{n} k \left\lfloor \frac{n+1-k}{d} \right\rfloor = \begin{cases} \lceil (n+1)(n-1)(2n+3)/24 \rceil & \text{if } d = 2 \\ \lceil (n+1)^2(n-2)/18 \rceil & \text{if } d = 3 \\ \lceil (n+1)(2n+1)(n-3)/48 \rceil & \text{if } d = 4 \\ \lceil (n+1)n(n-4)/30 \rceil & \text{if } d = 5 \end{cases}.$$

Quickies

1099. *Proposed by Mihàly Bencze, Braşov, Romania.*

Given $a > 0$, find all positive $x_1, \ldots, x_n$ satisfying the following system of equations.

$$\sqrt{a}\left(\sqrt{x_1} + \sqrt{x_2}\right) = \sqrt{(x_3 + a)(x_4 + a)}$$

$$\sqrt{a}\left(\sqrt{x_2} + \sqrt{x_3}\right) = \sqrt{(x_4 + a)(x_5 + a)}$$

$$\vdots$$

$$\sqrt{a}\left(\sqrt{x_{n-1}} + \sqrt{x_n}\right) = \sqrt{(x_1 + a)(x_2 + a)}$$

$$\sqrt{a}\left(\sqrt{x_n} + \sqrt{x_1}\right) = \sqrt{(x_2 + a)(x_3 + a)}.$$

1100. *Proposed by Richard Ehrenborg, University of Kentucky, Lexington, KY.*

Let L_1 and L_2 be two parallel lines in the Euclidean plane. Let $x_1, x_2, \ldots, x_m$ be m points on the line L_1 and $y_1, y_2, \ldots, y_n$ be n points on the line L_2. Assume that all the $m + n$ points are in general position, that is, no three line segments of the form $\overline{x_i y_j}$ intersect in a point strictly between L_1 and L_2. Into how many parts do the mn line segments $\overline{x_i y_j}$ divide the region between L_1 and L_2?

Solutions

2066. *Proposed by George Stoica, Saint John, New Brunswick, Canada.*

Is there a function $f : \mathbb{R} \to \mathbb{R}$ satisfying $|f(x + y)| \geq |f(x) + f(y)|$ for all $x, y \in \mathbb{R}$, with strict inequality for at least some x, y? How about a function satisfying the reverse inequality $|f(x + y)| \leq |f(x) + f(y)|$ everywhere, strictly somewhere?

Solution by Tom Jager, Calvin College, Grand Rapids, MI.

No function satisfies the first condition. Suppose that $|f(x + y)| \geq |f(x) + f(y)|$ for all x and y. We will show that $f(x + y) = f(x) + f(y)$ for all x and y, violating the condition that the inequality be strict for at least some x, y. First,

$$|f(0)| = |f(0 + 0)| \geq |f(0) + f(0)| = 2|f(0)|,$$

so $f(0) = 0$. Hence,

$$0 = |f(x + (-x))| \geq |f(x) + f(-x)|,$$

so $f(-x) = -f(x)$.

Suppose $f(x) \geq 0$ and $f(y) \geq 0$. Since

$$|f(x + y) - f(x)| \leq |f(y)| = f(y),$$
$$-f(y) \leq f(x + y) - f(x) \leq f(y), \text{ so}$$
$$f(x) - f(y) \leq f(x + y) \leq f(x) + f(y).$$

Similarly,

$$f(y) - f(x) \leq f(x + y) \leq f(x) + f(y).$$

Therefore,

$$0 \leq f(x + y) \leq f(x) + f(y).$$

We now have

$$f(x + y) = |f(x + y)| \geq |f(x) + f(y)| = f(x) + f(y).$$

Combining these inequalities, we obtain $f(x + y) = f(x) + f(y)$.

If $f(x) \leq 0$ and $f(y) \leq 0$, then $f(-x) \geq 0$ and $f(-y) \geq 0$. From the previous case, we have

$$f(x + y) = -(f(-x + (-y))) = -(f(-x) + f(-y)) = f(x) + f(y)$$

as desired.

Finally, if $f(x)$ and $f(y)$ have opposite signs, we may assume, without loss of generality, that $f(x) > 0$ and $f(y) < 0$. If $f(x + y) \geq 0$, then by the first case,

$$f(x + y) + f(-y) = f(x + y - y) = f(x), \text{ and hence}$$
$$f(x + y) = f(x) + f(y).$$

If $f(x + y) \leq 0$, then by the second case,

$$f(x + y) + f(-x) = f(x + y - x) = f(y), \text{ and hence}$$
$$f(x + y) = f(x) + f(y).$$

A function satisfying the second condition is given by

$$f(x) = \begin{cases} 0 & \text{if } x \text{ is rational} \\ 1 & \text{if } x \text{ is irrational} \end{cases}.$$

Note that

$$|f(x + y)| = |f(x)| + |f(y)| \text{ if either } x \text{ or } y \text{ is rational and}$$

$$|f(x + y)| < |f(x)| + |f(y)| \text{ if both } x \text{ and } y \text{ are irrational.}$$

Also solved by Robert Calcaterra, Kyle Gatesman, Eugene A. Herman, Albert Natian, and the proposer. There was one incomplete or incorrect solution.

2067. *Proposed by Elton Bojaxhiu, Eppstein am Taunus, Germany and Enkel Hysnelaj, Sydney, Australia.*

Chord $\overline{XY}$ of a circle $\mathcal{C}$ is not a diameter. Let P, Q be two different points strictly inside $\overline{XY}$ such that Q lies between P and X. Chord $\overline{MN}$ through P is perpendicular to the diameter of $\mathcal{C}$ through Q, where $MP < NP$. Prove that $(MQ - PQ) \cdot XY \geq 2 \cdot QX \cdot PY$, and characterize those cases in which equality holds.

Solution by Robert Calcaterra, University of Wisconsin, Platteville, Platteville, WI.
Let R be the intersection of $\overline{MN}$ and the diameter of C through Q. Choose a coordinate system so that C is the unit circle, Q is at $(a, 0)$, R is at $(b, 0)$, and P is at (b, c) with $c > 0$. Then M is at $(b, \sqrt{1 - b^2})$ and N is at $(b, -\sqrt{1 - b^2})$. Let $u = QX$, $v = PQ$, and $w = PY$. By the intersecting chords theorem,

$$w(u + v) = PY \cdot PX = MP \cdot PN = \left(\sqrt{1 - b^2} - c\right)\left(\sqrt{1 - b^2} + c\right) = 1 - b^2 - c^2.$$

Moreover,

$$MQ^2 = QR^2 + MR^2 = PQ^2 - PR^2 + MR^2 = v^2 - c^2 + (1 - b^2) = v^2 + uw + vw.$$

Note that the inequality we seek to establish is equivalent to

$$MQ^2 \cdot XY^2 - (2 \cdot QX \cdot PY + PQ \cdot XY)^2 \geq 0.$$

Now the left-hand side of this inequality is

$$(v^2 + uw + vw)(u + v + w)^2 - (2uw + v(u + v + w))^2 = w(u + v)(u - v - w)^2.$$

Observe that $w(u + v)(u - v - w)^2$ is never negative and only takes on the value 0 when $u = v + w$. Therefore, $(MQ - PQ) \cdot XY \geq 2 \cdot QX \cdot PY$ and equality holds if and only if $QX = PQ + PY$, i.e. when Q is the midpoint of $\overline{XY}$.

Also solved by the proposers. There was one incomplete or incorrect solution.

2068. *Proposed by Ovidiu Furdui and Alina Sîntămariăn, Technical University of Cluj-Napoca, Cluj-Napoca, Romania.*

Prove that the series

$$\sum_{n=1}^{\infty} \frac{3 \cdot 6 \cdots (3n)}{7 \cdot 10 \cdots (3n + 4)} \cdot \frac{1}{3n + 7}$$

converges, and find its sum.

Solution by Jake Viscusi (student),The Episcopal Academy, Newtown Square, PA.
Note that the series may be rewritten as

$$\frac{1}{3} \sum_{n=1}^{\infty} \left(\prod_{j=1}^{n} \frac{j}{\frac{4}{3} + j}\right) \frac{1}{\frac{4}{3} + n + 1}.$$

In view of this, we generalize to series of the form (for $c > 0$)

$$\sum_{n=1}^{\infty} \left(\prod_{j=1}^{n} \frac{j}{c+j} \right) \frac{1}{c+n+1} = \sum_{n=1}^{\infty} \frac{\Gamma(c+1)}{\Gamma(c+1)} \left(\prod_{j=1}^{n} \frac{j}{c+j} \right) \frac{1}{c+n+1}$$

$$= \sum_{n=1}^{\infty} \frac{\Gamma(c+1)\,\Gamma(n+1)}{\Gamma(c+n+2)}$$

$$= \sum_{n=1}^{\infty} B(c+1, n+1),$$

where the identity $\alpha \Gamma(\alpha) = \Gamma(\alpha + 1)$ has been used $n + 1$ times in the denominator of the right-hand side of the second equation and $B(\alpha, \beta)$ is the beta function. By definition,

$$\sum_{n=1}^{\infty} B(c+1, n+1) = \sum_{n=1}^{\infty} \int_0^1 (1-x)^c \, x^n \, dx$$

$$= \lim_{m \to \infty} \sum_{n=1}^{m} \int_0^1 (1-x)^{c-1} \left(x^n - x^{n+1} \right) dx$$

$$= \lim_{m \to \infty} \left(\int_0^1 (1-x)^{c-1} x \, dx - \int_0^1 (1-x)^{c-1} x^{m+1} \, dx \right)$$

(telescoping sum)

$$= \frac{1}{c(c+1)} - \lim_{m \to \infty} \int_0^1 (1-x)^{c-1} x^{m+1} \, dx.$$

Note, for any $0 \le x \le 1$, we have $x(1-x) \le \frac{1}{4}$, so

$$\lim_{m \to \infty} \int_0^1 (1-x)^{c-1} x^{m+1} dx \le \lim_{m \to \infty} \int_0^1 \left(\frac{1}{4} \right)^{c-1} x^{m+2-c} dx$$

$$= \lim_{m \to \infty} \frac{1}{(4)^{c-1}(m+3-c)} = 0.$$

Thus

$$\sum_{n=1}^{\infty} \left(\prod_{j=1}^{n} \frac{j}{c+j} \right) \frac{1}{c+n+1} = \frac{1}{c(c+1)}.$$

Setting $c = 4/3$ and multiplying by $1/3$, we arrive at the solution to the original problem: $3/28$.

Also solved by Ulrich Abel & Vitaliy Kushnirevych (Germany), Farrukh Ataev (Uzbekistan), Michel Bataille (France), Paul Bracken, Brian Bradie, Robert Calcaterra, Hongwei Chen, Pedro Acosta De León, Robert Doucette, Dmitry Fleischman, Kyle Gatesman, Marty Getz & Dixon Jones, Michael Goldenberg & Mark Kaplan, Russell Gordon, GWstat Problem Solving Group, Albert Natian, Northwestern University Problem Solving Group, Moubinool Omarjee (France), Adebola Omotajo & Habib Far, Randy Schwartz, Nicholas Singer, Albert Stadler (Switzerland), Michael Vowe (Switzerland), and the proposers.

2069. *Proposed by Eugene Delacroix, Lycee Therese dAvila, France and Su Pernu Mero, Valenciana GTO, Mexico.*

Three points are chosen uniformly and independently at random in the unit interval $[0, 1]$. These points divide the interval into four segments of lengths a, b, c, and d. Find the expected value and standard deviation of the random variable $X = \max(a, b, c, d)$.

Solution by Elton Bojaxhiu, Eppstein am Taunus, Germany and Enkel Hysnelaj, Sydney, Australia.
We will develop a general recurrence formula for the case of n points, $X_1, \ldots, X_n$, chosen uniformly and independently at random from $[0, 1]$. These n points divide $[0, 1]$ into $n + 1$ segments. Let X be the maximal length of those segments. Let $F_n(\lambda) = P(X \leq \lambda)$ be the distribution of X.

For the case of one point ($n = 1$), it is easy to see that

$$F_1(\lambda) = \begin{cases} 0 & \text{if } \lambda \leq \frac{1}{2} \\ 2\lambda - 1 & \text{if } \frac{1}{2} \leq \lambda \leq 1 \\ 1 & \text{if } \lambda \geq 1 \end{cases}.$$

For $n \geq 2$, we will derive a formula for F_n from that of F_{n-1}. Let $\lambda \in [0, 1]$. Since all n variables have the same distribution and are independent, we need to treat only the case when X_1 is the smallest variable. So

$$F_n(\lambda) = nP(X \leq \lambda, X_1 < X_2, X_1 < X_3, \ldots, X_1 < X_n)$$
$$= n\left(P(X \leq \lambda, X_1 < X_2, X_1 < X_3, \ldots, X_1 < X_n, X_1 \leq \lambda)\right.$$
$$\left.+P(X \leq \lambda, X_1 < X_2, X_1 < X_3, \ldots, X_1 < X_n, X_1 > \lambda)\right).$$

If $X_1 > \lambda$, then $X > \lambda$. Hence the second probability is 0. Therefore

$$F_n(\lambda) = nP(X \leq \lambda, X_1 < X_2, X_1 < X_3, \ldots, X_1 < X_n, X_1 \leq \lambda)$$
$$= nP(X \leq \lambda, X_1 \leq \lambda | X_1 < X_2, X_1 < X_3, \ldots, X_1 < X_n)$$
$$\cdot P(X_1 < X_2, X_1 < X_3, \ldots, X_1 < X_n)$$
$$= n\int_0^\lambda P(X \leq \lambda | X_2, X_3, \ldots, X_n > X_1 = x)(1 - x)^{n-1}dx.$$

But, $P(X \leq \lambda | X_2, X_3, \ldots, X_n > X_1 = x)$ is the probability that $X \leq \lambda$ when $X_2, X_3, \ldots, X_n$ are uniformly distributed from $[x, 1]$. Therefore

$$P(X \leq \lambda | X_2, X_3, \ldots, X_n > X_1 = x) = F_{n-1}\left(\frac{\lambda}{1 - x}\right).$$

Hence

$$F_n(\lambda) = n\int_0^\lambda F_{n-1}\left(\frac{\lambda}{1 - x}\right)(1 - x)^{n-1}dx.$$

We need to find $F_3(\lambda)$. Knowing $F_1(\lambda)$ and using the recurrence relation, we compute

that

$$F_2(\lambda) = \begin{cases} 0 & \text{if } \lambda \leq \frac{1}{3} \\ (3\lambda - 1)^2 & \text{if } \frac{1}{2} \geq \lambda \geq \frac{1}{3} \\ -3\lambda^2 + 6\lambda - 2 & \text{if } 1 \geq \lambda \geq \frac{1}{2} \\ 1 & \text{if } \lambda \geq 1 \end{cases}$$

and knowing $F_2(\lambda)$, a little more work yields

$$F_3(\lambda) = \begin{cases} 0 & \text{if } \lambda \leq \frac{1}{4} \\ (4\lambda - 1)^3 & \text{if } \frac{1}{3} \geq \lambda \geq \frac{1}{4} \\ -44\lambda^3 + 60\lambda^2 - 24\lambda + 3 & \text{if } \frac{1}{2} \geq \lambda \geq \frac{1}{3} \\ 4\lambda^3 - 12\lambda^2 + 12\lambda - 3 & \text{if } 1 \geq \lambda \geq \frac{1}{2} \\ 1 & \text{if } \lambda \geq 1 \end{cases}.$$

Since the support of X is non-negative, it is well known that

$$E(X) = \int_0^1 (1 - F_n(x))\, dx \text{ and } E(X^2) = 2\int_0^1 x(1 - F_n(x))\, dx.$$

From those formulas, we can easily compute that, in the case $n = 3$,

$$E(X) = \frac{25}{48} \approx 0.521,\ E(X^2) = \frac{83}{288} \approx 0.288,$$

$$\text{var}(X) = E(X^2) - (E(X))^2 = \frac{13}{768} \approx 0.017, \text{ and}$$

$$\sigma_X = \sqrt{\text{Var}(X)} = \sqrt{\frac{13}{768}} \approx 0.130.$$

Also solved by Robert A. Agnew, Robert Calcaterra, J. A. Grzesik, GWstat Problem Solving Group, John C. Kieffer, Peter McPolin (Northern Ireland), Carl M. Russell, Randy K. Schwartz, Richard Stong, Lawrence R. Weill, and the proposers. There was one incorrect or incomplete solution.

2070. *Proposed by Enrique Treviño, Lake Forest College, Lake Forest, IL.*

Fix a prime p. For any integer $n \geq p$, let S_n be the number of ways of coloring n points using p distinct colors, each at least once. Characterize those n such that S_n is not a multiple of p^2.

Editor's Note. The points are assumed to be distinguishable.

Solution by Bill Cowieson, Fullerton College, Fullerton, CA.
We claim that S_n is not a multiple of p^2 if and only if $n - 1$ is a multiple of $p - 1$.

For $1 \leq i \leq p$, let A_i be the set of colorings that omit color i. By the inclusion-exclusion principle,

$$S_n = p^n - |A_1 \cup \cdots \cup A_p|$$

$$= p^n + \sum_{k=1}^{p} (-1)^k \sum_{1 \leq i_1 < \cdots < i_k \leq p} |A_{i_1} \cap \cdots \cap A_{i_k}|$$

$$= p^n + \sum_{k=1}^{p-1} (-1)^k \binom{p}{k} (p-k)^n.$$

For $1 \le k \le p-1$,

$$\binom{p}{k} = \frac{p}{p-k}\binom{p-1}{k}, \quad \text{so} \quad \frac{S_n}{p} = p^{n-1} + \sum_{k=1}^{p-1}(-1)^k\binom{p-1}{k}(p-k)^{n-1}.$$

Note that

$$(p-k)^{n-1} \equiv (-1)^{n-1}k^{n-1} \pmod{p}$$

and

$$\binom{p-1}{k} = \frac{(p-1)(p-2)\cdots(p-k)}{1\cdot 2\cdots k} \equiv (-1)^k \pmod{p},$$

so

$$\frac{S_n}{p} \equiv (-1)^{n-1}\sum_{k=1}^{p-1}k^{n-1}.$$

Let a be a primitive root modulo p, so $\{a^1, a^2, a^3, \ldots, a^{p-1}\} \equiv \{1, 2, \ldots, p-1\}$ (mod p). It follows that

$$\frac{S_n}{p} \equiv (-1)^{n-1}\sum_{k=1}^{p-1}k^{n-1} \pmod{p} \equiv (-1)^{n-1}\sum_{i=1}^{p-1}(a^i)^{n-1} \pmod{p}$$

$$\equiv (-1)^{n-1}\sum_{i=1}^{p-1}b^i \pmod{p},$$

where $b = a^{n-1}$. Now $b \equiv 1 \pmod{p}$ if and only if $p-1$ divides $n-1$. In that case,

$$\frac{S_n}{p} \equiv (-1)^{n-1}(p-1) \equiv (-1)^n \not\equiv 0 \pmod{p}.$$

If $b \not\equiv 1 \pmod{p}$, we have the finite geometric series

$$\frac{S_n}{p} \equiv (-1)^{n-1}\sum_{i=1}^{p-1}b^i \pmod{p} \equiv (-1)^{n-1}(b^p - b)(b-1)^{-1} \equiv 0 \pmod{p}.$$

Thus, p^2 does not divide S_n if and only if $n-1$ is a multiple of $p-1$.

Editor's Note. A number of solvers noted that $S_n = p!S(n, p)$, where $S(n, p)$ is a Stirling number of the second kind.

Also solved by Armstrong Problem Solvers, Elton Bojaxhiu (Germany) & Enkel Hysnelaj (Australia), Robert Calcaterra, Neville Fogarty & Charles Samuels, Harris Kwong, Theophilus Pedapolu, Nicholas C. Singer, David Stone & John Hawkins, and the proposer. There was one incomplete or incorrect solutions.

Answers

Solutions to the Quickies from page 151.

A1099. Adding the equations, we obtain

$$\sum_{i=1}^{n} \left(\sqrt{(x_i + a)(x_{i+1} + a)} - \sqrt{a} \left(\sqrt{x_i} + \sqrt{x_{i+1}} \right) \right) = 0,$$

where the indices are taken modulo n. Now

$$\sqrt{(x_i + a)(x_{i+1} + a)} - \sqrt{a} \left(\sqrt{x_i} + \sqrt{x_{i+1}} \right) \geq 0 \Leftrightarrow$$

$$(x_i + a)(x_{i+1} + a) \geq a \left(\sqrt{x_i} + \sqrt{x_{i+1}} \right)^2 \Leftrightarrow$$

$$x_i x_{i+1} + a^2 \geq 2a \sqrt{x_i} \sqrt{x_{i+1}} \Leftrightarrow$$

$$\left(\sqrt{x_i x_{i+1}} - a \right)^2 \geq 0,$$

with equality if and only if $x_i x_{i+1} = a^2$. Since the terms in the equation above are non-negative and sum to zero, each term must be zero. Therefore we must have

$$x_1 x_2 = x_2 x_3 = \cdots x_n x_1 = a^2,$$

and hence

$$x_1 = x_3 = x_5 = \cdots = b > 0 \text{ and}$$

$$x_2 = x_4 = x_6 = \cdots = \frac{a^2}{b}.$$

If $n = 2k$, this satisfies the original system of equations.

If $n = 2k + 1$, then $x_n x_1 = a^2$ gives $b^2 = a^2$ and hence $b = a$. Therefore

$$x_1 = x_2 = \cdots = x_n = a$$

and this satisfies the original system of equations.

A1100. Without loss of generality, assume that L_1 and L_2 are vertical lines. If necessary, perturb the points $x_1, \ldots, x_m, y_1, \ldots, y_n$ such that no line segment $\overline{x_i y_k}$ is horizontal. Furthermore, assume that the points x_i are in order of their heights and similarly assume the same for the points y_k. Three cases can occur.

a. The lowest point in a region is at the intersection of two line segments. There are $\binom{m}{2}$ ways to select two points x_i and x_j, where $i < j$, and $\binom{n}{2}$ ways to select two points y_k and y_ℓ, where $k < \ell$. The lowest point will be the intersection of two line segments $\overline{x_i y_\ell}$ and $\overline{x_j y_k}$. There are $\binom{m}{2}\binom{n}{2}$ such regions.
b. The lowest point in a region is the intersection of a line segment and one of the lines L_1 and L_2. There are mn ways to select a pair x_i and y_k. Depending on which of the two points is lower, the lowest point will either lie on L_1 and $\overline{x_i y_k}$ or on L_2 and $\overline{x_i y_k}$. But only one of these two cases occurs. The number of such cases is $mn = \binom{m}{1}\binom{n}{1}$.
c. There is no lowest point. In this case the region is the unique region that is unbounded from below, giving $1 = \binom{m}{0}\binom{n}{0}$ regions.

Thus, the total number of regions is

$$\binom{m}{2}\binom{n}{2} + \binom{m}{1}\binom{n}{1} + \binom{m}{0}\binom{n}{0}.$$

REVIEWS

PAUL J. CAMPBELL, *Editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles, books, and other materials are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of mathematics literature. Readers are invited to suggest items for review to the editors.

Grechuk, Bogdan, *Theorems of the 21st Century*, Vol. 1, Springer, 2019; xvi+446 pp, $54.99. ISBN 978-3-030-19095-8.

Years ago, I used to give a lecture titled "Mathematics of the Last Five Years"—which of course needed to be updated on a regular basis. Author Grechuk, in a monumental undertaking done astonishingly well, now makes that easy: He has organized by year background, explanation, and commentary on half a dozen to a dozen new theorems for each year. This volume covers 2001 through 2010. It is limited to results that were communicated in the *Annals of Mathematics* (hence, e.g., Perelman's proof of the Poincaré conjecture is not covered). Apart from importance so-denominated, the other criterion for inclusion is a "sufficiently simple formulation which can be understood and appreciated by readers with at most an undergraduate (ideally high school) education." The emphasis is on the formulation, importance, and applications of each theorem; proofs are not discussed. I hope we do not have to wait 10 years for the volume covering 2011 through 2020; an annual report—perhaps published in the *Annals* itself—would be ideal.

West, Geoffrey, *Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economics, and Companies*, Penguin Press, 2017; iv+481 pp, $30. ISBN 978-1-59420-558-3.

This is a remarkably stimulating book about complexity of many kinds. It investigates and documents an enormous number of scaling relationships and demonstrates that biological characteristics of animals, plants, cities, and companies all follow common generic laws. For instance, measured in terms of dimensionless quantities, the growth curve is the same for all animals. The reason is that "birth, growth, and death are all governed" by dynamics determined by metabolic rate and minimization of energy. These consequences are the result of fractal networks. Author West addresses why humans cannot live a thousand years, why companies die but cities grow, and why the pace of innovation must continue to accelerate. Amid the explorations, he tries to find partial answers to profound questions such as, "Can we maintain a vibrant, innovative society driven by ideas and wealth creation, or are we destined to become a planet of slums, conflict, and devastation?"

Alexander, Amir, *Proof: How the World Became Geometrical*, Scientific American/Farrar, Straus and Giroux, 2019; iii+305 pp, $28. ISBN 978-0-374-25490-2.

"The world is deeply geometrical." Author Alexander begins with a chapter about confiscation by Louis XIV of an estate of one of his ministers (asserting without documenting that the superior geometrical order of its gardens was what led to its confiscation). He goes on to investigate the "exalted place" of geometry in the West as "an emblem of universal order and a tool of political power." "The universal principles of geometry, as presented at Versailles, were the deep foundations of royal supremacy." The book recounts how the "rational and eternal world of absolute truth" encapsulated in geometry shaped the landscape from the 1400s on—not just in social and political aspects, but particularly in the landscape of great gardens and designed cities such as Washington, D.C.

Karaali, Gizem, and Lily S. Khadjavi, *Mathematics for Social Justice: Resources for the College Classroom*, AMS/MAA Press, 2019; vii+277 pp, $55 ($41.25 for AMS/MAA members). ISBN 978-1-4704-4926-1.

Buell, Catherine A., and Bonnie Shulman (eds.), Special Issue on Mathematics for Social Justice, *PRIMUS* 29 (3–4) (2019) 205–419.

The book has five short reflective essays on mathematics and social justice, including sustainability and environmental justice, and 14 teaching modules. Half of the modules have no prerequisites and are suitable for a general education or liberal arts course in mathematics, and half are for precalculus, calculus, or differential equations. (Regrettably, there is no index.) A second volume will provide 17 modules for introductory statistics and quantitative reasoning courses. The special issue of *PRIMUS* includes further essays and modules, including "Unnatural disasters . . ." by Karaali and Khadjavi, which features two more modules suitable for calculus, one dealing with mercury poisoning in Iraqi grain and the other with methane release in southern California. Modules such as those bring details and context of real situations in social science and physical science into applying mathematics, thereby enriching the mathematics experience and enhancing some students' motivation and excitement about mathematics. However, much of the content of the essays (but not the modules) is about subsequent substantial elaboration and discussion of the social and political issues involved, not about the mathematics.

Boghosian, Bruce M., The inescapable casino, *Scientific American* 321 (5) (November 2019) 71–77.

A simple model shows that ordinary commerce, after a large number of "seemingly innocuous but subtly biased transactions," inevitably leads to wealth inequality. "The free market is essentially a casino that you can never leave . . . [Y]ou win some and you lose some, but the longer you stay in the casino, the more likely you are to lose." This three-parameter affine wealth model fits extremely well the Lorenz curves of the U.S. and other countries. "[F]ar from wealth trickling down to the poor, the natural inclination of wealth is to flow upward so that the 'natural' wealth distribution in a free-market economy is one of complete oligarchy . . . [O]nly a carefully designed mechanism for redistribution can compensate . . ."

Dawson, C. Bryan, Calculus limits unified and simplified, *College Mathematics Journal* 50 (5) (November 2019) 331–342.

This article argues for the introduction of infinitesimals in calculus, accompanies them with intuitive definitions and notation, and introduces a new, easy, and very natural method for evaluating limits. Such an approach has the advantages of expanding students' notion of number—integers, fractions, negative numbers, irrationals, why not infinitesimals?—and relieving them of the computational burden that gets in the way of understanding the concept of limit. In particular, the methodology lets students follow their natural inclination to plug in a value to try to evaluate $\lim_{x \to b} f(x)$; however, what they plug in is not b, but $b + \alpha$, where α is an arbitrary infinitesimal, and they then check if they get the same rendered real result regardless of α.

Richeson, David S., *Tales of Impossibility: The 2000-Year Quest to Solve the Mathematical Problems of Antiquity*, Princeton University Press, 2019; xii+436 pp, $29.95. ISBN 978-0-691-19296-3.

This book discusses the history and solution of the four problems of antiquity: squaring the circle (as well as rectifying it), doubling the cube, constructing regular polygons, and trisecting angles. Author Richeson clearly explains what it means to be impossible to solve a problem, cites other impossibility results, goes into detail about geometric constructions with various instruments, and discusses the defective proofs and the cranks that have turned up along the way. The discussions focus on geometric concepts and roots of polynomials (number fields get scant mention). Richeson devotes a chapter to the long-time lack of recognition of the achievement of Pierre Wantzel. At age 20, Wantzel was the first to prove that the duplication and trisection problems, as well as that of constructing every regular polygon, cannot be solved. A "deafening silence" and neglect for 100 years greeted publication of his results, despite the fact that they appeared together (in italics) on the same page in a well-circulated journal.