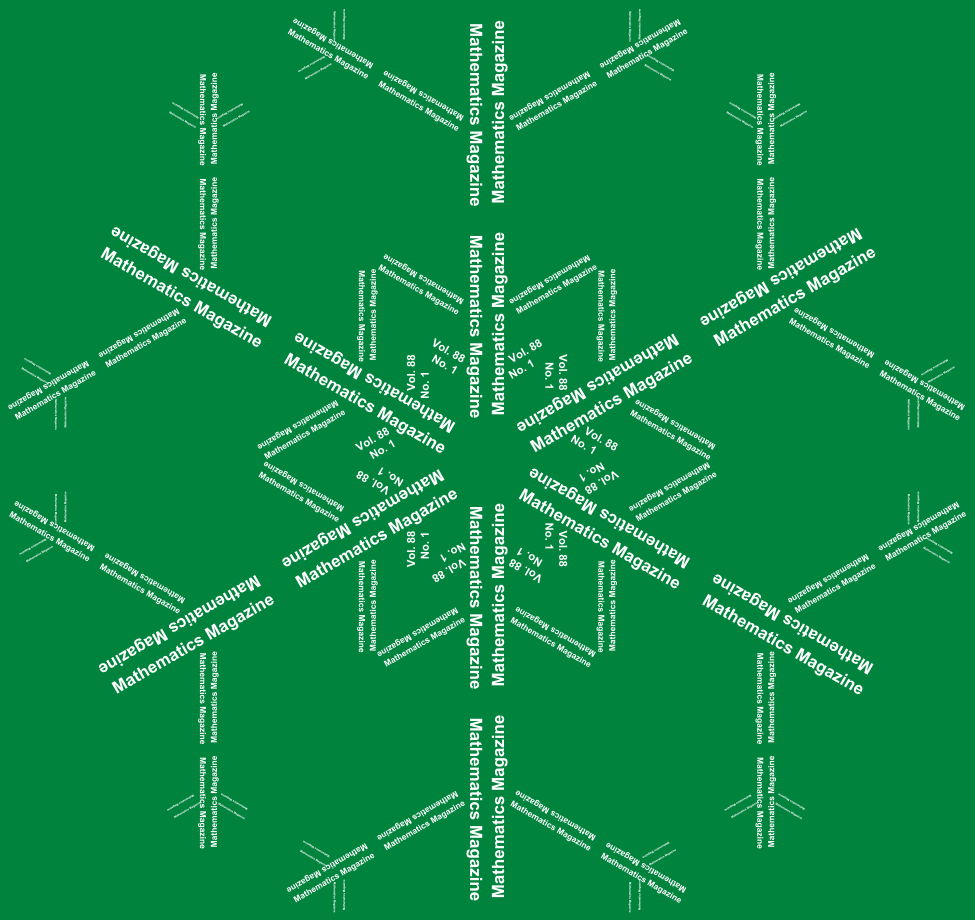
Vol. 88, No. 1, February 2015

MAA

# MATHEMATICS MAGAZINE




- Polynomials with closed Lill paths
- Visualizing curvature and torsion with bicycle wheels and ribbons
- The median value of a continuous function
- 2D, or not 2D, that is the crossword

## EDITORIAL POLICY

*Mathematics Magazine* aims to provide lively and appealing mathematical exposition. The *Magazine* is not a research journal, so the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for the Magazine. Articles should include examples, applications, historical background, and illustrations, where appropriate. They should be attractive and accessible to undergraduates and would, ideally, be helpful in supplementing undergraduate courses or in stimulating student investigations. Manuscripts on history are especially welcome, as are those showing relationships among various branches of mathematics and between mathematics and other disciplines.

Submissions of articles are required via the *Mathematics Magazine*'s Editorial Manager System. The name(s) of the author(s) should not appear in the file. Initial submissions in pdf or LaTeX form can be sent to the editor at www.editorialmanager.com/mathmag/.

The Editorial Manager System will cue the author for all required information concerning the paper. Questions concerning submission of papers can be addressed to the editor at mathmag@maa.org. Authors who use LaTeX are urged to use the Magazine article template. However, a LaTeX file that uses a generic article class with no custom formatting is acceptable. The template and the Guidelines for Authors can be downloaded from www.maa.org/pubs/mathmag.

*MATHEMATICS MAGAZINE* (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, NW, Washington, DC 20036 and Lancaster, PA, in the months of February, April, June, October, and December.

Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

Address advertising correspondence to

MAA Advertising
1529 Eighteenth St. NW
Washington, DC 20036
Phone: (202) 319-8461
E-mail: advertising@maa.org

Further advertising information can be found online at www.maa.org.

Change of address, missing issue inquiries, and other subscription correspondence can be sent to:

The MAA Customer Service Center
P.O. Box 91112
Washington, DC 20090-1112
(800) 331-1622
(301) 617-7800
maaservice@maa.org



Periodicals postage paid at Washington, D.C. and additional mailing offices.

Postmaster: Send address changes to Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, NW, Washington, DC 20036-1385.

Printed in the United States of America.

## COVER IMAGE

**Snowflake Calligram** by David A. Reimann, 2014

This snowflake calligram celebrates winter with the natural beauty found in the six-fold symmetry of snowflakes. Each arm of the snowflake contains the text "Mathematics Magazine" at three sizes.

Vol. 88, No. 1, February 2015

MAA

# MATHEMATICS MAGAZINE

# ARTICLES

## Polynomials with Closed Lill Paths

DAN KALMAN
American University
Washington, D.C. 20016
kalman@american.edu

MARK VERDI
Arlington, VA
mark.s.verdi@gmail.com

Lill's method is a striking 19th century technique for discovering real roots of polynomials with real coefficients. It is illustrated for the polynomial $p(x) = 4x^4 + 6x^3 + 5x^2 + 4x + 1$ in FIGURE 1. In the figure there are two rectangular paths. Along the outer path, shown in black, the lengths of the successive sides are given by the coefficients of $p$. By convention the path begins at the origin with its first leg along the positive $x$ axis. We call this the *Lill path* of $p$. The inner (gray) path also begins at the origin, its first leg making an angle $\theta$ with the positive $x$ axis and with vertices on successive legs of the outer path. If the final vertices of the inner and outer paths coincide, as they do in the figure, then one root of $p$ is given by $x = -\tan\theta$. While this appears totally unexpected at first exposure, the reader may dispel a bit of the mystery by considering the case $p(x) = ax + b$ with $a$ and $b$ positive.

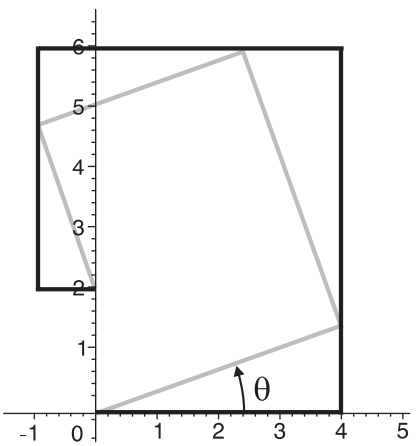



**Figure 1** Lill's method for $p(x) = 4x^4 + 6x^3 + 5x^2 + 4x + 1$.

Hull [**3**] provides dramatic imagery for the configuration of the two paths. Imagine that we are visiting a universe with a strangely modified law of reflection: the angle of incidence and the angle of reflection are *complementary*, rather than equal as in our

*Math. Mag.* **88** (2015) 3–10. doi:10.4169/math.mag.88.1.3. 
MSC: Primary 12Y05.

universe. Then the inner path is the trajectory of a projectile fired from the origin. Our goal is to aim in just the right direction $\theta$ so that the projectile ricochets off all the walls of the outer path and hits the final vertex. If we succeed, we obtain a root of $p$ as $-\tan\theta$.

In itself, Lill's method is an amazing result. But by connecting the geometry of paths and the algebra of polynomials in an unexpected way, it is also a springboard to many other insights and observations. For example, as recounted in Hull's paper, Beloch applied Lill's method to show that real roots of cubic equations can always be obtained using origami constructions. More generally, it is natural to explore connections between algebraic properties of a polynomial and geometric properties of its Lill path. Here we investigate the properties of polynomials whose Lill paths are closed.

Lill's interest in the method was apparently for finding approximations to roots of polynomials. That is certainly the interest expressed by Bixby, who published a pamphlet in 1879 in order to bring the method to the notice of American mathematicians and engineers. In a later publication [**1**], he explained his reasons for producing the pamphlet, describing Lill's method thus:

> From my own personal experience in such matters, [it] ... is the best graphical method yet developed, and far easier, quicker, and more exact, than any other graphical or mechanical method.

For more information about Lill's method and its history, see [**2**], [**3**], [**4**, pp. 13–24], and [**6**]. Animated demonstrations are provided at [**2, 5**].

**Polynomials with closed Lill paths.** Suppose a polynomial has a closed Lill path. That is, suppose the initial vertex and the final vertex coincide. What does that imply?

In the geometrically simplest case, a closed path can simply be a single line segment that is traced an even number of times. Examples include $ax^2 + a$, $ax^4 - a$, and $ax^6 + ax^4 + ax^2 + a$. But these are degenerate cases. If we insist on a path that is not confined to a line, then the simplest closed Lill path is a rectangle, as shown in FIGURE 2. That means $p(x)$ must have the form $ax^3 + bx^2 + ax + b$ for some $a$ and $b$. We see immediately that $p(x)$ factors as $(x^2 + 1)(ax + b)$. We can also obtain a rectangle using 5 edges, as in FIGURE 3, and the path polynomial is $p(x) = ax^4 + bx^3 + cx^2 + bx + e$. But as the symmetry of the rectangle shows that $c = e + a$, we can write

$$p(x) = ax^4 + bx^3 + ex^2 + ax^2 + bx + e = (x^2 + 1)(ax^2 + bx + e).$$

We invite the reader to apply similar reasoning for the path in FIGURE 4 to show the path polynomial is $(x^2 + 1)(ax^3 + bx^2 + ex + f)$.

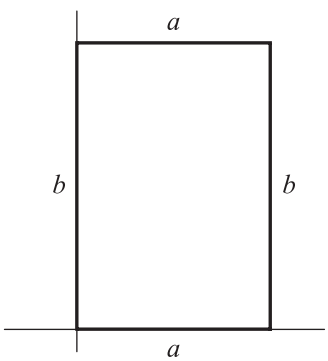



**Figure 2** A rectangular Lill path.

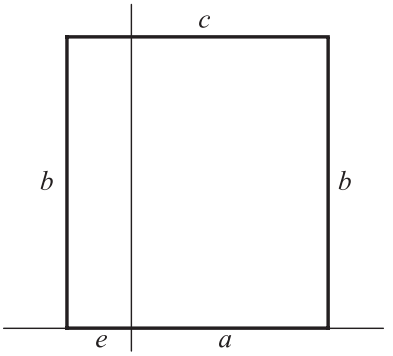



**Figure 3** Closed path for a quartic.

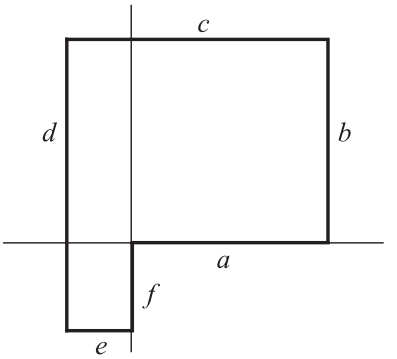



**Figure 4** Closed path for a quintic.

So far, we have considered only polynomials with positive coefficients. With one or more negative coefficients the method still applies, but we require a more general definition of the Lill path. The legs of a Lill path are oriented, with positive and negative directions. A convenient way to specify the positive directions is to use the unit vectors **i** and **j**, which point in the directions of the positive $x$ and $y$ axes, respectively. Then the positive directions on successive legs of the Lill path follow the cyclic pattern **i**, **j**, $-$**i**, $-$**j**, **i**, **j**, $-$**i**, $-$**j**, .... To construct the Lill path, for a coefficient $c$ we draw a line of length $|c|$ in the positive direction for positive $c$ and in the negative direction for negative $c$.

Now consider FIGURE 5. The third and fourth legs are oriented negatively, so the corresponding coefficients of the path polynomial are negative. In the figure, we have labeled each leg with its length, which is the opposite of the corresponding coefficient for legs three and four. With this notation, the path polynomial is $p(x) = ax^7 + bx^6 + cx^5 + dx^4 + ex^3 + fx^2 + gx + h$. Note that $g = a - c + e$ and $h = b - d + f$. Therefore we can write

$$p(x) = (ax + b)(x^6 + 1) + (cx + d)(x^4 - 1) + (ex + f)(x^2 + 1),$$

and since $x^2 + 1$ is a divisor of both $x^6 + 1$ and $x^4 - 1$ we again find that $x^2 + 1$ is a factor of the path polynomial.

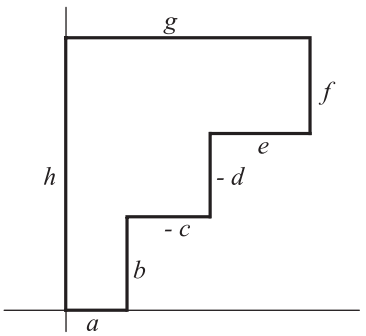



**Figure 5** Some coefficients of $p$ are negative.

By now it seems pretty clear that a polynomial with a closed Lill path should be divisible by $x^2 + 1$. That was observed in every case considered above, including the degenerate ones. But perhaps these examples are special in some way. Allowing self-intersection and various sign patterns for the coefficients, many different kinds of

closed paths can be constructed (see FIGURE 6). The reader may wish to show that in each of these cases the path polynomial is divisible by $x^2 + 1$, or to formulate and analyze other paths.

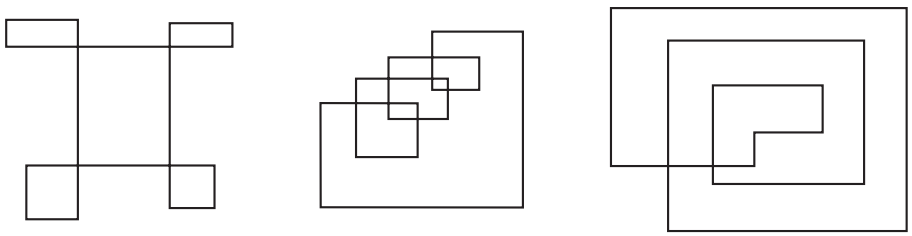


**Figure 6** Sample closed Lill paths.

Although the calculations vary from one example to another, they all point toward one conclusion:

**Closed Lill Path Theorem.** *The polynomial $p(x)$ has a closed Lill path if and only if $p(x)$ is divisible by $x^2 + 1$.*

Before proving this result, it will be algebraically simpler if we make one initial observation: the polynomial $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ has a closed Lill path if and only if the *reverse polynomial* $\text{rev}p(x) = a_n + a_{n-1}x + \cdots + a_0 x^n$ has a closed Lill path. For any polynomial $p$, the Lill path for the reverse polynomial has the same sequence of leg-lengths as the path for $p$, but in the opposite order. That is essentially the same as starting at the end of $p$'s Lill path and traversing it in the opposite direction. However, tracing $p$'s Lill path in reverse, the initial point is not necessarily at the origin, the initial direction is not necessarily that of **i**, and the right angle turns would all be made in the opposite directions. Thus, if in the forward direction, successive legs of the path are parallel to **i** and **j**, so that the path turns counter-clockwise, then in reverse the path would go from **j** to **i**, thus turning clockwise. So the Lill path for the reverse polynomial is not identical to the path for $p$ in reverse. But it seems apparent that the two paths are congruent.

We can argue the point somewhat more formally as follows. The Lill path for an arbitrary polynomial $p$ can be transformed into the Lill path for $\text{rev}p$ in three steps. First, translate the final vertex to the origin, $O$, if it is not at $O$ already. Second, rotate about $O$ (if necessary) so that the edge for $a_0$ (that is, the final edge in the original path) extends along the $x$ axis in the proper direction, parallel to **i** if $a_0 \geq 0$ and to $-\mathbf{i}$ otherwise. Third, reflect across the $x$ axis, leaving the $a_0$ leg fixed, but reversing the directions of all the right angle turns. Tracing the resulting path starting from the origin is identical to tracing the Lill path for $\text{rev}p$. And since translations, rotations, and reflections are isometries, the Lill path for $p$ is congruent to the Lill path for $\text{rev}p$. Consequently, either both are closed or neither is.

Next we reformulate the closed Lill path theorem as follows: The polynomial $p$ has a closed Lill path if and only if $i$ is a root of $\text{rev}p$. Here, in addition to the foregoing argument about $p$ and $\text{rev}p$, we are using the fact that $p$ has real coefficients. Thus $p(i) = 0$ iff $p(-i) = 0$ iff $p(x)$ is divisible by $(x - i)(x + i) = x^2 + 1$.

Now the theorem is reduced to a calculation. If $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ then $\text{rev}p(i) = a_n + a_{n-1}i + a_{n-2}i^2 + \cdots + a_0 i^n$. Let us envision this sum in the complex plane, considering each term $a_{n-k}i^k$ as a vector. To add them we start the initial term at the origin, at its end append the second term, then the third, and proceed

in this fashion until all the terms are strung out, with each term starting at the end of the preceding term. The vector sum is indicated by the terminal point of the final term, so $\mathrm{rev}p(i) = 0$ precisely when the final term ends at the origin. But the configuration of this sum of vectors is exactly the Lill path for $p$, because the successive powers of $i$, as vectors in the complex plane, form the sequence $\mathbf{i}, \mathbf{j}, -\mathbf{i}, -\mathbf{j}, \mathbf{i}, \mathbf{j}, -\mathbf{i}, -\mathbf{j}, \ldots$. Moreover, the condition for the Lill path to be closed is that the final vertex of the path is the origin. Thus we have seen that $\mathrm{rev}p(i) = 0$ if and only if the Lill path for $p$ is closed, which is what we wished to show.

**Angling for a generalization.** As observed in [**4**], Lill's result does not require rectangular paths. With a slight modification, it works equally well using inner and outer paths at which successive legs meet with a constant exterior angle $\phi$. As an illustration, FIGURE 7 displays two path configurations for $p(x) = 4x^4 + 6x^3 + 5x^2 + 4x + 1$. The first is the usual Lill diagram with $\phi = 90^\circ$. In the second $\phi = 80^\circ$, so that the (interior) angles between successive legs all equal $100^\circ$. Notice that in this arrangement, each triangle has one angle equal to $\theta$, one equal to $\pi - \phi$ (the interior angle at a vertex of the outer path), and one angle equal to $\phi - \theta$. In particular, the adjacent angles in successive triangles sum to $\phi$, just as they sum to $\pi/2$ in the standard Lill diagram.

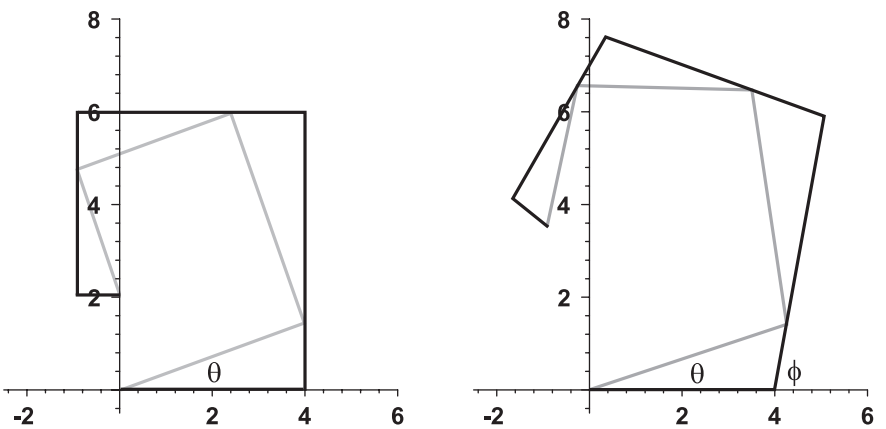



**Figure 7** Lill paths with $\phi = 90^\circ,\ 80^\circ$.

As the figure suggests, we can deform the first configuration to the second by enlarging each interior angle in the original Lill path. To visualize this, imagine that the legs of the outer path are hinged at the vertices, that the vertices of the inner path are attached rigidly to the outer path legs, and that the inner path legs are free to stretch or contract. Opening up each hinge by $10^\circ$ then transforms the first configuration to the second.

In the new configuration, we can still obtain a root of $p(x)$ from the diagram, this time as $x = -\sin(\theta)/\sin(\phi - \theta)$, where $\theta$ is again the angle between the first legs of the inner and outer paths. Thus, there is an associated $\phi$-Lill's method in which paths are constructed with exterior angle $\phi$, and we again obtain a root when the ends of the inner and outer paths coincide. For a proof of this result, see [**4**].

In this context, what is the significance of a closed $\phi$-Lill path? Apparently, if the $\phi_1$-Lill path is closed, then for a slightly different angle $\phi_2$ the $\phi_2$-Lill path will not be closed. So the polynomials with closed $\phi_1$-Lill paths will be different from those

with closed $\phi_2$-Lill paths. Is there an extension of the closed Lill path theorem to these modified Lill paths?

Again we can begin by looking at examples. With $\phi = 120^\circ$ the simplest nondegenerate closed path is an equilateral triangle. The path polynomial is $a(x^2 + x + 1)$ where each leg of the triangle has length $a$. More complicated examples can be analyzed using the geometry of parallelograms and equilateral triangles. For example, in FIGURE 8, we display a closed Lill path for a sixth degree polynomial with positive coefficients. The origin is marked with a filled square and the other vertices are indicated with filled circles, while intersections of edges are left unmarked. Letting $a$ be the side of the small central equilateral triangle, and $b$ the length of the first edge, we proceed to label all the line segments making up the figure, introducing additional parameters as needed. The path polynomial for this figure is thus

$$\begin{aligned} p(x) &= bx^6 + (a+b)x^5 + (a+b+c)x^4 + (a+c)x^3 \\ &\quad + (a+c+d)x^2 + (a+d)x + a + d \\ &= a(x^5 + x^4 + x^3 + x^2 + x + 1) + b(x^6 + x^5 + x^4) \\ &\quad + c(x^4 + x^3 + x^2) + d(x^2 + x + 1) \\ &= a(x^3 + 1)(x^2 + x + 1) + (bx^4 + cx^2 + d)(x^2 + x + 1), \end{aligned}$$

and so is evidently divisible by $x^2 + x + 1$. This and similar examples lead us to conjecture that a polynomial $p$ has a closed $120^\circ$-Lill path if and only if $p(x)$ is a multiple of $x^2 + x + 1$. By the same logic as before, these are polynomials having a root of $(-1 + i\sqrt{3})/2$.

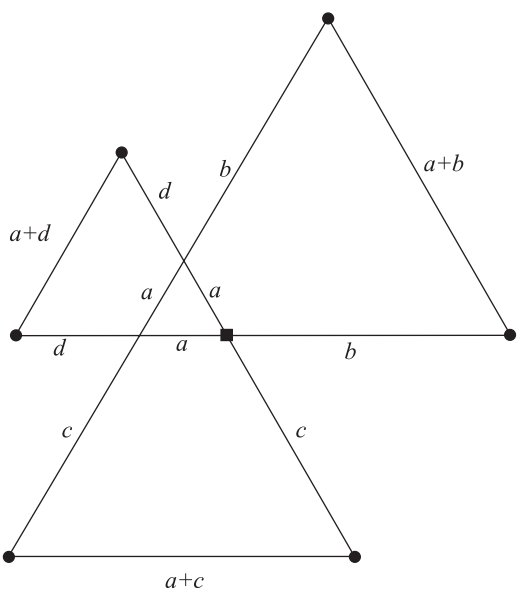



**Figure 8** A closed $120^\circ$-Lill path.

There is a clear analogy between this situation and the one we considered earlier. There, we looked at paths with right-angle turns, and polynomials with $i$ as a root. And in exponential form, we have $i = e^{i\pi/2}$, showing that multiplication by $i$ imposes a rotation through $\pi/2$ in the complex plane. Now we are interested in paths with turns of size $2\pi/3$, and polynomials having a root of $(-1 + i\sqrt{3})/2$. Again we can write the

root in exponential form, $e^{i(2\pi/3)}$, so that multiplying by this root imposes a rotation of $2\pi/3$ in the complex plane.

The general phenomenon is now apparent. For any angle $\phi$, we expect a real polynomial to have a closed $\phi$-Lill path iff $e^{i\phi} = \cos\phi + i\sin\phi$ is a root. In this case, since the roots come in complex conjugate pairs, the polynomial would be divisible by both $(x - \cos\phi - i\sin\phi)$ and $(x - \cos\phi + i\sin\phi)$, leading to the following formulation.

**Closed $\phi$-Lill Path Theorem.** *The polynomial $p(x)$ has a closed $\phi$-Lill path if and only if $p(x)$ is divisible by $x^2 - 2(\cos\phi)x + 1$.*

This result is easily proved by modifying the earlier proof in an obvious way. We can still work with the reversed polynomial, and interpret the $\phi$-Lill path as a vector diagram in the complex plane showing the evaluation of $p(e^{i\phi})$. However, there is an equivalent argument using rotation matrices that is also worth considering.

We now consider a rotation of the (real) plane as a linear transformation, and hence as multiplication by a matrix. Representing vectors as columns, a counter-clockwise rotation through an angle $\phi$ is represented by the matrix

$$R = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}.$$

In the argument to follow, we will be interested in the *minimal polynomial* of $R$. This is a monic polynomial $m$ of minimal degree for which $m(R) = \mathbf{0}$, the zero matrix. Moreover, it is known that $q(R) = \mathbf{0}$ for a polynomial $q$ if and only if $m(x)$ is a divisor of $q(x)$. Indeed, if division of $q(x)$ by $m(x)$ leaves a nonzero remainder $r(x)$, then $m$ is a divisor of $q - r$ and $q(R) = r(R)$ follows. But $r(R)$ cannot be zero because $r$ has lower degree than $m$. Thus, $q(R)$ can be zero only when there is no remainder $r$.

The minimal polynomial of $R$ can only be a linear polynomial if $R$ is a multiple of the identity matrix, and that occurs only when $\phi$ is an integer multiple of $\pi$. Otherwise, the minimal polynomial must be at least quadratic. On the other hand, every square matrix is annihilated by its own characteristic polynomial (that's the Cayley-Hamilton theorem), so for a $2 \times 2$ matrix the minimal polynomial can be at most quadratic. We see therefore that when $\phi$ is not an integer multiple of $\pi$, the minimal polynomial of the rotation matrix $R$ is precisely the characteristic polynomial

$$\det(xI - R) = \det\begin{bmatrix} x - \cos\phi & \sin\phi \\ -\sin\phi & x - \cos\phi \end{bmatrix} = x^2 - 2(\cos\phi)x + 1.$$

Proceeding to the main argument, let $p(x) = a_n x^n + \cdots + a_0$, and suppose $\phi$ is not an integer multiple of $\pi$. The legs of the $\phi$-Lill path are, in order, $a_n\mathbf{i}$, $a_{n-1}R\mathbf{i}$, $a_{n-2}R^2\mathbf{i}$, and so on. Starting from the origin, the final vertex is thus the vector sum of these legs, and hence equal to $\mathrm{rev}\,p(R)\mathbf{i}$. The path is closed if and only if this vector sum is zero.

Next observe that

$$\begin{aligned} \mathrm{rev}\,p(R)\mathbf{i} = 0 \quad &\text{iff} \\ R\,\mathrm{rev}\,p(R)\mathbf{i} = 0 \quad &\text{iff} \\ \mathrm{rev}\,p(R)R\mathbf{i} = 0 \quad & \end{aligned}$$

as $R$ and $\mathrm{rev}\,p(R)$ evidently commute. Furthermore, the vectors $\mathbf{i}$ and $R\mathbf{i}$ are linearly independent because $\phi$ is not an integer multiple of $\pi$. Thus, $\mathrm{rev}\,p(R)\cdot\mathbf{i}$ and $\mathrm{rev}\,p(R)\cdot R\mathbf{i}$ both vanish if and only if $\mathrm{rev}\,p(R)$ is the zero matrix. And this in turn occurs if and only if $\mathrm{rev}\,p(x)$ is a multiple of the minimal polynomial of $R$, namely, $x^2 - 2\cos\phi x + 1$.

We have now seen that $p$ has a closed $\phi$-Lill path if and only if $\operatorname{rev} p$ is a multiple of $x^2 - 2\cos\phi x + 1$. Applying this result to $\operatorname{rev} p$ shows that it has a closed $\phi$-Lill path if and only if $p$ is a multiple of $x^2 - 2\cos\phi x + 1$. But $p$ has a closed $\phi$-Lill path if and only if $\operatorname{rev} p$ does. Thus $p$ has a closed $\phi$-Lill path if and only if it is a multiple of $x^2 - 2\cos\phi x + 1$, as we wished to show.

**Additional questions.** There remain other aspects of closed Lill paths to explore. For example, a closed path remains closed under a cyclic permutation of the edges. In other words, if we traverse the path starting from a different vertex, we will still return to the starting point. In this way we define a new path polynomial. How is it related to the original? What algebraic conclusions can be drawn from this operation? We have already seen that polynomials with closed Lill paths are reversible. If they are also cyclically permutable, we obtain a dihedral group of path reorderings. What implications arise from this?

Going in another direction, what can be said about non-planar paths? Here, the idea of using rotation matrices rather than complex numbers seems particularly appropriate. For a fixed $3 \times 3$ rotation matrix $R$, create a path in $\mathbb{R}^3$ whose legs take the form $a_k R^k \mathbf{i}$. Is there a version of Lill's theorem in this setting? Does a closed path here have an analogous meaning to the one we found in the plane?

These are just a few possible areas to investigate. No doubt the reader can propose others.

**Acknowledgments** We thank Francis Su for suggesting this topic for investigation, and the American University Department of Mathematics and Statistics for supporting Verdi's summer research. After discovering the closed Lill path theorem, we realized that something very similar had been discussed in [**2**]. The extension to $\phi$-Lill paths is, as far as we know, original.

## REFERENCES

1. W. H. Bixby, *Graphical Method for Finding the Real Roots of Numerical Equations of Any Degree if Containing but One Variable*, West Point, 1879, http://www.maa.org/ume.
2. P. V. Bradford, *Visualizing Solutions to n-th Degree Algebraic Equations Using Right-Angle Geometric Paths: Extending Lill's Method of 1867*, http://web.archive.org/web/20130926234416/http://www.concentric.net/~pvb/ALG/rightpaths.html.
3. T. C. Hull, Solving cubics with creases: The work of Beloch and Lill, *Amer. Math. Monthly* **118** no. 4 (2011) 307–315, http://dx.doi.org/10.4169/amer.math.monthly.118.04.307.
4. D. Kalman, *Uncommon Mathematical Excursions: Polynomia and Related Realms*. MAA, Washington, DC, 2009.
5. D. Kalman, *Uncommon Mathematical Excursions: Polynomia and Related Realms: Supplementary Resources*. http://www.maa.org/ume.
6. Wikipedia contributors, Lill's method, *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Lill's_method.

**DAN KALMAN** (MR Author ID: 198950) received his Ph.D. from the University of Wisconsin in 1980. Before joining the mathematics faculty at American University in 1993, he worked for eight years in the aerospace industry in Southern California. Kalman is a past associate executive director and MD-DC-VA Governor of the MAA, author of two books published by the MAA, and frequent contributor to MAA journals. He delights in puns and word play of all kinds.

**MARK VERDI** (MR Author ID: 1074407) graduated from American University in 2014 with a degree in Applied Mathematics. He delved into Lill's theorem as a sophomore and remains an enthusiast of the aesthetics, applications, and history of mathematics. Mark has also performed research in oceanography and materials science, and he now works as a software engineer in Arlington, Virginia.

# Bend, Twist, and Roll: Using Ribbons and Wheels to Visualize Curvature and Torsion

FRED KUCZMARSKI
Shoreline Community College
Shoreline, WA 98133
fkuczmar@shoreline.edu

You are in a park with your bicycle, standing next to a slide whose center line is in the form of a cylindrical helix. Can you use the front wheel to measure the total curvature and torsion of the helix?

Before answering this question, we review curvature and torsion, emphasizing visual ways to think about these concepts. We introduce some simple but insightful constructions that played a role in the historical development of the subject but are seldom found in introductory texts today. These include such exotic sounding ideas as the hodograph, the indicatrices of a curve, and the Darboux vector. We explain these in due course, but for now we point out that they have one aspect in common; they all give visual and dynamic (in the sense of motion) ways to think about curvature and torsion.

We also use ribbons and bicycle wheels. For us, a ribbon is a surface made up of line segments. Constructing ribbons around curves gives dramatic ways to visualize curvature and torsion. The inspiration to use a bicycle wheel to measure curvature and torsion comes from Levi [**4**], who showed how to measure geodesic curvature with a bicycle wheel.

One note before we start. We assume all our curves come with a parameterization by arc length $\gamma : [0, \ell] \to \mathbb{R}^n$, with $n = 2$ or 3, and are *smooth* or sufficiently differentiable to allow the constructions that follow. At times, we find it helpful to distinguish between the function $\gamma$ and the *trace*, or the set of points $\Gamma = \{\gamma(s), s \in [0, \ell]\}$. So we might say "the vector $\mathbf{T}$ is tangent to $\Gamma$" to conjure up a picture of the trace with an attached tangent vector.

## The curvature of plane curves

It is 1975 and you are driving the back roads of Seattle. GPS has not yet been invented, but you have the latest technology, a compass glued to your dashboard. At a bend in the road, you simultaneously look at the compass needle and your speedometer. With some quick calculations, you determine that at that particular instant your speed is 44 ft/sec and the car is turning to the left at the rate of 22 deg/sec. So at that point $P$, the road bends at the rate of

$$(22 \text{ deg/sec}) \,/\, (44 \text{ ft/sec}) = 0.5 \text{ deg/ft} = \frac{\pi}{360} \text{ rad/ft}.$$

You have just calculated the *curvature* of the road at $P$. The curvature measures the rate at which the forward-pointing unit tangent $\mathbf{t}(s) = \gamma'(s)$ to $\gamma$ changes direction relative to arc length. (The prime denotes differentiation with respect to $s$ throughout the paper.) To better see how $\mathbf{t}$ changes direction, we cluster the unit tangent vectors by

*Math. Mag.* **88** (2015) 11–26. doi:10.4169/math.mag.88.1.11. 
MSC: Primary 53A04.

parallel translation so that their tails come together at a point $O$. Doing so removes the translational motion of the vectors and leaves their rotational motion. FIGURE 1 shows an example, where the tangent vectors are spaced at equal intervals of arc length along a logarithmic spiral.

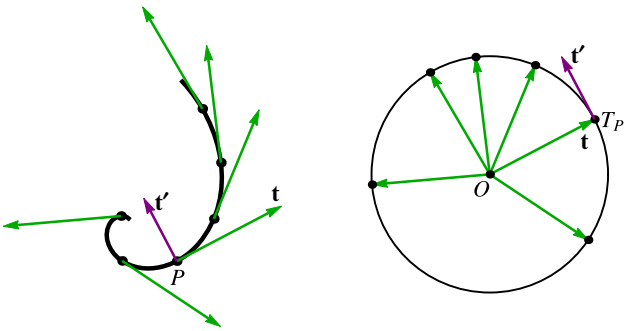



**Figure 1** The tangent cluster of a logarithmic spiral.

As $P = \gamma(s)$ moves along $\Gamma$, the corresponding tip $T_P$ of the clustered tangent vector moves along the unit circle $S^1$ centered at $O$. If we think of $s$ as denoting time, so that $P$ moves along $\Gamma$ with unit speed, then $T_P$ moves along $S^1$ with varying speed. Later we give the curvature $\kappa$ a sign, but for now, we define the absolute value of the curvature as the speed of $T_P$ so that $|\kappa| = |\mathbf{t}'|$. Since $T_P$ moves along a circle of unit radius, $|\kappa|$ measures the rate at which the unit tangent rotates with respect to $s$. The vector $\mathbf{t}' = \gamma''$, the *curvature vector*, may be thought of as the acceleration of a particle traversing the curve with unit speed. Differentiating the equation $\mathbf{t} \cdot \mathbf{t} = 1$ gives $\mathbf{t} \cdot \mathbf{t}' = 0$ so that $\mathbf{t}'$ is perpendicular to $\mathbf{t}$.

The idea to cluster the velocity vectors of a curve is due to William Hamilton (1806–1865) of quaternion fame. In his proof that the planets orbit the sun in elliptical orbits, he defined the *hodograph* as the curve traced by the tips of the clustered velocity vectors [**1**]. For a particle moving with constant speed, the hodograph is a circular arc. FIGURE 1 shows that for such a particle the acceleration and velocity vectors are always orthogonal.

Most calculus texts (cf. [**5**, p. 855]) force the curvature to be non-negative by defining it as $|\mathbf{t}'|$. We would like to give the curvature a sign so that a curve lying in a plane $\Pi$ embedded in space has positive curvature if it bends to the left and negative curvature otherwise. The sign depends on two arbitrary choices—the arc length parameterization that defines the forward direction (akin to putting a one-way sign on our Seattle road) and the viewpoint of the observer. What looks like a left turn to you will look like a right turn to me if we are on opposite sides of $\Pi$. We fix our viewpoint by choosing a constant unit vector $\mathbf{b}$ normal to $\Pi$, the *binormal*, and agree that we look at $\Gamma$ from the $\mathbf{b}$-side of $\Pi$, or the side into which $\mathbf{b}$ points.

In most textbooks, the *principal normal* $\mathbf{n}$ is defined as the unit vector in the direction of $\mathbf{t}'$ so that $\mathbf{t}' = \kappa\mathbf{n}$ (with $\kappa \geq 0$) and

$$\kappa = \mathbf{t}' \cdot \mathbf{n}. \tag{1}$$

The principal normal $\mathbf{t}'/|\mathbf{t}|$ points toward the concave (inward) side of $\Gamma$. But it is undefined at points where $\mathbf{t}' = \mathbf{0}$ and abruptly reverses direction at inflection points. Instead we define the smooth unit *normal* vector $\mathbf{n} = \mathbf{n}(s) = \mathbf{b} \times \mathbf{t}(s)$. This vector always points to our left as we move along the curve and may point toward the convex (outward) side. We still use (1) to define the curvature. Our definition agrees with the

textbook definition up to sign. At points where $\mathbf{n}$ and $\mathbf{t}'$ are parallel, $\Gamma$ turns left and $\kappa > 0$. At points where $\mathbf{n}$ and $\mathbf{t}'$ are antiparallel, $\Gamma$ turns right and $\kappa < 0$. FIGURE 2 (adapted from [**6**, p. 15]) shows an example.

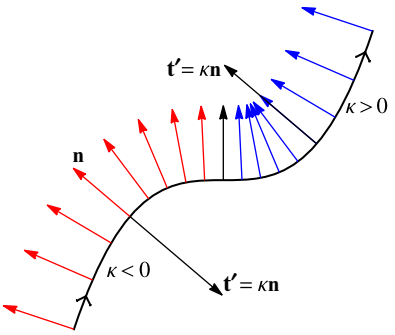



**Figure 2** Giving the curvature a sign.

## The curvature of space curves

The textbook definitions of the curvature and principal normal for a space curve are identical to those for a plane curve and come with the same problems. In these textbooks, the curvature is not signed and the principal normal is undefined at the zeros of $\gamma'' = \mathbf{t}'$, where it may abruptly change direction. Instead we first try to define our viewpoint by finding a smooth unit binormal $\mathbf{b} = \mathbf{b}(s)$, now a function of $s$, normal to both $\mathbf{t}(s)$ and $\mathbf{t}'(s)$. For a plane curve $\gamma$, this is always possible; we just choose $\mathbf{b}$ to be one of the two unit vectors normal to plane of $\Gamma$. But even for a $C^\infty$ space curve such a choice of $\mathbf{b}$ is not always possible. The $C^\infty$ curve

$$\gamma(t) = \begin{cases} (t, t^3, 0), & t \le 0 \\ (t, 0, t^3), & t > 0 \end{cases}$$

shows why. It can be reparameterized by arc length but does not admit a continuous vector field $\mathbf{b}$ that is normal to both $\mathbf{t}$ and $\mathbf{t}'$. Because $\mathbf{b}$ is normal to the $xy$-plane if $t < 0$ and normal to the $xz$-plane if $t > 0$, $\mathbf{b}$ cannot be continuously defined on any open interval containing the origin. The problem is that the principal normal does not *reverse* direction but rather rotates through an angle of $90^\circ$ at the isolated zero of $\gamma''$.

From now on, we assume $\gamma$ admits a smooth unit binormal $\mathbf{b}$ normal to both $\mathbf{t}$ and $\mathbf{t}'$. The existence of one such vector field $\mathbf{b}$ guarantees the existence of a second, namely $-\mathbf{b}$. To ensure there are no others, we assume the zeros of $\mathbf{t}' = \gamma''$ are isolated. In particular, $\Gamma$ contains no straight line segments.

We now proceed exactly as for plane curves and define $\mathbf{n} = \mathbf{b} \times \mathbf{t}$ and the curvature $\kappa = \mathbf{t}' \cdot \mathbf{n}$ as before. The curvature agrees with the textbook definition up to sign. The normal $\mathbf{n}$ agrees with the principal normal $\mathbf{N}$ up to sign at points where the latter is defined. We call the orthonormal frame $\{\mathbf{b}, \mathbf{t}, \mathbf{n}\}$ the *modified* Frenet frame, or just the Frenet frame. Unlike the *classical* Frenet frame $\{\mathbf{t}, \mathbf{N}, \mathbf{t} \times \mathbf{N}\}$, the Frenet frame is defined at points where $\mathbf{t}' = \mathbf{0}$ and varies continuously along the curve.

As $P$ moves along $\Gamma$ with unit speed, the tip $T_P$ of the clustered tangent vector moves along a curve $\Gamma_t$, the *tangent indicatrix*, or *tantrix* for short. The tantrix lies on the unit sphere $S^2$ and is just Hamilton's hodograph for a unit speed space curve.

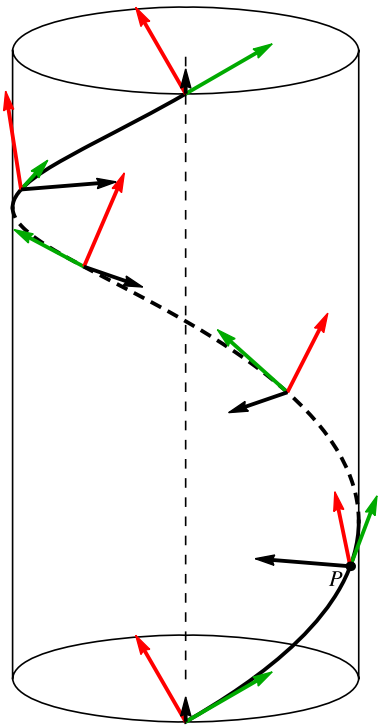



**Figure 3** A Frenet frame along a helix.

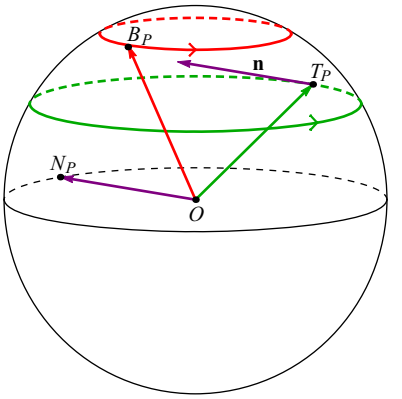



**Figure 4** Measuring the total curvature of a cylindrical helix.

The speed $|\mathbf{t}'|$ of $T_P$ records the rotation rate of $\mathbf{t}$ and gives the absolute value of the curvature. We define the forward direction along the tantrix by the vector $\mathbf{n}$ so that the curvature (1) is positive (negative) when $T_P$ moves along the tantrix in the forward (backward) direction.

FIGURE 3 shows a right cylindrical helix with a Frenet frame. In FIGURE 4, we cluster the frame by parallel translating the frame vectors so that their tails come together at $O$. As $P$ moves along $\Gamma$, the tips $N_P$ and $B_P$ of the clustered normal and binormal vectors trace out curves on $S^2$, the *normal* and *binormal indicatrices*. For now, we focus on the tantrix and the normal indicatrix. Since the normals to the helix point directly toward its axis, the normal indicatrix is a great circle that we take to be the equator of $S^2$. The tantrix is then a circle of latitude in the northern hemisphere. As the helix is traversed with unit speed, $T_P$ moves counterclockwise (as seen from the north pole) along the tantrix. There are several ways to see that the curvature of the helix is positive:

(a) The figure illustrates that the forward direction along the tantrix given by the vector $\mathbf{n}$ is also counterclockwise.

(b) Since $\mathbf{n}$ is the principal normal, $\mathbf{t}'$ is parallel to $\mathbf{n}$ and by (1) $\kappa > 0$.

(c) The helix clearly turns to the left from our viewpoint at the tip of $\mathbf{b}$.

The integral $\int_a^b \kappa(s)\, ds$ gives the *total curvature* of $\boldsymbol{\gamma}$ between $\gamma(a)$ and $\gamma(b)$. For a plane curve, the total curvature measures the directed angle of rotation from $\mathbf{t}(a)$ to $\mathbf{t}(b)$. For a space curve, the total curvature is still the sum (or integral) of the differential angles $d\theta = \kappa(s)\, ds$ between consecutive tangent vectors, but since the tangent vectors do not lie in a plane, the angle between $\mathbf{t}(a)$ and $\mathbf{t}(b)$ is not the total

curvature. We can visualize the total curvature by forming the *tangent cone* with vertex $O$ and generators $OT$ through the points $T$ of the tantrix. Rolling the tangent cone on a plane gives a circular sector whose radii $O'T'$ correspond to the generators $OT$. The rolling motion transfers the differential angles on the surface of the cone to the plane. The central angle of the sector thus gives the total curvature. Or we can simply flatten the tangent cone after cutting it open along the generator $\mathbf{t}(a)$. We illustrate this in FIGURE 5 for the tangent cone of a cylindrical helix.

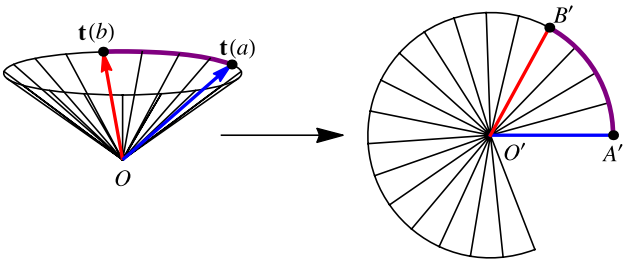



**Figure 5** Developing the tangent cone of a cylindrical helix on a plane.

We may also interpret the total curvature as the signed arc length of the tantrix from $\mathbf{t}(a)$ to $\mathbf{t}(b)$. For example, suppose the tangents to the helix of FIGURE 3 make an angle of $\phi$ with its axis. Then the tantrix is the circle of latitude $\pi/2 - \phi$. The total curvature of one turn of the helix is the circumference of this circle, or $2\pi \sin\phi$. If $\phi = \pi/2$, then the helix degenerates to a circle with total curvature $2\pi$.

The *curvature inflections* of a space curve, or $\kappa$*-inflections* for short, are the points where the curvature changes sign. At these points, $T_P$ stops its motion along the tantrix and reverses direction. The inflection points thus correspond to cusps on the tantrix. The tantrix is not differentiable at a cusp but still has a well-defined tangent line there. For a curve confined to a plane, a cusp occurs at an endpoint of an arc of a great circle as the arc folds back on itself.

FIGURE 6 shows an example. The cusp of the tantrix corresponds to the $\kappa$-inflection $Q(0, 0, 0)$ of the curve $\gamma(t) = (t^2, t^3, t^4)$, $-1 \le t \le 1$. The arrows on the tantrix point in the forward direction as determined by the vector $\mathbf{n}$. Since $\mathbf{n}$ is continuous, the forward directions either converge to or diverge from a cusp. In our example the forward directions diverge from the cusp. As $\Gamma$ is traversed in the forward direction from $P(\gamma(-0.65))$ to $Q$ to $R(\gamma(0.5))$, the corresponding point of the tantrix moves from $T_P$ to $T_Q$ to $T_R$. The curvature is thus negative at $P$ and positive at $R$. Since the tantrix is symmetric about its cusp, the total curvature vanishes.

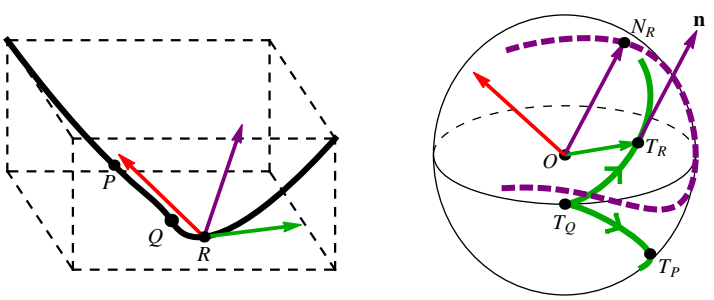



**Figure 6** A space curve (left) and its clustered Frenet frame (right) with the tantrix (solid) and normal indicatrix (dashed).

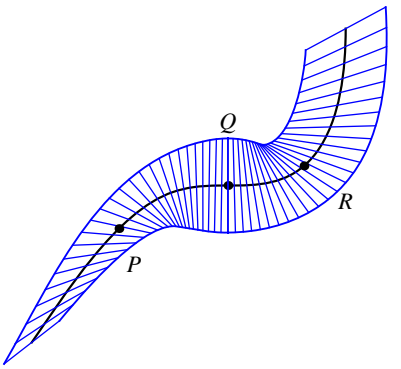



**Figure 7** Constructing a normal ribbon to visualize the $\kappa$-inflections of a space curve.

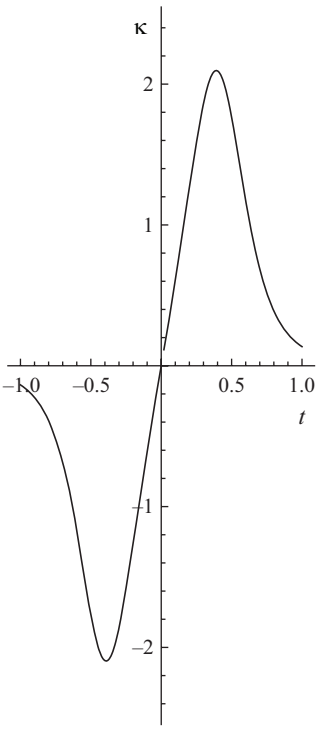



**Figure 8** The curvature function of the curve $\gamma(t) = (t, t^3, t^4)$.

It is difficult to identify $Q$ as a $\kappa$-inflection of the curve in FIGURE 6, even when we use Mathematica to vary our point of view. But if we construct a *normal ribbon* $\mathcal{R}$, or a surface consisting of short segments through the points of $\Gamma$ parallel to the principal normals, the inflection point becomes apparent. The vector $\mathbf{b}$ is normal to $\mathcal{R}$ and in FIGURE 7 our viewpoint is from the $\mathbf{b}$-side of $\mathcal{R}$. The curvature is negative at $P$ ($\Gamma$ bends right) and positive at $R$ ($\Gamma$ bends left). The bending of the normal segments on the ribbon suggests that the absolute curvature is greatest near $P$ and $R$. This is confirmed by the graph of the curvature function in FIGURE 8.

## Torsion

Just as the absolute curvature measures the rate at which the unit tangent $\mathbf{t}$ changes direction relative to arc length, the torsion measures the rate at which the unit binormal $\mathbf{b}$ changes direction. Since $\mathbf{b}$ is normal to $\mathbf{t}$ and $\mathbf{t}'$, it is normal to the plane spanned by these vectors at points where $\mathbf{t}' \neq \mathbf{0}$. Thus, the torsion measures the rate at which this plane changes direction, or the curve's attempt to twist out of this plane. It is zero for plane curves, where $\mathbf{b}$ is constant.

We just saw that the motion of the tip of the clustered tangent vector along the tantrix gives us a way to visualize the curvature. Similarly, the motion of the tip of the clustered binormal along the binormal indicatrix gives us information about the torsion. In fact, the twisting and turning of the Frenet frame along the curve tells us about both the curvature and torsion. We measure this motion by computing the derivatives $\mathbf{t}'$, $\mathbf{b}'$, and $\mathbf{n}'$.

To allow for the possibility of using different frames later on, we temporarily change notation and let $\{\mathbf{f}_1(s), \mathbf{f}_2(s), \mathbf{f}_3(s)\}$ be an orthonormal frame defined along the space curve $\gamma$. We wish to express each of the derivatives $\mathbf{f}_1'$, $\mathbf{f}_2'$, and $\mathbf{f}_3'$ as linear combinations of the vectors $\mathbf{f}_1$, $\mathbf{f}_2$, and $\mathbf{f}_3$. Let $\mathbf{f} = (\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)^T$ and $\mathbf{f}' = M\mathbf{f}$, for some $3 \times 3$ matrix $M$. We call $M = (m_{ij})$ the *frame matrix*. Its entries are functions of arc length along $\gamma$. To determine these functions, first note that since $\mathbf{f}_i \cdot \mathbf{f}_j = \delta_{ij}$, $\mathbf{f}_i' \cdot \mathbf{f}_j = -\mathbf{f}_j' \cdot \mathbf{f}_i$. Since the vectors $\mathbf{f}_i$ are orthonormal,

$$m_{ij} = \mathbf{f}_i' \cdot \mathbf{f}_j = -\mathbf{f}_j' \cdot \mathbf{f}_i = -m_{ji}. \tag{2}$$

Thus, $M$ is skew symmetric. In particular, its diagonal entries are zero.

For the Frenet frame, since $\mathbf{t}' \cdot \mathbf{b} = 0$, (2) implies that $\mathbf{b}' \cdot \mathbf{t} = 0$, and hence $\mathbf{b}' = (\mathbf{b}' \cdot \mathbf{n})\mathbf{n} = -\tau\mathbf{n}$, where

$$\tau = -\mathbf{b}' \cdot \mathbf{n} \tag{3}$$

is the *torsion*. We now have the *Frenet–Serret* equations that express the derivative of the Frenet frame in terms of the frame itself:

$$\begin{bmatrix} \mathbf{n}' \\ \mathbf{b}' \\ \mathbf{t}' \end{bmatrix} = \begin{bmatrix} 0 & \tau & -\kappa \\ -\tau & 0 & 0 \\ \kappa & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{n} \\ \mathbf{b} \\ \mathbf{t} \end{bmatrix}. \tag{4}$$

The negative sign in the definition (3) of torsion is arbitrary, but with this sign convention, a right-handed cyclindrical helix, like the one in FIGURE 3, has positive torsion. We use the binormal indicatrix $\Gamma_b$ to *see* why. As $P$ moves along $\Gamma$ with unit speed, the tip $B_P$ of the clustered binormal vector moves along $\Gamma_b$. If we define the forward direction along $\Gamma_b$ by $-\mathbf{n}$, the (signed) speed of $B_P$ gives the torsion. FIGURE 4 illustrates that the right-handed helix of FIGURE 3 has positive torsion since $B_P$ moves in the forward direction *opposite* the vector $\mathbf{n}$.

The cusps of the binormal indicatrix correspond to the *torsion inflections* ($\tau$-*inflections*), or the points where the torsion changes sign. At these points, $B_P$ comes to a stop as it reverses direction along $\Gamma_b$. Look ahead at FIGURE 12. It shows the binormal indicatrix of the curve of FIGURE 6. This curve has no $\tau$-inflections.

We now come to a rather curious point; the sign of the curvature is arbitrary, while that of the torsion is not. For curvature, this is intuitive. Reversing either our viewpoint or the forward direction interchanges left and right and therefore changes the sign of the curvature. For a proof, note that choosing the opposite viewpoint reverses the direction of $\mathbf{n} = \mathbf{b} \times \mathbf{t}$ and hence changes the sign of the curvature $\kappa = \mathbf{t}' \cdot \mathbf{n}$. Reversing the direction of travel also changes the sign of the curvature, for doing so reverses the direction of $\mathbf{n}$ but does *not* change the direction of $\mathbf{t}'$; the vector $\mathbf{t}'$ still points toward the inward side of $\Gamma$.

It would appear from (3) that the sign of the torsion also depends upon the orientation of the curve and our viewpoint. But this is not the case. For if we reverse the direction of $\mathbf{b}$, then $\mathbf{n} = \mathbf{b} \times \mathbf{t}$ also reverses direction, and the torsion does not change sign. Similarly, if we reverse the orientation, then $\mathbf{n} = \mathbf{b} \times \mathbf{t}$ and $\mathbf{b}'$ both reverse direction. This suggests that the sign of the torsion is a geometric property of $\Gamma$.

We use the normal ribbon to visualize this property. FIGURE 9 shows the space curve $\gamma(t) = (\cos t, \sin t, \cos 2t)$, $0 \le t \le 2\pi$, and its normal ribbon. The blue sections of the ribbon correspond to the points of $\Gamma$ with positive torsion. Point the thumb of your right hand along the curve in either direction and your fingers curl around $\Gamma$ in the same direction as do the normal segments. The red sections correspond to the points of negative torsion. The absolute torsion is greatest where the ribbon twists most

rapidly near the midpoints of each section. At the $\tau$-inflections $A$, $B$, $C$, and $D$, the ribbon flattens out as the torsion changes sign.

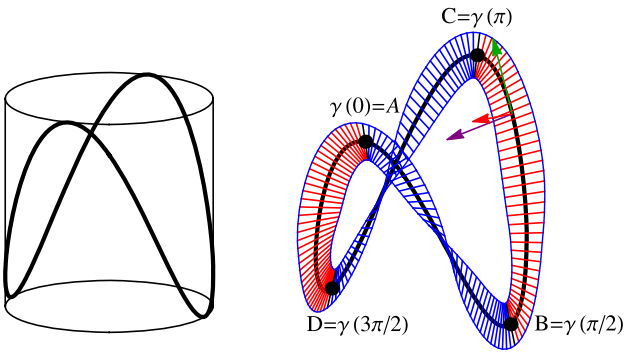



**Figure 9** Using the normal ribbon to visualize the $\tau$-inflections of a space curve.

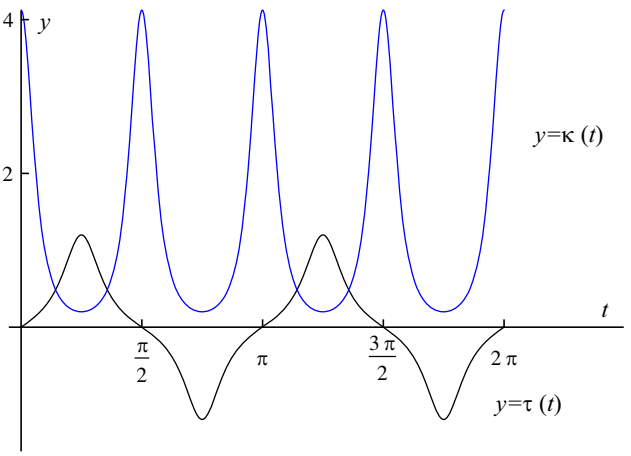



**Figure 10** The torsion and curvature functions for $\gamma(t) = (\cos t, \sin t, \cos 2t)$.

To see why the sense of the twisting of the normal ribbon gives the sign of the torsion, let's return to the Frenet equations (4) and focus on $\mathbf{n}' = \tau\mathbf{b} - \kappa\mathbf{t}$. By looking at how the ribbon twists around $\Gamma$, we pick up the torsion, or the component of $\mathbf{n}'$ normal to $\Gamma$ and parallel to $\mathbf{b}$. That a rightward bending corresponds to a *positive* torsion is explained by the fact that $\mathbf{n}' \cdot \mathbf{b} = \tau$ (and not $-\tau$). By looking at how rapidly the segments change direction on the surface of the ribbon, we see the component of $\mathbf{n}'$ parallel to the surface, or the absolute curvature. The rapid bending of the segments along the ribbon in FIGURE 9 near the $\tau$-inflections suggests that these points have the greatest absolute curvature. FIGURE 10 shows the curvature and torsion functions using the Frenet frame in FIGURE 9.

If the torsion is small, it may be difficult to identify the sense of the twisting of the normal ribbon. A plane image of a space curve may lead to the wrong conclusion about the sign of the torsion. To solve these problems, we turn to bicycle wheels. Our idea is simple. Mark one spoke of the wheel and move the wheel's center along $\Gamma$ while holding its axle parallel to the curve. As the wheel turns about its axle (in a manner we

explain later), the marked spoke traces out another ribbon along $\Gamma$, a *torsion ribbon*. This ribbon can be formed from a strip of paper and does not twist. The twisting of the normal ribbon about the torsion ribbon measures the torsion of $\Gamma$.

To justify these assertions, we need to understand how a bicycle wheel rotates as we carry its axle through space. Two ideas are crucial—angular velocity and the dual of a spherical curve. We start with the latter.

## The dual of a spherical curve

Let $\sigma = \sigma(s)$ be an oriented, piecewise-differentiable curve on $S^2$, with finitely many vertices, or points, where $\sigma'(s) = \mathbf{0}$. It might help to think of $\sigma(s)$ as the tantrix of the curve $\gamma(s)$, but for now, $s$ is just an arbitrary parameter. We assume the vertices of the trace $\Sigma$ are cusps so that $\Sigma$ has a well-defined tangent at each point. Define an orientation by choosing a vector field $\mathbf{T}$, one of the two possible smooth unit vector fields tangent to $\Sigma$. The dual $\sigma^*$ of $\sigma$ is the curve

$$\sigma^*(s) = \sigma(s) \times \mathbf{T}(s). \tag{5}$$

To complete the definition of $\sigma^*$, we define its forward direction to be $-\mathbf{T}$. This makes sense since

$$(\sigma^*)' = \sigma' \times \mathbf{T} + \sigma \times \mathbf{T}' = \sigma \times \mathbf{T}', \tag{6}$$

and $\mathbf{T}$, being normal to both $\sigma$ and $\mathbf{T}'$, is parallel or anti-parallel to $(\sigma^*)'$. It might seem odd to use $-\mathbf{T}$ instead of $\mathbf{T}$, but in this way, the dual of $\sigma^*$ is $\sigma$.

The dual of the *counterclockwise* circle $\Sigma$ of latitude $\phi$ north, with the forward direction counterclockwise as seen from the north pole, is the counterclockwise circle $\Sigma^*$ of latitude $\pi/2 - \phi$ north. As illustrated in FIGURE 11, the vectors $\mathbf{T}$ and $-\mathbf{T}$ define the same orientations on $\Sigma$ and $\Sigma^*$, respectively, since $\sigma(s)$ and $\sigma^*(s)$ lie on *opposite* sides of the $NS$-axis of $S^2$ through the poles. As $P = \sigma(s)$ moves along $\Sigma$ in the forward direction $\mathbf{T}$, $P^* = \sigma^*(s)$ moves along $\Sigma^*$ in the forward direction $-\mathbf{T}$. If we reverse the forward direction of $\Sigma$, then $\sigma^*$ reverses direction and $\Sigma^*$ is the circle of latitude $\pi/2 - \phi$ south. But the orientation of $\Sigma^*$ is still counterclockwise since $\sigma(s)$ and $\sigma^*(s)$ now lie on the *same* side of the $NS$-axis and the vectors $\mathbf{T}$ and $-\mathbf{T}$ define opposite orientations on $\Sigma$ and $\Sigma^*$. As $P$ moves clockwise along $\Sigma$ in the forward direction, $P^*$ moves clockwise along $\Sigma^*$ against the forward direction.

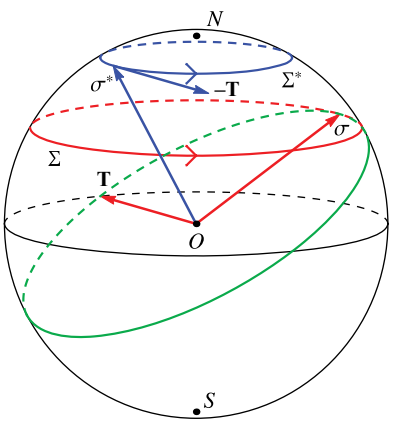



**Figure 11** The dual of a circle of latitude.

More generally, the tantrix (with forward direction $\mathbf{n}$) and the binormal indicatrix (with forward direction $-\mathbf{n}$) of a space curve are dual. The proof is immediate since $\mathbf{t} \times \mathbf{n} = \mathbf{b}$. FIGURE 12 (right) shows a great circle rolling on the tantrix of the curve of FIGURE 6 as its left pole traces the binormal indicatrix. The left half of the figure shows a view of the binormal indicatrix from the north pole and suggests that the cusps of a spherical curve correspond to the geodesic inflections of its dual. These are the points that look like inflections to a resident of the sphere.

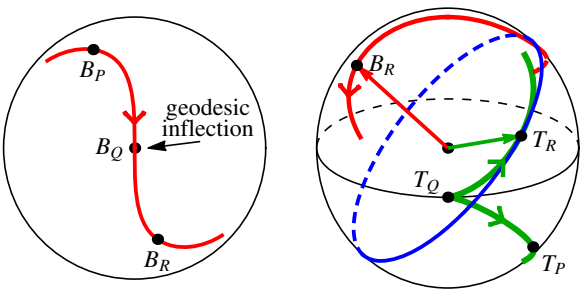



**Figure 12** The binormal indicatrix (left) and the binormal indicatrix as the dual of the tantrix (right).

## Angular velocity

Hold the axle of a bicycle wheel fixed (relative to the earth) and set the wheel spinning. The angular velocity $\omega$ of the wheel is a vector parallel to the axle that points in the direction of your thumb as you curl the fingers of your right hand around the wheel with the sense of rotation. The magnitude of $\omega$ is the rate of rotation. The velocity of a point $P$ moving with the wheel and remaining fixed relative to it is

$$\mathbf{v}_P = \omega \times \overrightarrow{OP}, \tag{7}$$

where $O$ is any point on the axis of rotation.

Now hold a baseball and turn it randomly while keeping its center $O$ fixed. Even if you try to move the ball smoothly, you probably make small rotations about fixed axes as you vary the axes. In fact, a theorem due to Euler states that at any moment a rigid body with a fixed point $O$ is rotating about an axis through $O$, the *instantaneous axis of rotation*. At that moment, the points of the axis are at rest and the velocity of any point $P$ fixed relative to the object is given by (7), where the angular velocity is defined as if the axis was fixed. In particular, the angular velocity is parallel to the instantaneous axis of rotation.

A body moving through space without a point fixed relative to the ambient space, like the Frenet frame moving along the helix in FIGURE 3, does not have an instantaneous axis of rotation. But we can determine its angular velocity by considering its motion relative to a point $O$ fixed on the body. Then $O$ is a fixed point of the relative motion, and we define the angular velocity as before. The velocity of a point $P$ of the object *relative* to $O$ is given by (7).

We use the Frenet equations to determine the angular velocity $\omega_n$ of the Frenet frame as a curve is traversed with unit speed. Since $\mathbf{t}' = \omega_n \times \mathbf{t} = \kappa\mathbf{n}$, $\mathbf{n}$ is orthogonal to $\omega_n$. Since also $\mathbf{n}' = \omega_n \times \mathbf{n}$, $\{\omega_n, \mathbf{n}, \mathbf{n}'\}$ is a positively oriented triple of mutually orthogonal vectors. Finally, since $\mathbf{n}$ is a unit vector,

$$\omega_n = \mathbf{n} \times \mathbf{n}' = \mathbf{n} \times (-\kappa\mathbf{t} + \tau\mathbf{b}) = \tau\mathbf{t} + \kappa\mathbf{b}. \tag{8}$$

The vector $\omega_n$ is sometimes called the Darboux vector, after the French mathematician Jean Gaston Darboux (1842–1917). For example, the angular velocity of the Frenet frame of the helix in FIGURE 3 is constant. For the Darboux vector $\omega_n = \mathbf{n} \times \mathbf{n}'$ points from the center of $S^2$ to the north pole and has constant magnitude $|\omega| = \sqrt{\kappa^2 + \tau^2}$. As the helix is traversed with unit speed, its clustered Frenet frame (FIGURE 4) rotates about the axis through the poles at a constant rate. But for a general curve, the Darboux vector varies with $s$ as both the direction and magnitude of $\omega_n$ change.

## Rotating wheels and frames

The frame matrix for the Frenet frame (4) is sparse; there are zeros *off* the diagonal ($m_{32} = m_{23} = 0$ or $m_{31} = m_{13} = 0$). While a general frame matrix is not sparse, we can use a bicycle wheel to generate all orthonormal frames with similarly sparse frame matrices. Hold the axle of the wheel and move its center $O$ along $\Gamma$ from $\gamma(0)$ to $\gamma(\ell)$. Label one end of the axle $A$ and the end of a selected spoke $S$, with $\left|\overrightarrow{OA}\right| = \left|\overrightarrow{OS}\right| = 1$. As $P$ moves along $\Gamma$, the orthogonal unit vectors $\mathbf{f}_1 = \overrightarrow{OA}$, $\mathbf{f}_2 = \overrightarrow{OS}$, and $\mathbf{f}_3 = \mathbf{f}_1 \times \mathbf{f}_2$ define a positively oriented orthogonal frame (*POOF*) along $\gamma$. We call any POOF along $\gamma$ that can be generated in this way a *bicycle frame*.

We make two assumptions about our wheel to ensure that the corresponding frame matrix is sparse:

(a) The wheel is free to rotate about its axle without friction, and

(b) the wheel starts at rest relative to its axle.

The first condition ensures that there is no torque about the axle and that the component of the wheel's angular velocity about the axle is constant. By (b) this constant is zero. The marked spoke has a tendency to point in a fixed direction and changes direction only as much as required to stay in the plane of the wheel.

This idea is borne out in the sparseness of the corresponding frame matrix. Since the angular velocity $\omega_{f_1}$ of the bicycle frame is parallel to the plane of the wheel, $\mathbf{f}_2' = \omega_{f_1} \times \mathbf{f}_2 = -k_2\mathbf{f}_1$ and $\mathbf{f}_3' = \omega_{f_1} \times \mathbf{f}_3 = -k_3\mathbf{f}_1$, for some functions $k_2$ and $k_3$. This gives the bicycle frame equations

$$\begin{bmatrix} \mathbf{f}_1' \\ \mathbf{f}_2' \\ \mathbf{f}_3' \end{bmatrix} = \begin{bmatrix} 0 & k_2 & k_3 \\ -k_2 & 0 & 0 \\ -k_3 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \mathbf{f}_3 \end{bmatrix}. \tag{9}$$

We leave it as an exercise to show conversely that any *POOF* with a sparse frame matrix is a bicycle frame. For example, the Frenet equations (4) imply that the Frenet frame is the bicycle frame generated by holding the axle parallel to $\mathbf{n}$ and choosing the marked spoke to initially point in the direction $\mathbf{f}_2(0) = \mathbf{b}(0)$.

Finally, since $\{\omega_{f_1}, \mathbf{f}_1, \mathbf{f}_1'\}$ is a positively oriented orthogonal frame with $\mathbf{f}_1' = \omega_{f_1} \times \mathbf{f}_1$ and $|\mathbf{f}_1| = 1$, the angular velocity $\omega_{f_1}$ of the wheel and any of its associated frames is

$$\omega_{f_1} = \mathbf{f}_1 \times \mathbf{f}_1' = -k_3\mathbf{f}_2 + k_2\mathbf{f}_3. \tag{10}$$

## Rolling wheels

To see the connection between bicycle wheels and the dual of a spherical curve, move the marked end $A$ of a wheel's axle along the oriented spherical curve $\sigma(s)$, keeping the center of the wheel fixed at the center $O$ of $S^2$. We assume as before that $\sigma(s)$ is piecewise differentiable with finitely many cusps and that the unit tangent vector field $\mathbf{T}$ gives the orientation. We claim the wheel rolls without slipping on the dual $\Sigma^*$ as we move $A$ along $\Sigma$ (see [**3**, pp. 537-538] for a more general result). FIGURE 13 shows a simple example where $\Sigma$ is a counterclockwise circle of latitude. Our proof follows Levi [**4**].

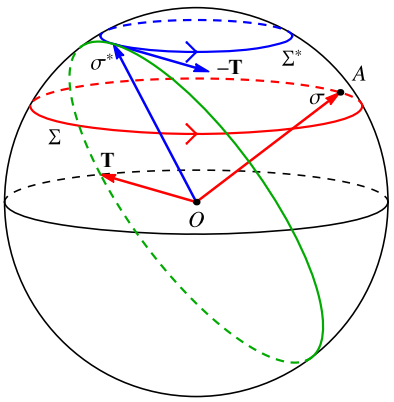



**Figure 13** A bicycle wheel rolling on the dual $\Sigma^*$ as its axle moves along $\Sigma$.

To get a feel for why we need to assume the wheel is frictionless, suppose instead that the wheel in FIGURE 13 was welded to its axle. Then the wheel would slide rather than roll along $\Sigma^*$. If there was friction at the axle, then the wheel would slide and roll, much like a bowling ball does before it makes the transition to pure rolling.

To show that our wheel rolls on $\Sigma^*$, it suffices to show that the wheel and $\Sigma^*$ are tangent at $\sigma^*$ and that the point of the wheel in contact with $\Sigma^*$ is at rest. To show the former, note that since $(\sigma^*)^* = \sigma$, the great circle tangent to $\Sigma^*$ at $\sigma^*$ has $\sigma$ as one of its poles (compare FIGURES 13 and 11).

To show the latter, we take the parameter $s$ to be time and calculate the angular velocity $\omega_\sigma$ of the wheel. From (10), it follows that

$$\omega_\sigma = \sigma \times \sigma'.$$

If $\sigma' = \mathbf{0}$, then the entire plane of the wheel is at rest, but otherwise the instantaneous axis of rotation is the line through $O$ parallel to $\omega_\sigma$. But since $\sigma'$ is parallel or anti-parallel to $\mathbf{T}$, this axis passes through the corresponding point of the dual $\sigma^* = \sigma \times \mathbf{T}$. Thus, the wheel rolls on $\Sigma^*$.

## The curvature wheel

We return to the question posed in the introduction and show how to use a bicycle wheel to measure the total curvature of a space curve $\gamma(s)$. With the wheel's center at $\gamma(0)$, hold its axle parallel to $\mathbf{b}(0)$ and mark a spoke. Any spoke will do, but we prefer to mark the spoke parallel to $\mathbf{t}(0)$. Now move the center of the wheel along $\Gamma$, keeping

its axle parallel to the binormal $\mathbf{b}(s)$. We call this wheel the *curvature wheel* since the directed angle $\theta = \theta(s)$ from the marked spoke to the forward direction $\mathbf{t}(s)$ measures the total curvature from $\gamma(0)$ to $\gamma(s)$.

For a plane curve, this is obvious. The marked spoke points in the fixed direction $\mathbf{t}(0)$ and $\theta(s)$ measures the angle from $\mathbf{t}(0)$ to $\mathbf{t}(s)$. The nonzero torsion of a space-curve forces the wheel to rotate and the marked spoke to change direction, but $\theta(s)$ still measures the total curvature. To see why, we define the *curvature frame* $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ with $\mathbf{f}_1(s) = \mathbf{b}(s)$ and $\mathbf{f}_2(0) = \mathbf{t}(0)$. We give the continuous function $\theta$ the correct sign by requiring that $\theta(0) = 0$ and

$$\mathbf{t} = \mathbf{f}_2 \cos\theta + \mathbf{f}_3 \sin\theta. \tag{11}$$

To show $\theta'(s)$ measures the curvature, we use (10) to compute the angular velocity of the curvature frame relative to the fixed three-dimensional space in which it rotates as

$$\omega_b = \mathbf{b} \times \mathbf{b}' = \mathbf{b} \times (-\tau\mathbf{n}) = \tau\mathbf{t}. \tag{12}$$

By (8) the angular velocity of the Frenet frame relative to the curvature frame is

$$\omega_{n/b} = \omega_n - \omega_b = (\tau\mathbf{t} + \kappa\mathbf{b}) - \tau\mathbf{t} = \kappa\mathbf{b}.$$

Thus, the rate at which the Frenet frame rotates relative to the curvature frame measures the curvature. The net rotation of the Frenet frame gives the total curvature.

FIGURE 14 shows a right cylindrical helix whose tangents make a $60^\circ$ angle with its axis, along with seven equally spaced curvature wheels (we see the sixth wheel from the $-\mathbf{b}$-side of the wheel). The figure suggests that the unit tangent rotates counterclockwise through a bit less than one revolution relative to the marked spoke in one turn of the helix so that the total curvature is a bit less than $2\pi$.

Imagine a second copy of the wheel, the *clustered* curvature wheel. It stays parallel to the moving wheel but its center is fixed. As the marked end of its axle traces out the binormal indicatrix on $S^2$, the wheel rolls on the tantrix (the dual of the binormal indicatrix) and its net roll records the total curvature. FIGURE 15 shows the wheel rolling on the tantrix as it records the curvature of the helix of FIGURE 14. Since the wheel rolls in the forward direction $\mathbf{n}$, the total curvature is the circumference of the tantrix, or $2\pi \sin 60^\circ \sim 1.7\pi$.

A nearby park has a slide in the form of a cylindrical helix that makes about a $60^\circ$ angle with the vertical. To determine its total curvature, we could measure the angle of rotation of a wheel as we hold its axle parallel to $\mathbf{b}$ and go down the slide. But we made the same measurement at home by holding the axle at a $30^\circ$ angle to the vertical (the angle of the binormal) and turning in a circle. A marked spoke did indeed rotate through a little less than one revolution.

We can also imagine the plane of the clustered wheel rolling on the tangent cone. As it rolls the wheel acts like a protractor, adding the differential angles between successive tangents. The angle between the marked spoke and $t(s)$ measures the corresponding angle on the surface of the cone between $\mathbf{t}(0)$ and $\mathbf{t}(s)$, and hence the total curvature. See FIGURE 16 and compare with FIGURE 5.

FIGURE 12 (right) shows an example where the tantrix has cusps. As the axle moves along the binormal indicatrix in the forward direction the clustered curvature wheel rolls along the tantrix in the negative direction, then stops and reverses direction at the cusp as the curvature changes sign.

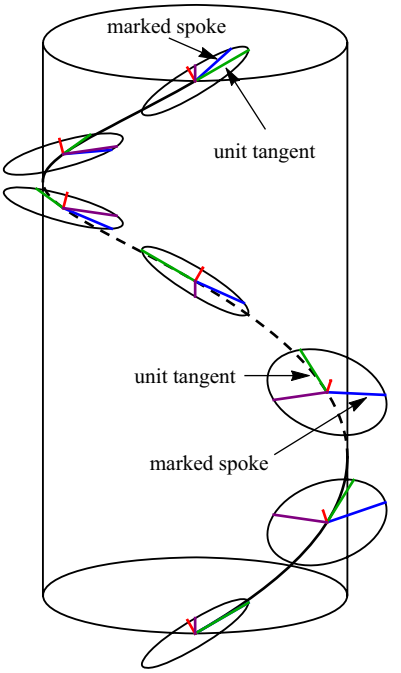



**Figure 14** The curvature wheel moving along a helix.

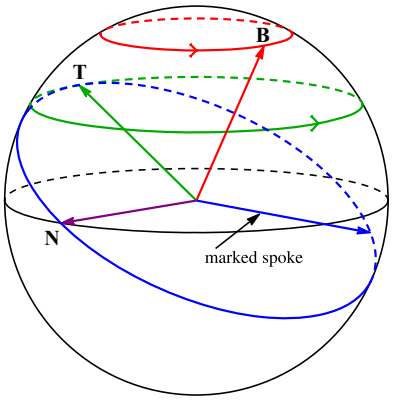



**Figure 15** The clustered curvature wheel of a helix rolling on the tantrix.

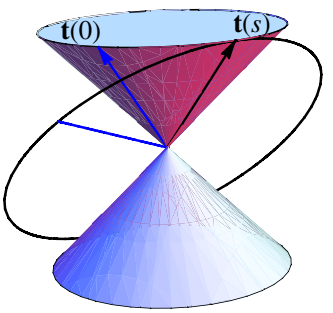



**Figure 16** The curvature wheel rolling on the tangent cone.

## The torsion wheel

Suppose instead we point the axle of the wheel in the direction of $\mathbf{t}$ as we carry its center along $\Gamma$. We call this wheel the *torsion wheel* since the rate of rotation of the Frenet frame relative to the wheel measures the torsion of $\gamma$. As the axle of the clustered torsion wheel moves along the tantrix, the wheel rolls on the binormal indicatrix and the net roll records the total torsion.

To verify this we calculate the angular velocity of the torsion wheel as

$$\omega_t = \mathbf{t} \times \mathbf{t}' = \mathbf{t} \times \kappa\mathbf{n} = \kappa\mathbf{b}. \tag{13}$$

Thus,

$$\omega_{n/t} = \omega_n - \omega_t = (\tau\mathbf{t} + \kappa\mathbf{b}) - \kappa\mathbf{b} = \tau\mathbf{t},$$

as expected.

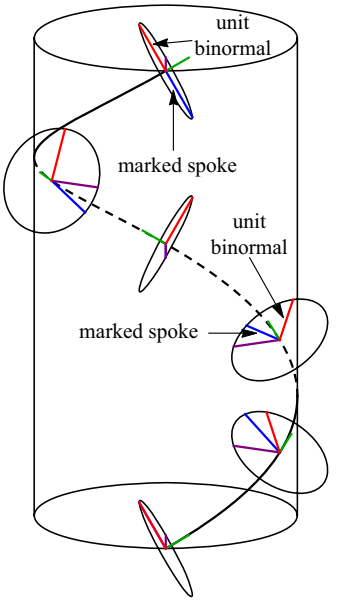



**Figure 17** The torsion wheel moving along a helix.

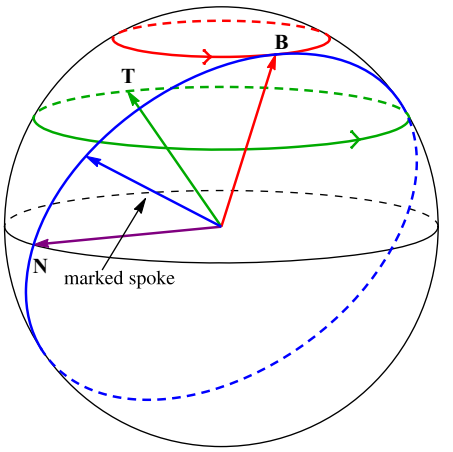



**Figure 18** The clustered torsion wheel of a helix rolling on the binormal indicatrix.

FIGURE 17 shows the torsion wheel for the right cylindrical helix of FIGURE 14. We chose the marked spoke to initially coincide with $\mathbf{b}(0)$ so that the directed angle from the marked spoke to $\mathbf{b}(s)$ records the total torsion. If we view the wheel from the *opposite* side of $\mathbf{t}$ as in the figure, the vector $\mathbf{b}$ rotates *clockwise* relative to the marked spoke when the torsion is positive.

FIGURE 18 shows the clustered wheel rolling on the binormal indicatrix for the helix of FIGURE 17. The total torsion of one turn of the helix is the length of the binormal indicatrix, in this case the circumference of the circle of latitude $\phi = \pi/3$, or $2\pi\cos(\pi/3) = \pi$. The torsion is positive since the clustered wheel rolls along the binormal indicatrix in the forward direction $-\mathbf{n}$.

For a more dramatic representation of the total torsion, we construct a *torsion ribbon*, the ruled surface left behind by a marked spoke of the torsion wheel as its center moves along $\Gamma$. If we choose $\mathbf{f}_2(0) = \mathbf{n}(0)$, the directed angle from $\mathbf{f}_2(s)$ to $\mathbf{n}(s)$ measures the total torsion, now visible as the twisting of the normal ribbon about the torsion ribbon. FIGURE 19 shows these ribbons for our helix. The torsion is positive by the right-hand rule. FIGURE 20 shows the torsion ribbon alone. The tangent planes to the torsion ribbon are constant along each of the ribbon's segments (exercise). Thus, the ribbon is a *developable surface* [**6**, pp. 66–72] and can be formed from a strip of paper.

We should point out that any frame that rotates with the torsion wheel is a Bishop frame [**2**]. In the short time since their introduction in 1975, Bishop frames have become quite popular, especially among computer scientists. See [**2**] for the paper that introduced these frames.

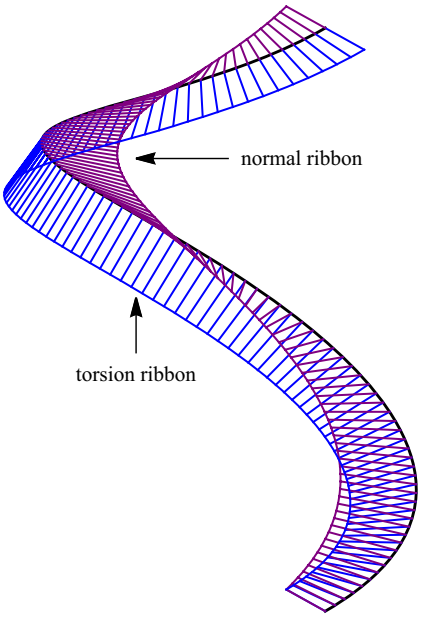



**Figure 19** The normal ribbon twisting about the torsion ribbon.

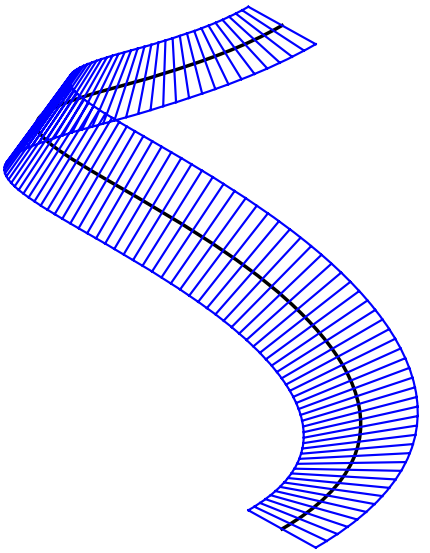


**Figure 20** A torsion ribbon for a helix.

**Acknowledgment** I would like to thank Art West for his many insightful suggestions that did much to improve an earlier version of this paper. I would also like to thank the referee for many helpful suggestions and Mike Gaul for taking the time to talk about many of the ideas in this paper. Most of all, thanks to my wife Mei chun Chang for her constant support.

## REFERENCES

1. T. A. Apostolatos, Hodograph: A useful geometric tool for solving some difficult problems in dynamics, *Am. J. Phys.* **71** (2003) 261–266, http://dx.doi.org/10.1119/1.1527948.
2. R. L. Bishop, There is more than one way to frame a curve, *Amer. Math. Monthly* **82** (1975) 246–251, http://dx.doi.org/10.2307/2319846.
3. R. C. Hibbeler, *Engineering Mechanics: Dynamics*. Eleventh edition. Pearson Prentice Hall, Upper Saddle River, NJ, 2007.
4. M. Levi, A "bicycle wheel" proof of the Gauss–Bonnet theorem, *Expo. Math.* **12** (1994) 145–164.
5. J. Stewart, *Calculus: Early Transcendentals*. Seventh edition. Brooks/Cole, Belmont, CA, 2012.
6. D. J. Struik, *Lectures on Classical Differential Geometry*. Second edition. Addison-Wesley Publishing, Reading, MA, 1961.

**Summary.** We describe classical ways to think about the signed curvature and torsion of space curves using spherical indicatrices to record the twisting and turning of the Frenet frame. By constructing ribbons around curves, we show how to visualize their curvature and torsion. Finally, we use bicycle wheels to measure these quantities.

**FRED KUCZMARSKI** (MR Author ID: 944646) received his Ph.D. from the University of Washington in 1995 under the guidance of Paul Goerss. His passion for geometry was later sparked by James King. He currently teaches at Shoreline Community College. In addition to thinking about mathematics, he enjoys baking bread and hiking in Washington's North Cascades.

# Solving the KO Labyrinth

ALISSA S. CRANS
ROBERT J. ROVETTI
JESSICA VEGA
Loyola Marymount University
Los Angeles, CA 90045
acrans@lmu.edu
rrovetti@lmu.edu
jvega32@gmail.com

The *KO Labyrinth*, pictured in FIGURE 1, is a colorful spherical puzzle with 26 chambers, some of which can be connected via internal holes through which a small ball can pass when the chambers are aligned correctly. The puzzle can be realigned by performing physical rotations of the sphere in the same way one manipulates a Rubik's Cube, which alters the configuration of the puzzle. There are two special chambers: one where the ball is put into the puzzle and one where it can exit. The goal is to navigate the ball from the entrance to the exit.

**Figure 1** The *KO Labyrinth.*

We will explore questions related to solving the puzzle, both as originally intended and under modified rules. We first consider a "goal-directed" player who is motivated to reach the end of the maze as quickly as possible and show that the shortest path through the maze takes only 10 moves. Next, we turn to a "random" player who wanders aimlessly through the puzzle. Using two different techniques, we show that such a player makes, on average, about 340 moves before reaching the end of the maze. We

*Math. Mag.* **88** (2015) 27–36. doi:10.4169/math.mag.88.1.27. 
MSC: Primary 05C12, Secondary 05C40; 05C81.

also determine the most- and least-visited chambers. Then, we pose an analogue of the gambler's ruin problem and separately consider whether one can solve the puzzle if we consider certain chambers to be off-limits. We conclude with comments and questions for future investigation.

## The KO graph

The 26 chambers of the puzzle have printed labels consisting of a letter and a number: *A1* through *A8*, *B1* through *B12*, and *C1* through *C6*. The chambers can be referenced in terms of corners, edges, and faces (as is often done with the Rubik's Cube). The *A chambers* are the eight "corners" of the sphere; they can be connected via internal holes only to the *B chambers* that are the 12 center "edges." The *C chambers* are the six "faces" that can be connected via internal holes only to the *B* chambers. There is no central chamber. The ball, via external holes, enters the maze through *C1* and exits through *C6*.

The *KO Labyrinth* can be manipulated similarly to the Rubik's Cube, which allows for a multitude of puzzle configurations; however, our interest here is not in how the puzzle itself is aligned but rather in how the chambers can be connected. Therefore, we define a *move* to be the act of directly passing the ball through a hole from one chamber to another, where such an act may require a number of Rubik's Cube-like rotations of the actual sphere before it can be executed. Insertion and extraction of the ball through the exterior holes are not considered moves. Not all pairs of chambers can be connected to allow a direct move even when the chambers can be placed next to each other because there is not always a hole. By inspecting the puzzle to determine all possible moves, we generated what we call the *KO graph*, shown in FIGURE 2. The 26 vertices represent the chambers of the puzzle, with two vertices joined by an edge if it is possible to move the ball between those two chambers.

A solution to the puzzle is thus a path from vertex *C1* to vertex *C6*. More formally, a *path* in a graph is a sequence of distinct vertices where each pair of consecutive vertices is joined by an edge. Recall that the *length* of a path is the number of edges involved and the *distance* between two vertices is the length of a shortest path between them. In addition, a graph is *connected* if every pair of vertices is connected by a path, and *disconnected* otherwise.

This representation of the KO graph is highly structured, with a vertical axis composed mostly of *B* chambers. It is symmetric about this axis except for the edge connecting *B1* and *C3*. The KO graph is connected.

There are pairs of chambers in the maze that function identically. We call such pairs of chambers "twins." In the KO graph, vertex $x$ is a *twin* of vertex $y$ if and only if $x$ and $y$ are adjacent to the same vertices. There are seven sets of twin vertices: $\{A1, A2\}$, $\{A3, A4\}$, $\{A5, A6\}$, $\{A7, A8\}$, $\{B2, B3\}$, $\{B7, B8\}$, and $\{B10, B11\}$.

Now that we have an understanding of the puzzle and a convenient means of representing it mathematically, we will explore its features as a puzzle from two standpoints: goal-directed play and random play.

## The goal-directed player

We begin by considering an intelligent player who has some knowledge of graph theory and uses that to her advantage to navigate through the maze using the least number of moves. We can use a *breadth-first search* or *Dijkstra's algorithm* (see [**2**], for example) to find the shortest path. Either of these iterative processes can be used to compute the distance between the exit vertex *C6* and every other vertex in the KO

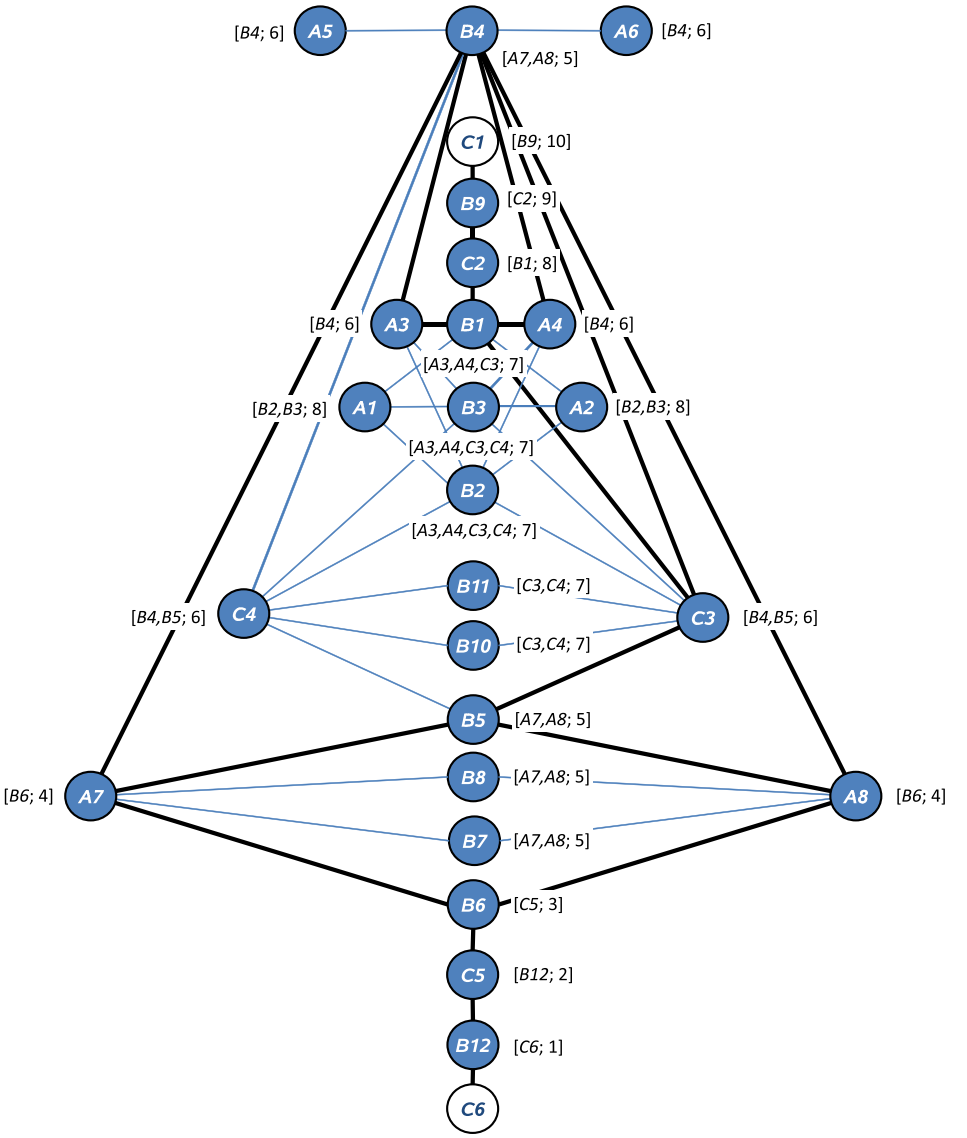



**Figure 2** KO graph. The labels are in the form [preceding vertices in the shortest path(s) from *C6*; distance to *C6*] and are produced by Dijkstra's algorithm. The bolded subgraph contains all shortest paths from the entrance chamber *C1* to the exit chamber *C6*.

graph. These distances are listed in Column I of TABLE 1; the distance from *C1* to *C6* is ten. An advantage to using Dijkstra's algorithm is that it also enables us, for each vertex, to keep track of the preceding vertex in the shortest path, or vertices if there are multiple shortest paths, from *C6*. For each vertex in the KO graph in FIGURE 2, the label gives the preceding vertices in the shortest path(s) from *C6* and distance to *C6*. For example, the label next to *B1* tells us that it is a distance of seven from *C6* and that there are actually at least three shortest paths from *B1* to *C6* that involve one of *A3*, *A4*, or *C3*.

To determine the total number of ten-move paths, we consider a subgraph of the original KO graph, appearing in bold in FIGURE 2, which is constructed by beginning at *C1* and systematically traversing the graph only through the vertices indicated by the

TABLE 1: Results by chamber

I: Distance to *C6* when starting at this chamber

II: Expected number of visits to this chamber when starting at *C1*

III: Expected number of moves before reaching *C6* when starting at this chamber

IV: Probability of reaching *C6* in the win-or-lose scenario, starting at this chamber

| Chamber | I | II | III | IV |
|---|---|---|---|---|
| *A*1 | 8 | 12.18 | 333.33 | 0.44 |
| *A*2 | 8 | 12.18 | 333.33 | 0.44 |
| *A*3 | 6 | 15.97 | 330.74 | 0.45 |
| *A*4 | 6 | 15.97 | 330.74 | 0.45 |
| *A*5 | 6 | 3.80 | 322.95 | 0.47 |
| *A*6 | 6 | 3.80 | 322.95 | 0.47 |
| *A*7 | 4 | 17.50 | 302.00 | 0.51 |
| *A*8 | 4 | 17.50 | 302.00 | 0.51 |
| *B*1 | 7 | 25.24 | 333.55 | 0.42 |
| *B*2 | 7 | 23.91 | 331.72 | 0.45 |
| *B*3 | 7 | 23.91 | 331.72 | 0.45 |
| *B*4 | 5 | 30.39 | 321.95 | 0.47 |
| *B*5 | 5 | 14.81 | 316.05 | 0.49 |
| *B*6 | 3 | 9.00 | 261.00 | 0.58 |
| *B*7 | 5 | 7.00 | 303.00 | 0.51 |
| *B*8 | 5 | 7.00 | 303.00 | 0.51 |
| *B*9 | 9 | 12.41 | 341.55 | 0.14 |
| *B*10 | 7 | 7.81 | 329.11 | 0.46 |
| *B*11 | 7 | 7.81 | 329.11 | 0.46 |
| *B*12 | 1 | 2.00 | 89.00 | 0.86 |
| *C*1 | 10 | 7.21 | 342.55 | – |
| *C*2 | 8 | 10.41 | 338.55 | 0.28 |
| *C*3 | 6 | 27.48 | 328.60 | 0.46 |
| *C*4 | 6 | 23.27 | 327.61 | 0.46 |
| *C*5 | 2 | 4.00 | 176.00 | 0.72 |

labels produced by Dijkstra's algorithm. We see that there are eight ten-move paths: two by way of *C3* and *B5* and six by way of *B4*.

Two pairs of twins, $\{A3, A4\}$ and $\{A7, A8\}$, appear in these ten-move paths. Twin vertices are redundant in that, in any path, any twin vertex may be replaced by its twin without altering anything else about the path. Removing one twin from $\{A7, A8\}$ eliminates half of our ten-move paths. Eliminating one twin from both pairs $\{A3, A4\}$

and {*A7*, *A8*} leaves three shortest paths. Even if we eliminate the entire pair {*A3*, *A4*}, we still have four ten-move solutions. However, removing the entire pair {*A7*, *A8*} causes far more harm as it disconnects the graph and renders the puzzle unsolvable. This observation inspires other questions of connectivity, such as identifying which other chambers are absolutely necessary in order to solve the puzzle, which we will consider later on.

## The random player

We now turn our attention to a simpler player who, not being very insightful, can consider only one move at a time and makes those moves randomly. Recall that we use the word *move* to refer to the KO ball directly passing from one chamber to another and not the number of realignments of the sphere. This simple player has knowledge of only two things: (1) which chamber currently holds the ball (i.e., the current *state*) and (2) what the possible next moves are (i.e., which states the ball can move to). We then wonder: Will the ball ever reach the exit chamber? How many moves, on average, will it take?

**A simulation** There are many ways in which moves can be chosen randomly. We assume our player moves the ball to any adjacent state with equal probability. Each move is made without regard to which moves have occurred previously. This allows for the possibility that the ball may double-back on itself or even travel in a cycle (although the probability that such cycles continue indefinitely is zero). The probabilistic route a player might take along the KO graph constitutes what is known as a *random walk* on the graph, a special case of a Markov chain.

We would like to know how many moves are needed, on average, to complete the maze in a random walk. More generally, we ask for the expected (average) number of moves it takes to get from chamber $i$ to chamber $j$ over many independent trials. In answering this, we will use the $26 \times 26$ *adjacency matrix* $A$ of our graph, in which the rows and columns correspond to the chambers in the order *A1–A8*, *B1–B12*, *C1–C6* and the $(i, j)$ entry is 1 if vertex $i$ is adjacent to vertex $j$, and 0 otherwise. Note that $A$ is symmetrical and every diagonal entry is 0.

We can now approximate the answer to our question via the following algorithm:

(1) Begin with $i = 21$, which corresponds to vertex *C1*. Set a counter variable $c = 0$.

(2) Randomly generate an integer $j$ from 1 to 26. This is the vertex that we will try to move to.

(3) If $a_{ij} = 0$, this means that vertex $i$ and vertex $j$ are not adjacent, so we cannot move and thus we return to Step 2. If $a_{ij} = 1$, we move from vertex $i$ to vertex $j$ and increment the counter $c$ by 1.

(4) If $j = 26$, we have reached the end of the maze (vertex *C6*); we report the value of $c$ and end the simulation. Otherwise, we replace the value of $i$ with the value of $j$ and return to Step 2.

When a simulation ends, the value of the counter $c$ gives the number of random moves that were required to reach *C6* starting in *C1*. After one million repeated trials, the average number of moves was 341.79 with a standard deviation of 318.84, and the longest simulation was 3,564 moves. In 162 simulations (0.0162% of the total), one of the ten-move paths was found, a somewhat rare event. In terms of solving the puzzle efficiently, it is clear that random guessing is unlikely to pay off, and clever maneuvers by a goal-directed player are better.

**The end is near** As the number of trials in our simulation increases, our computed average number of moves will approach its theoretical expected value. The theoretical value can be computed directly using some matrix algebra. We form the *transition matrix* $P$ from the adjacency matrix $A$ by dividing each entry $a_{ij}$ by the sum of the entries in the $i$th row; $p_{ij}$ is then the probability that a player currently in state $i$ will advance to state $j$ on the next move. Since our player is required to move the ball and not remain in her current chamber, each row of $P$ sums to 1 and each diagonal element $p_{ii}$ is 0.

The transition matrix $P$ represents a nonterminating random walk on a connected graph, which we show the KO graph to be when we consider questions of connectivity later on. As constructed, the KO graph allows a walk that reaches state *C6* to leave that state on the next move and continue on indefinitely. However, because reaching state *C6* for the first time amounts to solving the puzzle, we modify the KO graph so that the edge leading to state *C6* is *unidirectional*, disallowing a return to state *B12*. In the language of random walks, this transforms vertex *C6* into an *absorbing state* that traps the walk permanently. Any state that is not absorbing is now called a *transient state* and is visited only a finite number of times before the walk is eventually trapped in an absorbing state. The matrix $P$ is concurrently modified by setting $p_{26,j} = 0$ for all $j \neq 26$, and $p_{26,26} = 1$.

How long will it take to reach the absorbing state $j = 26$? This is an example of a *first passage time* problem in which we compute the expected number of moves it takes to "pass into" a state for the first time. We approach the problem by considering how many times we are expected to visit the various transient states before being absorbed. We construct the $25 \times 25$ submatrix $Q$ of $P$ by deleting the 26th row and column of $P$ so that $Q$ contains the transient states only. A well-known result [**1**] from the theory of Markov chains is that the $(i, j)$ entry of the matrix $M = (I - Q)^{-1}$ is the expected number of visits to transient state $j$, conditional on starting from state $i$, before being absorbed. If our player starts at the entrance state *C1*, then the corresponding 21st row of $M$ gives the expected number of visits to each of the remaining states before reaching the end of the maze. These values are given in Column II of TABLE 1.

We note some obvious patterns and symmetries: chamber *B12*, the penultimate chamber, is visited the least number of times, and twin chambers such as $\{A5, A6\}$ are visited the same number of times. The most-visited chamber, with 30.39 visits on average, is *B4*, which also happens to be the chamber with the highest degree. And, perhaps frustratingly, we expect to visit chamber *C1* (where we began) more than seven times before finally reaching the end!

The row sums of $M$ provide the expected total number of moves to all transient states before eventually finishing (conditional on starting in the state corresponding to each row), and the sum of the 21st row of $M$, 342.55, is the expected number of moves starting from *C1*. The mean number of moves computed in our earlier simulation (341.79) can be considered as a single statistical sample (of size $n$ =1,000,000) from a population whose true mean is 342.55. Interestingly, the 95% confidence interval around the sample mean is [341.17, 342.41], which does not capture the true mean, an outcome that should occur in 5% of sampled trials.

What if we start from a state other than the beginning chamber? Intuitively, we expect the total number of moves to be less than 342.55 since we are "closer to the finish line," and this bears out upon examination of all the row sums of $M$ (see Column III of TABLE 1). Note the relative constancy of these numbers; starting in almost any chamber, between 302 and 342 random moves are required to reach the end. The exceptions occur if we start in chambers *B6*, *B12*, or *C5*, which are the closest to the end of the maze. Yet we are perhaps mildly annoyed to see that even if we start in chamber *B12* (only one move away!) we are nevertheless expected to take 89 moves to finish. Such is the nature of random decision making.

## Changing the rules

Considering variations on a theme arises naturally in many applications and allows for a deeper mathematical exploration. For example, as a way to make the puzzle more challenging, we can consider whether one can still solve the game when certain chambers are off limits. First, however, we explore a novel use of the *KO Labyrinth* by using it to play a simple game modeled after a well-known problem from probability.

**Will I win or will I lose?** In the study of random walks, the classic gambler's ruin problem designates two states as targets to be reached and considers the probability of reaching one before the other. Analogously, we can consider starting the KO ball in the "middle" of the puzzle and play a simple game in which we "win" by reaching the exit chamber *C6* before reaching the entrance chamber *C1* (in which case we "lose"). What, then, is the probability of winning given a particular starting state? To answer this, first designate *C1* as a second absorbing state by removing its corresponding row and column from the submatrix $Q$ as previously defined to yield $Q'$. Next, define the $24 \times 2$ matrix $S$ using the rows of $P$ corresponding to the transient states and the columns of $P$ corresponding to the absorbing states. Then, $W = (I - Q')^{-1} S$ is a $24 \times 2$ matrix in which $w_{ij}$ is the probability that the ball reaches the $j$th absorbing state first, conditional on starting from the $i$th transient state.

Column IV of TABLE 1 contains the values $w_{i2}$, the probabilities of winning conditional on starting in state $i$. The probabilities are roughly the same for a majority of starting states, a consequence of the "forgetfulness" property common to all Markov-type systems, which says that only the current state, and not past history, determines future behavior. Once the player reaches any of the "inner" states in the KO graph, play proceeds as though it had started there. Only initial states near the beginning or end, such as *B9* or *B12*, have a distinct advantage in terms of reaching one particular outcome or another. Interestingly, only 7 out of 24 chambers overall are winning states (wherein one is more likely to reach the end of the puzzle if starting in those chambers) and are, not surprisingly, the ones closest to the exit chamber *C6* in the graph.

How does the probability of winning relate to the distance between a given starting chamber and the exit chamber? In FIGURE 3 we see that the probability of winning

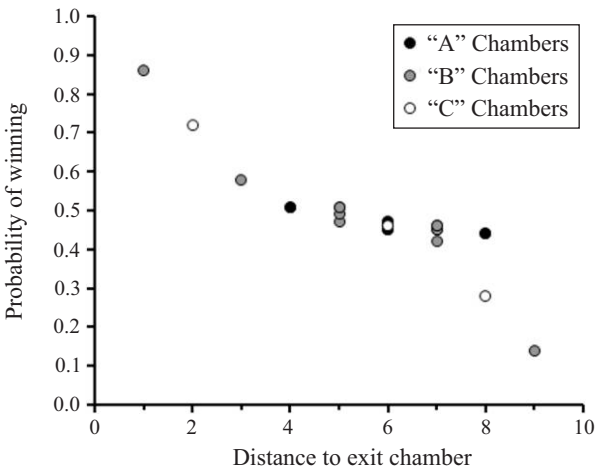



**Figure 3** Probability of winning the gambler's ruin analogue as a function of the starting chamber's distance to the end.

quickly falls to approximately 0.5 as the distance approaches 4. At a distance of 8, we notice an unusual divergence; a player starting in either of the twins *A1* and *A2* has a much higher probability of winning than one starting in *C2*, even though all three are equidistant from *C6*. How can this be? We note that *A1*, *A2*, and *C2* are *not* equidistant from the absorbing (and losing) state *C1*. Because we are considering nongoal-directed behavior, a random walk starting at *C2* is much more likely to be captured by *C1* and lose.

**Off limits!** Returning to our goal-directed player, suppose she now wants to make the puzzle more challenging for herself by declaring certain chambers off limits. The natural question to ask is whether this prohibits her from finding a path through the maze. We can address this question in terms of *cut sets*, that is, sets of vertices that, when deleted from the KO graph, result in a disconnected graph.

Connectivity of the KO graph can be determined by visual inspection; however, a more general method especially suitable for larger graphs or for repeated use is to compute the associated *distance matrix*, $D$, in which $d_{ij}$ is the distance from vertex $i$ to vertex $j$:

$$D = A + \sum_{k=2}^{N_v - 1} k(R^k - R^{k-1})$$

where $A$ is our adjacency matrix, $R = A + I$, $N_v$ is the number of vertices represented in $A$, and the matrix powers are computed as *Boolean products*. If $d_{ij} = 0$ for any $i \neq j$ then there is no path from $i$ to $j$ and the graph is disconnected. For the original KO graph, the entries of the 26th row of $D$ are the distances produced by Dijkstra's algorithm appearing in Column I of TABLE 1.

Considering cut sets of size $n = 1$, a visual inspection of the KO graph reveals that removing any of the singletons *B9*, *C2*, *B1*, *B6*, *C5*, *B12*, or *B4* disconnects the graph. Moreover, we notice that deleting any of the degree-one vertices *C1*, *C6*, *A5*, or *A6* does not disconnect the KO graph. Thus, before proceeding to larger cut sets, we make the decision to temporarily remove the eight vertices *A5*, *A6*, *B9*, *B12*, *C1*, *C2*, *C5*, and *C6* from the KO graph as these can either disconnect, or be disconnected from, the graph trivially and call the subsequent adjacency matrix $A'$. Once we have chosen to ignore these vertices, removing the remaining singletons *B1*, *B4*, or *B6* no longer disconnects the graph.

Moving on to cut sets of size $n > 1$, we examine under $A'$ the $\binom{18}{n}$ unique sets of $n$ vertices that can be deleted. Let $\Delta_n$ be the set of cut sets of size $n$ and $|\Delta_n|$ its cardinality. To construct $\Delta_n$, we systematically choose each subset of size $n$ from the graph, modify the adjacency matrix $A'$ by removing the rows and columns corresponding to that subset, and determine connectedness by computing the associated distance matrix.

Removal of cut sets up to $n = 5$ produces over 12,000 subgraphs whose connectedness needs to be determined. We can reduce the number of computations required by noticing that any candidate set that is a superset of an already-identified cut set will either not be a cut set itself or be a cut set by virtue of containing a cut set. The latter case is in some ways trivial, and we decide to restrict our attention to "novel" cut sets, requiring that no member of $\Delta_n$ be a superset of any member of $\Delta_{n-1}$. TABLE 2 gives the novel cut sets up to $n = 5$ (no larger cut sets exist). We note that only one cut set, $\{C3, C4\}$, can disconnect the graph but still allow the player to complete the maze. As expected, the vertices having the highest degrees are the ones that repeatedly show up in this table.

TABLE 2: Disconnecting Sets

| $n$ | $\lvert\Delta_n\rvert$ | $\Delta_n$ (Cut sets of size $n$) |
|---|---|---|
| 2 | 3 | $\{\{A7, A8\}, \{B4, B5\}, \{C3, C4\}\}$ |
| 3 | 1 | $\{\{B1, B2, B3\}\}$ |
| 4 | 1 | $\{\{B2, B3, B4, C3\}\}$ |
| 5 | 2 | $\{\{A1, A2, A3, A4, C3\}, \{A3, A4, B2, B3, C3\}\}$ |

## Further questions

The KO graph offers undergraduates familiar with graph theory many interesting features for further exploration. For example, notions of the graph's *complexity*, such as its girth, chromatic number, diameter, circumference, variance of vertex degrees, and average distance between vertices can be computed. In addition, while we used the distance matrix to show that the KO graph is connected, students can use an alternate method such as creating a spanning tree of the graph. We can also ask how many total solution paths exist and how many of those do not revisit a vertex. Finally, students can pose questions in the spirit of the Traveling Salesman problem, such as determining the fewest number of vertices that must be removed from the KO graph in order to have a Hamilton path from the entrance chamber *C1* to the exit chamber *C6*.

Other questions involve the mechanics of the puzzle: In our discussion we have defined a "move" as directly passing the ball from one chamber to another, ignoring the physical rotations required to facilitate such a move. If we now consider them, is it possible to find the least number of rotations necessary to perform each of our moves? If so, can we add these as *weights* on the edges of the KO graph and then use Dijkstra's algorithm to determine the least number of rotations required to solve the puzzle? Finally, those interested in puzzle design might also wonder: Does either the expected number of random moves needed to solve the *KO Labyrinth* or the fact that most chambers are nonwinning states under random play serve as a reasonable measure of the difficulty level of the puzzle? How can we use this analysis to create a more challenging maze, and how does that relate to the complexity properties of the graph?

## REFERENCES

1. G. F. Lawler, *Introduction to Stochastic Processes*. Chapman & Hall, Boca Raton, 2006.
2. A. Tucker, *Applied Combinatorics*. John Wiley & Sons, New Jersey, 2007.

**Summary.** The KO Labyrinth is a colorful spherical puzzle with 26 chambers, some of which can be connected via holes through which a small ball can pass when the chambers are aligned correctly. The puzzle can be realigned by performing physical rotations of the sphere in the same way one manipulates a Rubik's Cube, which alters the configuration of the puzzle. The goal is to navigate the ball from the entrance chamber to the exit chamber. We find the shortest path through the puzzle using Dijkstra's algorithm and explore questions related to connectivity of puzzle with the adjacency matrix, distance matrix, and first passage time analysis. We show that the shortest path through the maze takes only 10 moves, whereas a random walk through the maze requires, on average, about 340 moves before reaching the end. We pose an analogue of the gambler's ruin problem and separately consider whether we are able to solve the puzzle if certain chambers are off limits. We conclude with comments and questions for future investigation.

**ALISSA S. CRANS** (MR Author ID: 676843) earned her B.S. in mathematics from the University of Redlands and her Ph.D. in mathematics from the University of California, Riverside under the guidance of John Baez.

She is currently an associate professor of mathematics at Loyola Marymount University, an associate director of Project NExT, and a co-organizer of the Pacific Coast Undergraduate Mathematics Conference. Alissa was a recipient of the 2011 MAA Hasse Prize and Alder Award and has been an invited speaker at the MAA Carriage House, the Museum of Mathematics, and several MAA sectional meetings. She enjoys playing the clarinet with the Santa Monica College wind ensemble, running, biking, dancing, baking, and traveling.

**ROBERT J. ROVETTI** (MR Author ID: 1079109) earned his B.S. from Pacific Union College (Napa Valley, CA) in 1994. After working as a research analyst at the biotechnology firm Amgen Inc., he went on to earn his Ph.D. in biomathematics from the David Geffen School of Medicine at the University of California, Los Angeles, in conjunction with the Cardiovascular Research Laboratory. In 2008 he joined the mathematics department at Loyola Marymount University where he is currently an associate professor. He actively encourages students to work on applied problems that arise in various fields and has supervised several undergraduate research projects.

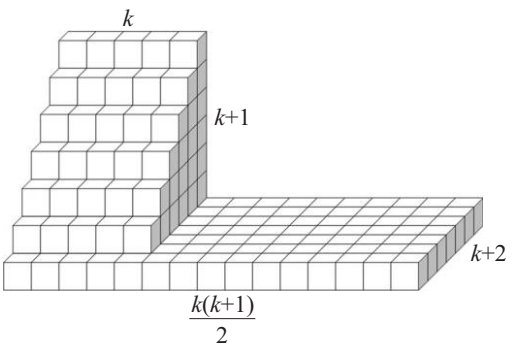
**JESSICA VEGA** (MR Author ID: 955203) earned dual bachelor's degrees in mathematics and theological studies from Loyola Marymount University in 2010. She currently works in Milwaukee, WI, as outcomes coordinator for a nonprofit natural healing center. Other interests include serving on the board of directors for Riverwest Yogashala, as well as teaching yoga in the Iyengar method, facilitating antiracism workshops, and performing and songwriting with Universal Love Band.

# Proof Without Words: Sums of Products of Three Consecutive Integers

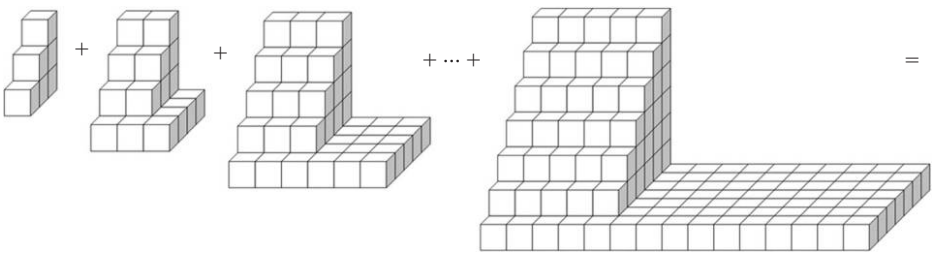
HASAN UNAL
Yildiz Technical University
Istanbul, Turkey
hunal@yildiz.edu.tr

$$\sum_{k=1}^{n} k(k+1)(k+2) = 1 \times 2 \times 3 + 2 \times 3 \times 4 + \cdots + n \times (n+1) \times (n+2)$$

$$= \left[\frac{n(n+1)(n+2)(n+3)}{4}\right]$$

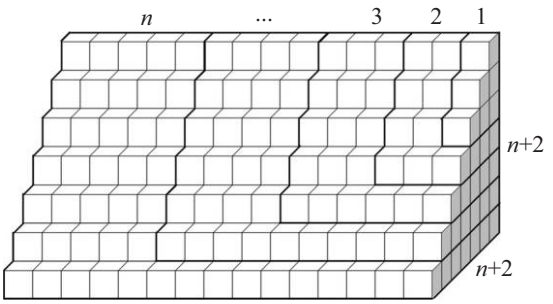



$$k\frac{(k+1)(k+2)}{2} + \frac{k(k+1)}{2}(k+2) = k(k+1)(k+2)$$




$$1 \times 2 \times 3 + 2 \times 3 \times 4 + 3 \times 4 \times 5 + \cdots + n \times (n+1) \times (n+2)$$

*Math. Mag.* **88** (2015) 37–38. doi:10.4169/math.mag.88.1.37. 
MSC: Primary 05A19, Secondary 11B75; 00A66.




$$\frac{n(n+1)}{2} \times \frac{(n+2)(n+3)}{2} = \left[\frac{n(n+1)(n+2)(n+3)}{4}\right]$$

**DR. HASAN UNAL** (MR Author ID: 1059881) received his masters and Ph.D. from Florida State University in 2005. He is currently working at Yildiz Technical University as an Associate Professor in İstanbul, Turkey. He is interested in visualization in teaching and learning mathematics.

# The Median Value of a Continuous Function

IRL C. BIVENS
Davidson College
Davidson, NC 28035
irbivens@davidson.edu

BENJAMIN G. KLEIN
Davidson College
Davidson, NC 28035
beklein@davidson.edu

Virtually every modern calculus textbook contains a discussion of the mean or average value

$$f_{\text{ave}} = \frac{1}{b-a}\int_a^b f(x)\,dx \tag{1}$$

of a continuous function $f$ over a closed interval $[a, b]$. Students learn applications of the average value as well as the result that $f_{\text{ave}}$ is an actual value of the function $f$ (the mean-value theorem for integrals). Noticeably absent however is any mention of the *median* value of $f$. In fact, outside of probability and statistics, one will search in vain for the concept within the undergraduate mathematics curriculum. What has prevented the median from migrating into the calculus class?

At least part of the answer is that the most direct route to the median is through the concept of "measure." Recall that to find the median value of a list of $n$ numbers we arrange the list in increasing order and either take the middle entry if $n$ is odd or average the two middle entries if $n$ is even. It immediately follows that at most half the numbers in the list are greater than the median and at most half are less. However, to extend the concept of "at most half" to the values of a function requires us to have some interpretation of measure. Another problem with the direct approach is that it can be difficult to see why anyone would ever care about the median in the first place. Sure, it might be nice to define it for theoretical purposes, but is the median actually good for anything?

In this article we propose an elementary definition of the median value of a continuous function that can be understood by a student in a first-year calculus course. Furthermore, this definition will allow us to solve immediately some interesting area minimization problems. Exploration of the mathematics associated with the median is an excellent way for students to develop their analytic skills. To this end we have provided a collection of exercises that supplement this article on the Mathematics Magazine website, at http://www.maa.org/publications/periodicals/mathematics-magazine/mathematics-magazine-supplements. Some of the exercises ask the reader to fill in missing steps from proofs or to explore alternative approaches and arguments. Other exercises are more open-ended and could be developed into interesting undergraduate research projects. (Although our article can be read independently of the exercises, we believe that every reader will profit from working at least some of them.)

Our article is organized as follows. First, we'll use a limit to define the median value of a continuous function and we'll preview the role that the median plays in area minimization. Next, we'll consider some examples of area minimization problems for

*Math. Mag.* **88** (2015) 39–51. doi:10.4169/math.mag.88.1.39. 
MSC: Primary 26A06, Secondary 26A42; 26A51.

which the median provides a quick solution. After that, we will use area minimization to prove the existence of the median value for any continuous function on a bounded interval. Then we'll describe an alternative characterization of the median value in terms of Lebesgue measure. This characterization will allow us to interpret the median value of a continuous function in terms of the median value of a random variable. The area minimization property of the median then can be viewed as a special case of an interesting minimization principle in probability.

## The median value

Suppose that $f$ is a continuous function on a closed interval $[a, b]$. Partition $[a, b]$ into $n$ equal subintervals and let $x_1^*, x_2^*, \ldots, x_n^*$ signify the midpoints of these subintervals. Let ave$(n)$ and med$(n)$ denote the respective mean and median of the $n$ function values, $f(x_1^*), f(x_2^*), \ldots, f(x_n^*)$. Most calculus textbooks contain a proof that

$$f_{\text{ave}} = \lim_{n \to +\infty} \text{ave}(n). \tag{2}$$

Motivated by equation (2) we define the median value of $f$ on $[a, b]$ to be

$$f_{\text{med}} = \lim_{n \to +\infty} \text{med}(n) \tag{3}$$

provided this limit exists.

For a simple example, suppose that $f$ is monotonic (that is, nonincreasing or nondecreasing) on $[a, b]$. Then it follows easily from (3) that $f_{\text{med}} = f((a + b)/2)$. In other words, to find the median we simply evaluate the function at the midpoint of $[a, b]$ (FIGURE 1). For example, the median value of $f(x) = x^2$ on $[0, 2]$ is 1.

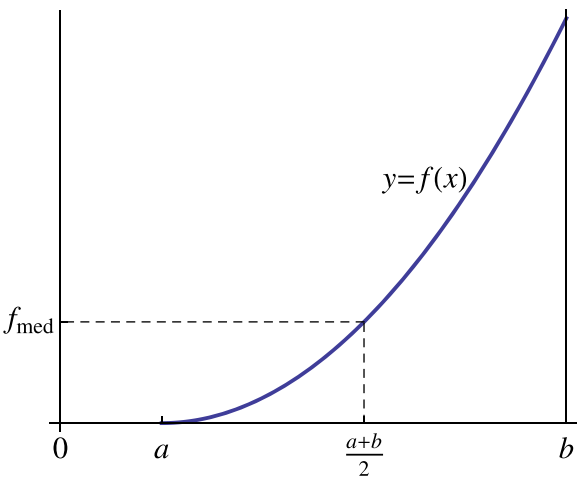



**Figure 1** The median value of a monotonic function.

More generally, the same conclusion holds if the graph of $f$ lies on or above the horizontal line $y = f((a + b)/2)$ on one side of the vertical line $x = (a + b)/2$, and the reverse is the case on the other side of this vertical line (FIGURE 2). For instance, the median value of $f(x) = x^2$ on $[-1, 3]$ is also 1.

In these simple cases the median can be seen to exist by inspection. We will prove that $f_{\text{med}}$ always exists and yields the unique horizontal line $y = f_{\text{med}}$ that minimizes the area

$$A(t) = \int_a^b |f(x) - t|\, dx \tag{4}$$

between a horizontal line $y = t$ and the graph of $y = f(x)$ over $[a, b]$ (FIGURE 3).

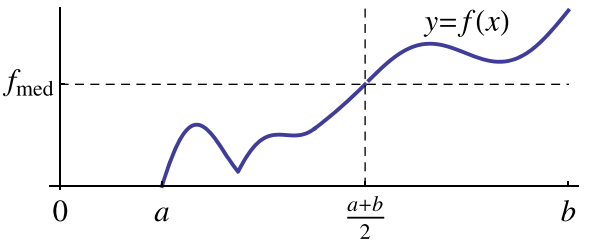



**Figure 2** $f_{\text{med}} = f((a+b)/2)$.

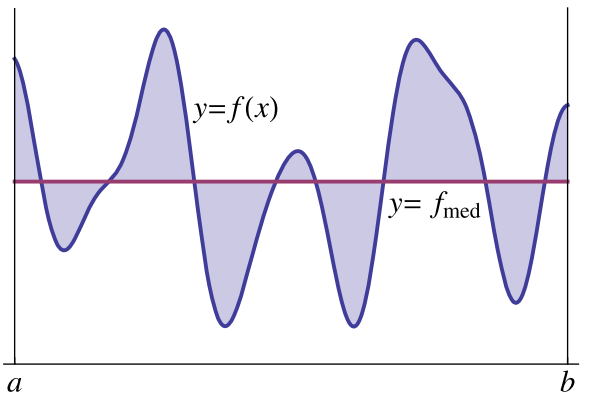



**Figure 3** The line $y = f_{\text{med}}$ yields minimum shaded area.

The reader may find this conclusion to be surprising in the case of FIGURE 2. Here the area minimizing horizontal line depends on the value of $f$ at the midpoint of $[a, b]$, but is otherwise independent of $f$ (FIGURE 4).

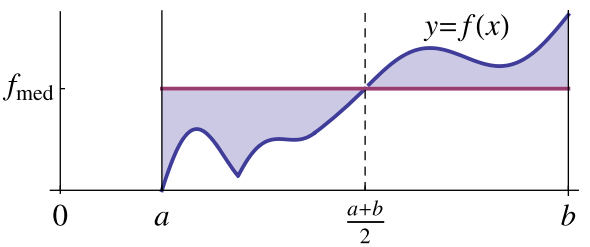



**Figure 4** Minimum shaded area.

Shouldn't the area minimizing line depend upon more than a single value of $f$?

More generally we will show that the existence of $f_{\text{med}}$ is guaranteed for any function $f$ that is continuous on a bounded interval. This will ensure the existence of a median value for functions such as $\sin(1/x)$ or $\sin(1/x)/x$ on $(0, 1]$. Furthermore, as long as the relevant areas remain finite, the horizontal line $y = f_{\text{med}}$ will continue to be area minimizing.

## The median and area minimization

The connection between $f_{\text{med}}$ and area minimization results from an important minimizing property possessed by the median value of any list $y_1, y_2, \ldots, y_n$ of real numbers. If $M$ denotes the median value of the list, then for every real number $t$

$$\sum_{k=1}^{n} |y_k - M| \leq \sum_{k=1}^{n} |y_k - t|. \tag{5}$$

The basic idea behind inequality (5) is simple. If more than half the numbers in the list are greater than some real number $t$, then increasing $t$ slightly will reduce the size of the sum $\sum_{k=1}^{n} |y_k - t|$. Likewise, if more than half the numbers are less than $t$, then decreasing $t$ slightly will reduce the sum. The minimum value of the sum is achieved when at most half the numbers in the list are greater than $t$ and at most half the numbers are less than $t$. Hence, the minimum value of $\sum_{k=1}^{n} |y_k - t|$ is achieved if $t = M$. (A more thorough discussion of this result is given by Farnsworth in [**3**].)

We will use this property of the median to derive an inequality between midpoint approximations of integrals. Suppose that $f$ is a continuous function on a closed interval $[a, b]$. Let $x_1^*, x_2^*, \ldots, x_n^*$ denote the midpoints of the subintervals of a regular partition of $[a, b]$ of order $n$, and let

$$\mathcal{M}_n(t) = \frac{b-a}{n} \sum_{k=1}^{n} |f(x_k^*) - t|$$

denote the corresponding midpoint approximation of $A(t) = \int_a^b |f(x) - t|\, dx$. If $\text{med}(n)$ is the median of the $n$ function values $f(x_1^*), f(x_2^*), \ldots, f(x_n^*)$, then it follows from (5) that for any real number $t$

$$\sum_{k=1}^{n} |f(x_k^*) - \text{med}(n)| \leq \sum_{k=1}^{n} |f(x_k^*) - t|.$$

Multiplying both sides of this inequality by $(b-a)/n$ we obtain the inequality

$$\mathcal{M}_n(\text{med}(n)) \leq \mathcal{M}_n(t) \tag{6}$$

for every natural number $n$ and for every real number $t$.

Inequality (6) is the linchpin between the median and area minimization. For example, in FIGURE 1 we have $f_{\text{med}} = f((a+b)/2) = \text{med}(2n-1)$. It then follows from (6) that $\mathcal{M}_{2n-1}(f_{\text{med}}) \leq \mathcal{M}_{2n-1}(t)$ for any real number $t$ and any natural number $n$. Since $\mathcal{M}_{2n-1}(f_{\text{med}}) \to A(f_{\text{med}})$ and $\mathcal{M}_{2n-1}(t) \to A(t)$ as $n \to +\infty$, we conclude that $A(f_{\text{med}}) \leq A(t)$. Hence, the line $y = f_{\text{med}}$ minimizes the area between the graph of $f$ and a horizontal line over $[a, b]$. (This argument also justifies FIGURE 4.)

The examples that follow provide some additional applications of the median.

**Example 1.** Find a horizontal line such that the area between that line and the graph of $y = f(x) = \frac{1}{2}(x + \sin 3x)$ over the interval $[a, b] = [0, 2\pi]$ is a minimum.

*Solution.* Note from FIGURE 5 that the graph of $f$ is centrally symmetric about its midpoint $(\pi, \pi/2)$. It follows that the set of function values $f(x_1^*), f(x_2^*), \ldots, f(x_n^*)$ is centrally symmetric about $y = \pi/2$. Hence $\text{med}(n) = \pi/2$ for all $n$ and (6) becomes $\mathcal{M}_n(\pi/2) \leq \mathcal{M}_n(t)$. Letting $n \to +\infty$ shows that $A(\pi/2) \leq A(t)$ for all $t$ and thus the horizontal line $y = \pi/2$ is area minimizing.

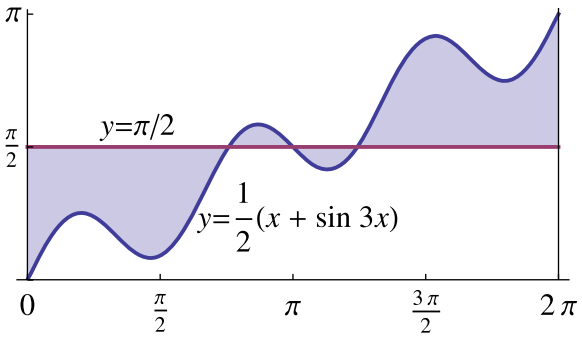



**Figure 5** Minimum shaded area.

More generally, if the graph of a continuous function $f$ over $[a, b]$ is centrally symmetric about its midpoint $((a + b)/2, f((a + b)/2))$, then the area minimizing horizontal line will have equation $y = f((a + b)/2)$. (Can the reader imagine any other simple way to prove this result?)

**Example 2.** Wile E. Coyote has donned roller skates and a jetpack in an effort to catch the Road Runner. If things go according to plan, then $x$ seconds after igniting the jetpack Mr. Coyote will have traveled a distance of $6x^2$ feet in a straight line from his initial position and he will catch the Road Runner at time $x = 30$. However, experience has taught Wile E. that things will go disastrously wrong at some random moment $0 \le x \le 30$. In preparation for the inevitable, he decides to stash a first aid kit somewhere along his anticipated route. Where should he place the first aid kit?

*Solution.* Let $0 \le X \le 30$ denote the moment at which things go awry. If we assume that all such moments are equally likely, then $X$ is a continuous random variable uniformly distributed over [0, 30] with constant probability density function $g(x) = 1/30$. Let $Y = 6X^2$ denote the distance from Mr. Coyote's initial position at the moment disaster strikes and let $t$ denote the distance from his initial position at which he stashes the first aid kit. He needs to choose $t$ so as to minimize the expected distance

$$E[|Y - t|] = \frac{1}{30}\int_0^{30} |6x^2 - t|\, dx$$

between $t$ and the composite random variable $Y$. Since $6x^2$ is an increasing function on [0, 30], the minimizing value is $t = 6 \cdot 15^2 = 1350$ and he should place the first aid kit a distance of 1350 feet from his initial position. (The expected distance between the first aid kit and the point of disaster is $E[|Y - 1350|] = 1350$ feet. The *mean* value of $Y$ is 1800. Were he to choose the mean instead, the expected distance would be $E[|Y - 1800|] = 800\sqrt{3} \approx 1386$ feet.)

By making a change of variables the median can be used to solve area minimization problems with respect to other one-parameter families of curves.

**Example 3.** Find the (nonvertical) line through the origin that minimizes the area between the graph of $y = x^2$ and the line over the interval [0, 1].

*Solution.* If $t$ is the slope of the line, then the area in question is given by

$$\int_0^1 |x^2 - tx|\, dx = \int_0^1 |x - t|x\, dx = \int_0^{1/2} |\sqrt{2u} - t|\, du$$

(using the substitution $u = x^2/2$, $du = x\,dx$). The third integral displayed is $A(t)$ for the function $f(u) = \sqrt{2u}$ on the interval $[0, 1/2]$. Since $f$ is an increasing function on this interval the area minimizing value of $t$ is $t = f_{\text{med}} = f(1/4) = 1/\sqrt{2}$. FIGURE 6 displays our solution to the original problem.

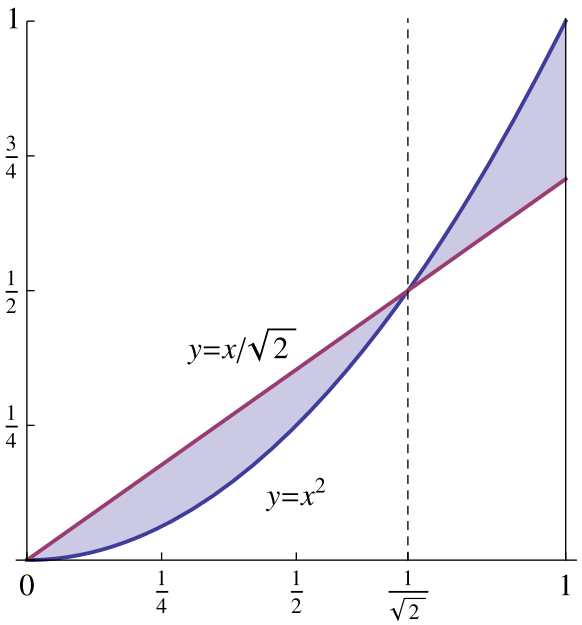



**Figure 6** Minimum shaded area.

## The median value of a continuous function on a bounded interval

In the last section we used the existence of the median to solve some area minimization problems. In this section we reverse course and use area minimization to prove the existence of the median for any function that is continuous on a bounded interval. The proof will rely on some unexpectedly simple behavior of the area function $A(t)$.

Suppose that $f$ is a continuous function on $[a, b]$ and a line $y = t$ does not intersect the graph of $f$. Then it is clear that $A(t)$ will decrease if we move the line towards the graph of $f$ and $A(t)$ will increase if we move the line away from the graph. However, it is less evident how $A(t)$ changes when the line and the graph of $f$ intersect. For example, consider the shaded area in FIGURE 7.

As we increase the value of $t$ from that displayed in the figure, some of the shaded regions will increase in area while others will decrease. It is not immediately clear what the net effect of all these increases and decreases will be. Surprisingly, no matter how unruly the graph of the continuous function $f$, there will be a strict decrease in shaded area as the horizontal line rises to the position displayed in FIGURE 3, after which there will be a strict increase in shaded area.

To see why, recall that a real-valued function on $\mathbb{R}$ is *convex* provided that over any interval $[t_0, t_1]$ the graph of the function lies on or below the corresponding chord of the graph over $[t_0, t_1]$. For fixed $x$, the integrand in the definition (4) of $A(t)$ is a shifted absolute value function and thus is a convex function of $t$. It then follows from the linearity and order properties of the definite integral that $A(t)$ is also convex and hence continuous. (See Roberts and Varberg [**6**] for properties of convex functions.) It

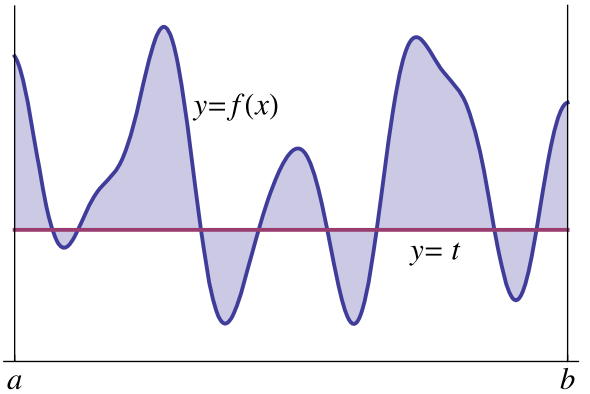



**Figure 7** How does the shaded area depend upon $t$?

is also clear from the definition of $A(t)$ that

$$\lim_{t\to-\infty} A(t) = +\infty \quad \text{and} \quad \lim_{t\to+\infty} A(t) = +\infty$$

and thus $A(t)$ will have an absolute minimum value.

Using the continuity of $f$ we now argue that the minimum value of $A(t)$ is achieved at a unique real number $t_m$. Note that if $t_0 < f(x) < t_1$, then the graph of $t \to |f(x) - t|$ lies strictly below the corresponding chord over $(t_0, t_1)$. Since $f$ is continuous the area function $A(t)$ inherits this strictness property from $|f(x) - t|$. More precisely, if there exists a value of $x$ in $[a, b]$ such that $t_0 < f(x) < t_1$, then the graph of $A(t)$ will lie strictly below the corresponding chord over $(t_0, t_1)$. Suppose now that the minimum of $A(t)$ is achieved at two distinct values $t_0 < t_1$. It then follows from convexity that the graph of $A(t)$ will coincide with the corresponding chord over $[t_0, t_1]$. On the other hand, if no point on the graph of $f$ is between the horizontal lines $y = t_0$ and $y = t_1$, then by continuity the graph of $f$ must lie on or below the line $y = t_0$ or must lie on or above the line $y = t_1$. In either case, one of the two lines $y = t_0$ or $y = t_1$ will not be area minimizing, a contradiction. Hence, there must exist $a \le x \le b$ such that $t_0 < f(x) < t_1$. But then the graph of $A(t)$ has to lie strictly below the corresponding chord over $(t_0, t_1)$, another contradiction. We conclude that $A(t)$ achieves its minimum value at a unique real number $t_m$. (Readers who prefer a more analytic argument are referred to the article supplement. Alternatively, see the argument by Noah [**5**] where a multivariable uniqueness result is proved.)

It now follows from convexity that $A(t)$ is strictly decreasing on the interval $(-\infty, t_m]$ and is strictly increasing on the interval $[t_m, +\infty)$. More generally, these same conclusions hold for any continuous function $f$ on a bounded interval as long as $f$ is absolutely integrable on the interval. (That is, for which the integral $\int_a^b |f(x)|\,dx$, proper or improper, exists and is finite.) In summary, we have the following result.

**Theorem 1.** *Suppose that a function $f$ is continuous and absolutely integrable on a bounded interval. Then the area function $A(t)$ is convex (and hence continuous everywhere) and satisfies*

$$\lim_{t\to-\infty} A(t) = +\infty \quad \text{and} \quad \lim_{t\to+\infty} A(t) = +\infty.$$

*Furthermore, there exists a unique real number $t_m$ such that the area function is strictly decreasing on the interval $(-\infty, t_m]$ and is strictly increasing on the interval $[t_m, +\infty)$. In particular, $A(t)$ is minimized uniquely at $t = t_m$.*

Assume that $f$ is a bounded continuous function on an open interval $(a, b)$. Our next goal is to use inequality (6) and Theorem 1 to prove that $f_{\text{med}}$ exists and is equal to the unique parameter value $t = t_m$ that minimizes $A(t)$. We see from Theorem 1 that to prove

$$\lim_{n\to+\infty} \text{med}(n) = t_m \tag{7}$$

it suffices to prove that

$$\lim_{n\to+\infty} A(\text{med}(n)) = A(t_m).$$

That is, for every $\varepsilon > 0$, we must show that there exists a natural number $N$ such that

$$A(t_m) \le A(\text{med}(n)) < A(t_m) + \varepsilon \quad \text{for all} \quad n \ge N. \tag{8}$$

Let $\varepsilon > 0$ be given. Since $\lim_{n\to+\infty} \mathcal{M}_n(t) = A(t)$ it follows that for any fixed value of $t$ there exists a natural number $N$ such that

$$|A(t) - \mathcal{M}_n(t)| < \frac{\varepsilon}{2} \quad \text{if} \quad n \ge N. \tag{9}$$

In fact, $N$ can be chosen such that inequality (9) is valid for all real numbers $t$. (When it comes to approximating an integral with a Riemann sum, an integrand of $|f(x) - t|$ is not all that different from an integrand of $f(x)$. For a desired degree of accuracy, a choice of $N$ that "works" for the integral of $f$ also works for any $A(t)$.) Assume we make such a choice of $N$.

Taking $t = t_m$ in (9) yields

$$\mathcal{M}_n(t_m) < A(t_m) + \frac{\varepsilon}{2} \quad \text{if} \quad n \ge N \tag{10}$$

and taking $t = \text{med}(n)$ in (9) yields

$$A(\text{med}(n)) < \mathcal{M}_n(\text{med}(n)) + \frac{\varepsilon}{2} \quad \text{if} \quad n \ge N. \tag{11}$$

Then, if $n \ge N$ we have

$$\begin{aligned}
A(t_m) &\le A(\text{med}(n)) && \text{(since } A(t_m) \text{ is the minimum value of } A(t)) \\
&< \mathcal{M}_n(\text{med}(n)) + \varepsilon/2 && \text{(by inequality (11))} \\
&\le \mathcal{M}_n(t_m) + \varepsilon/2 && \text{(by inequality (6) with } t = t_m) \\
&< (A(t_m) + \varepsilon/2) + \varepsilon/2 && \text{(by inequality (10))} \\
&= A(t_m) + \varepsilon.
\end{aligned}$$

These inequalities yield (8), which completes the proof of (7). We conclude that if $f$ is a bounded continuous function on an open interval $(a, b)$, then $f_{\text{med}}$ exists and minimizes $A(t)$.

Next, suppose that $f$ is an unbounded continuous function on an open interval $(a, b)$. Then $f$ is not Riemann integrable and we cannot apply the previous argument directly to $f$. Fortunately, what saves us is the discrete median's well-known resistance to "outliers." The idea is to "truncate" $f$ near the endpoints $a$ and $b$ so as to produce a bounded continuous function $h$ on $(a, b)$ with the property that $\text{med}(n)$ is the same for both $f$ and $h$. Then

$$f_{\text{med}} = \lim_{n\to+\infty} \text{med}(n) = h_{\text{med}}$$

exists. In addition, if the improper integral $\int_a^b |f(x)|\,dx$ converges, then $A(t)$ is defined for all real numbers $t$ and $t = t_m = f_{\text{med}}$ minimizes $A(t)$. (Refer to the article supplement for details, at http://www.maa.org/publications/periodicals/mathematics-magazine/mathematics-magazine-supplements.)

In summary we have the following theorem.

**Theorem 2.** *The median value $f_{\text{med}}$ exists for any continuous function $f$ on a bounded interval. Furthermore, if $f$ is absolutely integrable on the interval, then $t = t_m = f_{\text{med}}$ is the unique parameter value that minimizes the area function $A(t)$.*

Theorem 2 can be used to explain some puzzling behavior associated to area minimization problems similar to that of Example 3. Suppose that $f$ is any continuous function on [0, 1] and that $l$ denotes the nonvertical line through the origin that minimizes the area between the graph of $f$ and the line over [0, 1]. If the graph of $f$ and $l$ intersect once over (0, 1), then they will cross one another at their point of intersection, and that point must lie on the vertical line $x = 1/\sqrt{2}$.

Conversely, suppose the graph of $f$ and a line $y = tx$ cross one another exactly once at a point on the vertical line $x = 1/\sqrt{2}$. Then the area between the graph of $f$ and the line will be a minimum (FIGURE 8).

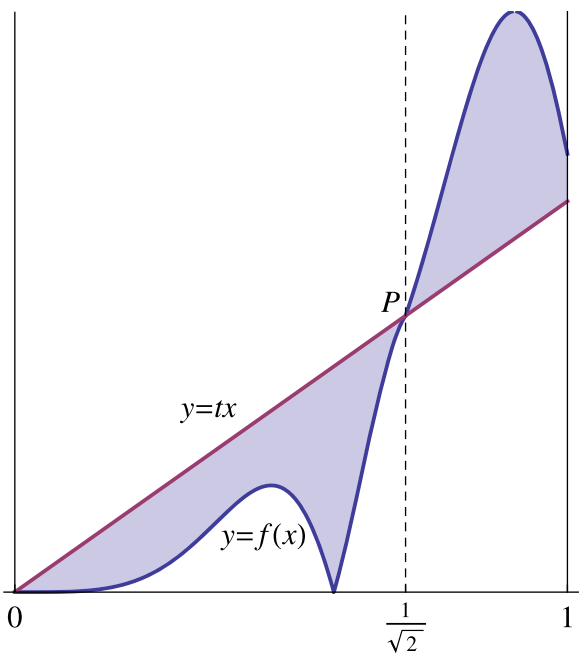



**Figure 8** Minimal area between the curve $y = f(x)$ and a line $y = tx$.

In either scenario, when $f$ is nonnegative this minimal area can be expressed as the positive difference between areas to the left and to the right of the line $x = 1/\sqrt{2}$ (FIGURE 9). It was the authors' attempts to understand the surprising phenomena illustrated in FIGURES 8 and 9 that led to the present article. (The justification of these results is one of the exercises in the article supplement.)

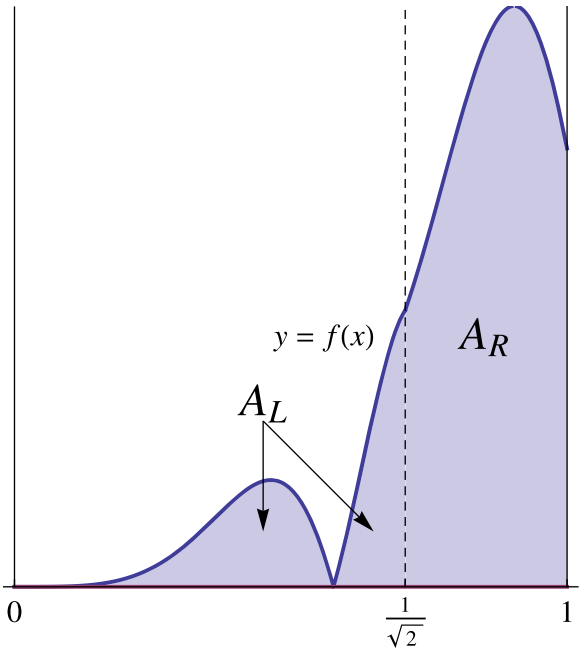



**Figure 9** The minimal area in FIGURE 8 is equal to $A_R - A_L$.

## The median and measure

Assume that $f$ is continuous on an interval with endpoints $a < b$. In this section we provide an alternative characterization of the median in terms of the measure of certain open sets that are naturally associated to $f$. Define two families of sets

$$\text{Above}(t) = f^{-1}((t, \infty)) = \{x \in (a, b) \mid f(x) > t\}$$
$$\text{Below}(t) = f^{-1}((-\infty, t)) = \{x \in (a, b) \mid f(x) < t\}.$$

It follows from the continuity of $f$ that for a given value of $t$, Above$(t)$ and Below$(t)$ are each open sets of real numbers. In general Above$(t)$ is the union of a countable collection of disjoint open intervals, and we let $m[\text{Above}(t)]$ denote the sum of the lengths of these intervals. (Here "sum" is interpreted as the sum of an infinite series if Above$(t)$ is equal to the union of a countably infinite collection of disjoint open intervals. More formally, $m[\text{Above}(t)]$ is the Lebesgue measure of Above$(t)$.) Similar comments and notation apply to Below$(t)$ and we have the following characterization of the median.

**Theorem 3.** *Suppose that $f$ is a continuous function defined on an interval with endpoints $a < b$. Then the median value $f_{\text{med}}$ is characterized by the pair of inequalities*

$$m[\text{Below}(f_{\text{med}})] \le \frac{b-a}{2} \quad \text{and} \quad m[\text{Above}(f_{\text{med}})] \le \frac{b-a}{2}.$$

(A proof of Theorem 3 is developed in the article supplement, at http://www.maa.org/publications/periodicals/mathematics-magazine/mathematics-magazine-supplements.)

Note that the inequalities in Theorem 3 agree with our intuition; at most half of the values of $f$ are below the median value and at most half are above. In addition, the inequalities in Theorem 3 explain the insensitivity of $f_{\text{med}}$ to changes in $f$. Changes

in $f$ that leave the sets Below($f_{\text{med}}$) and Above($f_{\text{med}}$) unaltered have no effect on the value of $f_{\text{med}}$. Contrast this with the value of $f_{\text{ave}}$, which is very sensitive to changes in the values of $f$.

Readers familiar with measure theory will recognize the inequalities in Theorem 3 as the definition of a median value of a measurable function. However, it's important to emphasize that this measure theoretic definition is guaranteed to agree with our limit definition (3) only if the function is continuous. For example, consider the function on $[0, 1]$ that is 0 on the rational numbers and 1 on the irrational numbers. According to our definition the median value of this function is 0, while according to the measure theoretic definition the median value is 1.

Theorem 3 provides us with a useful alternative to (3) when we wish to compute a median value.

**Example 4.** Find the median value of $f(x) = \sin x$ on the interval $[-\pi/2, \pi]$.

*Solution*. Since every horizontal line intersects the graph of $f$ a finite number of times, both inequalities in Theorem 3 become equalities, and either equality yields the median. Referring to FIGURE 10 we see that given any value of $0 < t < 1$ we have $m[\text{Above}(t)] = \pi - 2r_t$. We seek the value of $t$ such that $m[\text{Above}(t)] = \pi - 2r_t = (\pi - (-\pi/2))/2 = 3\pi/4$. Hence, $r_t = \pi/8$ and $f_{\text{med}} = \sin(\pi/8) = \sqrt{2 - \sqrt{2}}/2$.

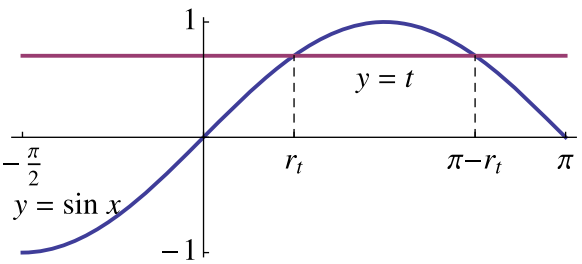



**Figure 10** Above$(t) = (r_t, \pi - r_t)$.

## The median value and probability

Suppose that $Y$ denotes a random variable. A *median value* of $Y$ is defined to be any real number $\alpha$ such that $P[Y < \alpha] \leq 1/2$ and $P[Y > \alpha] \leq 1/2$. (In general such a median value need not be unique.) If the expectation $E[|Y|]$ is finite, then the expectation $E[|Y - t|]$ is known as the average absolute deviation function or as the first-order absolute moment of $Y$ about $t$. It is a well-known principle in probability that the average absolute deviation function is minimized by any median value of $Y$. This principle is illustrated in Example 2 with the placement of Mr. Coyote's first aid kit. Our solution to Example 2 shows that the average absolute deviation function is minimized by $t = 1350$. On the other hand, $\alpha = 1350$ is also the (unique) median value of the random variable $Y$ in that example since $P[Y < 1350] = 1/2 = P[Y > 1350]$.

Although well known, proofs of this minimizing property of the median can be surprisingly hard to find. A proof that appears in Cramér's classic text [**1**] uses identities that involve Riemann–Stieltjes integrals. However, the justification of these identities is left as an exercise. A similar argument appears as an exercise in Hogg and Craig [**4**] using Riemann integration along with the assumption that $Y$ has a probability density function. A complete proof appears in DeGroot [**2**] within the context of loss functions.

The area minimization property of $f_{\text{med}}$ can be interpreted as a special case of this probability principle. Suppose that $f$ is a continuous absolutely integrable function

on an interval with endpoints $a < b$. Let $X$ denote a random variable that is uniformly distributed over the interval and let $Y = f(X)$ denote the associated composite random variable. (In practical terms, $Y$ is a random variable that can be simulated using a uniform random number generator $X$ and the function $f$.) Since

$$P[Y < t] = \frac{m[\text{Below}(t)]}{b-a} \quad \text{and} \quad P[Y > t] = \frac{m[\text{Above}(t)]}{b-a},$$

the inequalities $P[Y < \alpha] \le 1/2$ and $P[Y > \alpha] \le 1/2$ are equivalent to

$$m[\text{Below}(\alpha)] \le \frac{b-a}{2} \quad \text{and} \quad m[\text{Above}(\alpha)] \le \frac{b-a}{2}.$$

We conclude from these inequalities and Theorem 3 that $\alpha$ is a median value of $Y$ if and only if $\alpha = f_{\text{med}}$. Hence, the minimization of $A(t)$ by $t = f_{\text{med}}$ is equivalent to the minimization of the average absolute deviation function

$$E[|Y - t|] = \frac{1}{b-a}\int_a^b |f(x) - t|\, dx = \frac{A(t)}{b-a}$$

by the median value of $Y$.

## Additional approaches, applications, and analysis

In this article we proved the existence of $f_{\text{med}}$ by using a minimization property of the discrete median. But other properties can be employed to prove existence. For example, if all the entries in a list of numbers are changed by a small amount, then the median also changes by a small amount. One of our referees used this result to create a beautifully simple proof that $f_{\text{med}}$ exists for a continuous function on a closed interval. On the other hand, if most of the entries in a list are nearly the same, then changing a few entries by large amounts still only changes the median by a small amount. (In the terminology of statistics, the median is a "resistant" measure of central tendency.) Adding this property to our referee's argument extends the proof to a continuous function on an open interval. Yet another, (measure theoretic) existence proof for $f_{\text{med}}$ uses the fact that the discrete median is in the middle of a list.

Example 3 showed that the median can be applied to area minimization with respect to a nonhorizontal family of lines. More generally, the median can be applied directly to area minimization with respect to numerous other one-parameter families of curves.

All these ideas (and many more) are developed in the article supplement. It is our hope that this article, and the accompanying supplement on the Magazine's website, will inspire readers to make their own investigations into the fascinating concept of the median.

**Acknowledgment** We thank the referees for their careful review and the editor for his patient guidance in matters both mathematical and grammatical. (Special thanks go to the referee who showed us the extremely clever existence proof for the median value of a continuous function on a closed interval.) Their insightful comments helped us to clarify the goals of our article and to direct our exposition accordingly.

## REFERENCES

1. H. Cramér, *Mathematical Methods of Statistics.* Princeton Univ. Press, Princeton, NJ, 1999.
2. M. H. DeGroot, *Optimal Statistical Decisions.* Wiley Classics Library, New York, 2004, http://dx.doi.org/10.1002/0471729000.

3. D. L. Farnsworth, The sum of absolute deviations and the sum of squared deviations, *Amer. Math. Monthly* **97** (1990) 911–912, http://dx.doi.org/10.2307/2324330.
4. R. Hogg, A. Craig, *Introduction to Mathematical Statistics*. Fourth edition. Macmillian Publishing, New York, 1978.
5. S. G. Noah, The median of a continuous function, *Real Anal. Exchange* **33** (2007/2008) 269–274.
6. A. W. Roberts, D. E. Varberg, *Convex Functions*. Academic Press, New York, 1973.

**Summary.** Using a limit we define the median value of a continuous function on a bounded interval and we explicate the connection between the median and area minimization. With the help of Lebesgue measure we show that this connection can be viewed as a special case of an interesting minimization principle in probability.

**IRL BIVENS** (MR Author ID: 37440) received his A.B. degree from Pfeiffer College and his Ph.D. from the University of North Carolina at Chapel Hill. After teaching at Pfeiffer College and Rice University he joined the Faculty of Davidson College in 1982. At Davidson he has taught a variety of courses including differential geometry, a course on mathematics and magic, and a popular seminar on the history of mathematics. For relaxation he swims, reads, and attempts to juggle.

**BEN KLEIN** (MR Author ID: 336593) received his B.A. degree in mathematics from the University of Rochester and then earned an M.A. and Ph.D. (under G.A. Hedlund) in mathematics from Yale University. After four years in the Department of Mathematics at New York University, he moved to North Carolina in 1971 and served on the Faculty of Davidson College for thirty-seven years. He retired from full-time teaching in 2008 but has taught, off and on, on a part-time basis ever since. He has been active in the Mathematical Association of America (including service as Chair of the Southeastern Section), the North Carolina Council of Teachers of Mathematics, and the Advanced Placement Calculus Program.

**ACROSS**

1. Up and at'em
6. Title of Bette Midler's debut album, a reference to her stage persona: The Divine ____
11. "You've got mail" co.
14. Playful prank
15. Greek letter used by Wolfram
16. Letters on the 6 key of a telephone
17. * $x^2 + y^2 - z^2 = -1$, e.g.
19. ____ Mae Brown (Whoopi Goldberg's role in "Ghost")
20. Neither's partner
21. Proven conjecture
23. * $x^2 + y^2 + 2z^2 = 1$, e.g.
28. Dust busters, for short
29. 80s rocker Lewis, 30s Senator Long, and Black Panther founder Newton
30. Crawl sideways on all fours
34. Cathedral recess
35. Alternative name for a manatee
36. What all the starred clues and answers are
42. How soccer games often, but baseball games rarely, end
43. Unix shell that replaced Bourne
44. Place of horror and torture in European legend, but nowadays a fan website for the Arizona Diamondbacks
47. Explorer Polo
48. A linear one between normed spaces is bdd. iff it is cts.
49. * $x^2 + y^2 = 1$, e.g.
51. Archbishop Tutu
55. Adjective that some use for mathematicians over 40 (or 30, or whatever)
56. Kellogg's Cracklin' ____ Bran
57. * $z = y^2 + x^2$ and $z = y^2 - x^2$, e.g.
63. www.maa.org, for example
64. Phrase you might utter in response to an Ace-high hand: "Well, I've got ____"
65. Boredom
66. Prestigious UK public research univ., whose fictitious alumni include Pres. Josiah Bartlett ("The West Wing") and James Bond's father
67. Mini-version of 21-Across
68. Scolded onomatopoetically

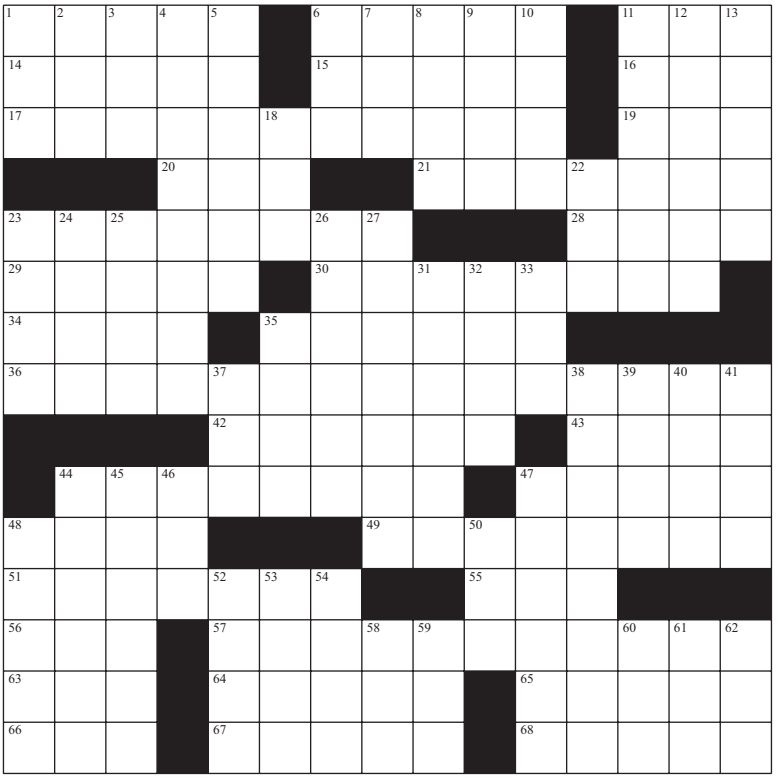
**DOWN**

1. "Oh, good heavens!", in German: "____ du lieber!"
2. Proof technique: counting in two ____s
3. Foursquare or Instagram, for instance
4. Having great, focused vision
5. Mistakes on homework
6. Chairman ____
7. Sick
8. Classic Steven Wright joke: "I spilled ____ remover on my dog; now he's gone"
9. Dog breed: ____ tzu
10. Fashioned
11. Unprincipled
12. Up next, in sports slang
13. Rich soils
18. Frat buddy
22. Eggs, in the lab
23. NBA star who also starred in "Kazaam"
24. Kind of platter at a Polynesian restaurant
25. Disney song sung by Lady's friend Peg: "____ Tramp"
26. Polar global-warming concern
27. Severe
31. Mental sharpness
32. A bad lecture might be one
33. Conservation grp. with a panda logo
35. Opposite/Hypotenuse
37. Former NBA player Smits nicknamed The Dunkin' Dutchman
38. Lets go, as in a potential proof
39. "Pick a ____ . . . "
40. Verb suffix for fluor-
41. MIT's Peter who found a quantum computing algorithm that factors in $O((\log N)^3)$ time
44. Servings of asparagus or pickles
45. Maker of Butterfingers
46. Either of two rays that make up an angle
47. Cereal plant grown as food
48. Namesake for a brand of nonalcoholic beer
50. Bathroom, in Cambridge
52. October birthstone
53. Back of the neck
54. Unit of mass equaling 1/16 ounce, and unit of volume equaling 1/8 fluid ounce
58. Order between "ready" and "fire"
59. A famous "Seinfeld" has Kramer inventing this, but designed for men
60. It used to be kept in wells
61. Required to be turned in, as with homework
62. ____ Vicious: name shared by the Sex Pistols's bassist and a pro wrestler

*Math. Mag.* **88** (2015) 52–53. doi:10.4169/math.mag.88.1.52. 

# 2D or not 2D?

BRENDAN W. SULLIVAN
Emmanuel College
Boston, MA 02115
sullivanb@emmanuel.edu





Clues start at left, on page 52. The solution is on page 71.

Extra copies of the puzzle, in both .pdf and .puz (AcrossLite) formats, can be found at www.maa.org/mathmag/supplements.

# The James Function

CHRISTOPHER N. B. HAMMOND
WARREN P. JOHNSON
Connecticut College
New London, CT 06320
cnham@conncoll.edu
wpjoh@conncoll.edu

STEVEN J. MILLER
Williams College
Williamstown, MA 01267
sjm1@williams.edu

## 1 Introduction

In his 1981 *Baseball Abstract* [**8**], Bill James posed the following problem: suppose two teams $A$ and $B$ have winning percentages $a$ and $b$, respectively, having played equally strong schedules in a game such as baseball where there are no ties. If $A$ and $B$ play each other, what is the probability $p(a, b)$ that $A$ wins?

This question is perhaps more relevant to other sports, because in baseball the outcome is particularly sensitive to the pitching matchup. (In 1972, the Philadelphia Phillies won 29 of the 41 games started by Steve Carlton, and 30 of the 115 games started by their other pitchers.) The answer is quite interesting, even if its applicability is somewhat limited by the tacit assumption of uniformity.

For $0 < a < 1$ and $c > 0$, define $q_c(a)$ by

$$a = \frac{q_c(a)}{q_c(a) + c}. \tag{1.1}$$

James calls $q_{\frac{1}{2}}(a)$ the log5 of $a$, and does not consider any other values of $c$. Under the assumption of uniformity, he claims that $p(a, b)$ would be given by the function

$$P(a, b) = \frac{q_{\frac{1}{2}}(a)}{q_{\frac{1}{2}}(a) + q_{\frac{1}{2}}(b)}. \tag{1.2}$$

In this context, we take uniformity to mean that a team's likelihood of defeating another team is determined only by their winning percentages. For example, this assumption ignores the impact of the starting pitchers and precludes the situation where one team has a tendency to do particularly well or particularly poorly against another team.

This technique is sometimes called the log5 method of calculating $p(a, b)$, although we will avoid using this name as there is nothing obviously logarithmic about it. It is easy to see from (1.1) that

$$q_c(a) = \frac{ca}{1 - a}.$$

*Math. Mag.* **88** (2015) 54–71. doi:10.4169/math.mag.88.1.54. 
MSC: Primary 00A05.

Substituting this expression into (1.2), we see that

$$P(a,b) = \frac{a(1-b)}{a(1-b)+b(1-a)}, \tag{1.3}$$

not only for $c = \frac{1}{2}$ but for any positive $c$. The explicit form of $P(a,b)$ was first given by Dallas Adams [**8**], who also christened it the *James function*. It makes sense to extend the James function to values of $a$ and $b$ in the set $\{0, 1\}$, except when $a = b = 0$ or $a = b = 1$. In these two cases we would not have expected to be able to make predictions based on winning percentages alone. Moreover, both cases would be impossible if the two teams had previously competed against each other.

James's procedure can be interpreted as a twofold application of the general method known as the *Bradley–Terry model* (or sometimes the *Bradley–Terry–Luce model*). If $A$ and $B$ have worths $w(A)$ and $w(B)$, respectively, the probability that $A$ is considered superior to $B$ is

$$\pi(A, B) = \frac{w(A)}{w(A)+w(B)}.$$

Despite the attribution of this model to Bradley and Terry [**1**] and to Luce [**10**], the basic idea dates back to Zermelo [**15**]. The question, of course, is how to assign the "right" measure for the worth of $A$ in a particular setting. In chess, for instance, it is common to express the worth of a player as $10^{R_A/400}$, where $R_A$ denotes the player's Elo rating (see [**4**]). (The rating of chess players is the question in which Zermelo was originally interested. Good [**5**], who also considered this problem, seems to have been the first to call attention to Zermelo's paper.) Another example is James's so-called Pythagorean model (introduced in [**7**, p. 104] and discussed further in [**12**]) for estimating a team's seasonal winning percentage, based on the number $R$ of runs it scores and the number $S$ of runs it allows. In this case, the worth of the team is $R^2$ and the worth of its opposition is $S^2$.

In the construction of the James function, we can view the measure of a team's worth as being obtained from the Bradley–Terry model itself. We begin by assigning an arbitrary worth $c > 0$ (taken by James to be $\frac{1}{2}$) to a team with winning percentage $\frac{1}{2}$. Equation (1.1) can be construed as an application of the Bradley–Terry model, where the worth of a team is determined by the assumption that its overall winning percentage is equal to its probability of defeating a team with winning percentage $\frac{1}{2}$. Equation (1.2) represents a second application of the Bradley–Terry model, where each team has an arbitrary winning percentage and the measure of its worth comes from the previous application of the model.

This area of study, which is usually called the theory of paired comparisons, has focused from the outset on the question of inferring worth from an incomplete set of outcomes [**15**]. (See [**2**] for a thorough treatment, as well as [**3**] and [**14**] for additional context.) James, on the other hand, takes the worths to be known and uses them to determine the probability of the outcomes. We will adopt a similar point of view, emphasizing a set of axiomatic principles rather than a specific model.

James's justification [**8**] for his method does not invoke the Bradley–Terry model, but rather the fact that the resulting function $P(a,b)$ satisfies six self-evident conditions:

1. $P(a,a) = \frac{1}{2}$.
2. $P(a, \frac{1}{2}) = a$.
3. If $a > b$, then $P(a,b) > \frac{1}{2}$, and if $a < b$, then $P(a,b) < \frac{1}{2}$.
4. If $b < \frac{1}{2}$, then $P(a,b) > a$, and if $b > \frac{1}{2}$, then $P(a,b) < a$.

5. $0 \le P(a, b) \le 1$, and if $0 < a < 1$, then $P(a, 0) = 1$ and $P(a, 1) = 0$.
6. $P(a, b) + P(b, a) = 1$.

Condition (1) pertains to the situation where two different teams have the same winning percentage (as opposed to a single team competing against itself). To avoid contradicting (5), condition (4) should exclude the cases where $a = 0$ and $a = 1$. We will call this set, with this slight correction, the *proto-James conditions*. (James originally referred to them as "conditions of logic.") In addition to presenting some empirical evidence for (1.2), James makes the following assertion.

**The James Conjecture 1. (1981)** *The James function $P(a, b)$ is the only function that satisfies all six of the proto-James conditions.*

Jech [**9**] independently proposed a similar, albeit shorter list of conditions. Although he did not consider the James conjecture, he was able to prove a uniqueness theorem pertaining to a related class of functions.

The purpose of this paper is to examine the mathematical theory underlying the James function and to demonstrate that the James conjecture is actually false. In fact, we will introduce and study a large class of functions that satisfy the proto-James conditions.

While the proto-James conditions are certainly worthy of attention, we prefer to work with a slightly different set. The following conditions apply to all points $(a, b)$ with $0 \le a \le 1$ and $0 \le b \le 1$, except for $(0, 0)$ and $(1, 1)$:

(a) $P(a, \frac{1}{2}) = a$.
(b) $P(a, 0) = 1$ for $0 < a \le 1$.
(c) $P(b, a) = 1 - P(a, b)$.
(d) $P(1 - b, 1 - a) = P(a, b)$.
(e) $P(a, b)$ is a nondecreasing function of $a$ for $0 \le b \le 1$ and a strictly increasing function of $a$ for $0 < b < 1$.

We shall refer to conditions (a) to (e) as the *James conditions*. Condition (d), which is not represented among the proto-James conditions, simply states that the whole theory could be reformulated using losing percentages rather than winning percentages, with the roles of the two teams reversed. Together with condition (c), it is equivalent to saying $P(1 - a, 1 - b) = 1 - P(a, b)$, which may seem more natural to some readers. It should be clear from (1.3) that the James function satisfies James conditions (a) to (d). We will verify condition (e) in Section 3.

It is fairly obvious that the James conditions imply the proto-James conditions. Condition (a) is identical to condition (2). Condition (c) is condition (6), which implies (1) by taking $b = a$. Condition (e) is stronger than (3) and (4), and in concert with (1) and (2) implies them both. Combined with (c) or (d), it also implies that $P(a, b)$ is a nonincreasing function of $b$ for $0 \le a \le 1$ and a strictly decreasing function of $b$ for $0 < a < 1$. Finally, (b) implies the second of the three parts of (5). Together with (c), it also implies that $P(0, b) = 0$ if $0 < b \le 1$. By taking $b = 0$ in (d) and replacing $1 - a$ with $b$, condition (b) further implies that $P(1, b) = 1$ if $0 \le b < 1$, and this together with (c) gives $P(a, 1) = 0$ for $0 \le a < 1$, which is (a hair stronger than) the third part of (5). These facts, combined with (e), show that $0 < P(a, b) < 1$ when $0 < a < 1$ and $0 < b < 1$, which implies the first part of (5).

We will focus our attention on functions that satisfy the James conditions, and hence also the proto-James conditions. See [**6**], the online supplement to this paper, for an example of a function that satisfies the proto-James conditions but not the James conditions.

## 2 Verification of the James function

While the Bradley–Terry model is practically ubiquitous, its applicability to this situation is not obvious from an axiomatic perspective. We now present a self-contained proof that, under an intuitive probabilistic model in which $a$ and $b$ are the probabilities of success in simultaneous Bernoulli trials, the James function $P(a, b)$ represents the probability $p(a, b)$. This model satisfies the assumption of uniformity discussed in Section 1. The following argument was discovered by the third-named author several years ago [**11**], but has not previously appeared in a formal publication.

**Theorem 1.** *The probability $p(a, b)$ that a team with winning percentage $a$ defeats a team with winning percentage $b$ is given by the James function*

$$P(a, b) = \frac{a(1-b)}{a(1-b) + b(1-a)},$$

*except when $a = b = 0$ or $a = b = 1$, in which case $p(a, b)$ is undefined.*

*Proof.* Let teams $A$ and $B$ have winning percentages $a$ and $b$, respectively. Independently assign to each of $A$ and $B$ either a 0 or 1, where $A$ draws 1 with probability $a$ and $B$ draws 1 with probability $b$. If one team draws 1 and the other 0, the team with 1 wins the competition. If both teams draw the same number, repeat this procedure until they draw different numbers.

The probability that $A$ draws 1 and $B$ draws 0 on any given turn is clearly $a(1-b)$, while the opposite occurs with probability $b(1-a)$. The probability that $A$ and $B$ both draw 1 is $ab$, and the probability that they both draw 0 is $(1-a)(1-b)$. Hence

$$ab + (1-a)(1-b) + a(1-b) + b(1-a) = 1. \tag{2.1}$$

It follows that $0 \le ab + (1-a)(1-b) \le 1$ and $0 \le a(1-b) + b(1-a) \le 1$ whenever $0 \le a \le 1$ and $0 \le b \le 1$.

We can conclude the argument in either of two ways. Since the probability that $A$ and $B$ draw the same number is $ab + (1-a)(1-b)$, in which case they draw again, $p(a, b)$ must satisfy the functional equation

$$p(a, b) = a(1-b) + [ab + (1-a)(1-b)]\, p(a, b).$$

The only case in which we cannot solve for $p(a, b)$ is when $ab + (1-a)(1-b) = 1$. By (2.1), this situation only occurs when $a(1-b) + b(1-a) = 0$, which implies that either $a = b = 0$ or $a = b = 1$. Otherwise, $p(a, b) = P(a, b)$.

Alternatively, we may observe that the probability that $A$ wins on the $n$th trial is

$$a(1-b)\,[ab + (1-a)(1-b)]^{n-1},$$

and so the probability that $A$ wins in at most $n$ trials is

$$a(1-b) \sum_{k=1}^{n} [ab + (1-a)(1-b)]^{k-1}.$$

As $n$ tends to $\infty$, this expression yields a convergent geometric series unless $ab + (1-a)(1-b) = 1$. Using (2.1), we again obtain the James function. ∎

This proof relies on a particular model for the relationship between winning percentages and the outcome of a competition. Under different assumptions about this relationship, it seems possible that we would obtain other approximations for $p(a, b)$. Any such function would presumably also satisfy the James conditions.

## 3 Properties of the James function

In this section we will consider several important properties of the James function. We begin by computing the partial derivatives of $P(a, b)$, which will lead to an observation originally due to Dallas Adams. Note that

$$\frac{\partial P}{\partial a} = \frac{b(1-b)}{[a(1-b)+b(1-a)]^2}, \tag{3.1}$$

which shows that the James function satisfies condition (e), and also

$$\frac{\partial P}{\partial b} = \frac{-a(1-a)}{[a(1-b)+b(1-a)]^2}. \tag{3.2}$$

Furthermore, we have

$$\frac{\partial^2 P}{\partial a^2} = \frac{-2b(1-b)(1-2b)}{[a(1-b)+b(1-a)]^3},$$

so that, as a function of $a$, it follows that $P(a, b)$ is concave up for $\frac{1}{2} < b < 1$ and concave down for $0 < b < \frac{1}{2}$. Similarly,

$$\frac{\partial^2 P}{\partial b^2} = \frac{2a(1-a)(1-2a)}{[a(1-b)+b(1-a)]^3}.$$

Adams makes an interesting remark relating to the mixed second partial derivative

$$\frac{\partial^2 P}{\partial a \partial b} = \frac{a-b}{[a(1-b)+b(1-a)]^3}. \tag{3.3}$$

It follows from (3.3) that $\frac{\partial P}{\partial a}$, viewed as a function of $b$, is increasing for $b < a$ and decreasing for $b > a$, so it is maximized as a function of $b$ when $b = a$. Since $\frac{\partial P}{\partial a}$ is positive for every $0 < b < 1$, it must be most positive when $b = a$. Alternatively, (3.3) tells us that $\frac{\partial P}{\partial b}$, viewed as a function of $a$, is increasing for $a > b$ and decreasing for $a < b$, so it is minimized as a function of $a$ when $a = b$. Since $\frac{\partial P}{\partial b}$ is negative for every $0 < a < 1$, we conclude that it is most negative when $a = b$.

Adams interprets these facts in the following manner: since $P(a, b)$ increases most rapidly with $a$ when $b = a$ (and decreases most rapidly with $b$ when $a = b$), one should field one's strongest team when playing an opponent of equal strength [**8**]. Once again, this observation is perhaps more interesting in sports other than baseball, where the star players (other than pitchers) play nearly every game when healthy, although James points out that Yankees manager Casey Stengel preferred to save his ace pitcher, Whitey Ford, for the strongest opposition. It seems particularly relevant to European soccer, where the best teams engage in several different competitions at the same time against opponents of varying quality, and even the top players must occasionally be rested.

In principle, there are two ways to increase the value of $P(a, b)$: by increasing $a$ or by decreasing $b$. Under most circumstances, a team can only control its own quality and not that of its opponent. There are some situations, however, such as the Yankees signing a key player away from the Red Sox, where an individual or entity might exercise a degree of control over both teams. Similarly, there are many two-player games (such as Parcheesi and backgammon) in which each player's move affects the position of both players. In any such setting, it is a legitimate question whether the

priority of an individual or team should be to improve its own standing or to diminish that of its adversary.

Recall that the gradient of a function signifies the direction of the greatest rate of increase. The next result, which has apparently escaped notice until now, follows directly from equations (3.1) and (3.2).

**Proposition 2.** *For any point $(a, b)$, except where $a$ and $b$ both belong to the set $\{0, 1\}$, the gradient of the James function $P(a, b)$ is a positive multiple of the vector*

$$\langle b(1-b), -a(1-a)\rangle.$$

*In other words, to maximize the increase of $P(a, b)$, the optimal ratio of the increase of $a$ to the decrease of $b$ is $b(1-b) : a(1-a)$.*

One consequence of this result is that when two teams have identical winning percentages, the optimal strategy for increasing $P(a, b)$ is to increase $a$ and to decrease $b$ in equal measure. The same fact holds when two teams have complementary winning percentages. In all other situations, the maximal increase of $P(a, b)$ is achieved by increasing $a$ and decreasing $b$ by different amounts, with the ratio tilted towards the team whose winning percentage is further away from $\frac{1}{2}$. In the extremal cases, when one of the two values $a$ or $b$ belongs to the set $\{0, 1\}$, the optimal strategy is to devote all resources to changing the winning percentage of the team that is either perfectly good or perfectly bad. This observation is somewhat vacuous when $a = 1$ or $b = 0$, since $P(a, b)$ is already as large as it could possibly be, although the strategy is entirely reasonable when $a = 0$ or $b = 1$. It also makes sense that the gradient is undefined at the points $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$, since these winning percentages do not provide enough information to determine how much one team must improve to defeat the other.

If $P(a, b) = c$, it is easy to see that $a(1-b)(1-c) = (1-a)bc$, which implies the next result.

**Proposition 3.** *If $0 < a < 1$, then $P(a, b) = c$ if and only if $P(a, c) = b$. In other words, for a fixed value of $a$, the James function is an involution.*

The practical interpretation of this result is simple to state, even if it is not intuitively obvious: if team $A$ has probability $c$ of beating a team with winning percentage $b$, then team $A$ has probability $b$ of beating a team with winning percentage $c$. The James conditions already imply this relationship whenever $b$ and $c$ both belong to the set $\{0, 1\}$ or the set $\{\frac{1}{2}, a\}$. Nevertheless, it is not evident at this point whether the involutive property is a necessary consequence of the James conditions. (Example 6 will provide an answer to this question.)

Proposition 3 has two further implications that are worth mentioning. The first is a version of the involutive property that holds for a fixed value of $b$:

If $0 < b < 1$, then $P(a, b) = 1 - c$ if and only if $P(c, b) = 1 - a$.

The second is that the level curves for the James function (that is, the set of all points for which $P(a, b) = c$ for a particular constant $c$) can be written

$$b = P(a, c) = \frac{a(1-c)}{a(1-c) + c(1-a)} \tag{3.4}$$

for $0 < a < 1$. These level curves are the concrete manifestation of a straightforward principle: if a team $A$ improves by a certain amount, there should be a corresponding amount that a team $B$ can improve so that the probability of $A$ defeating $B$ remains

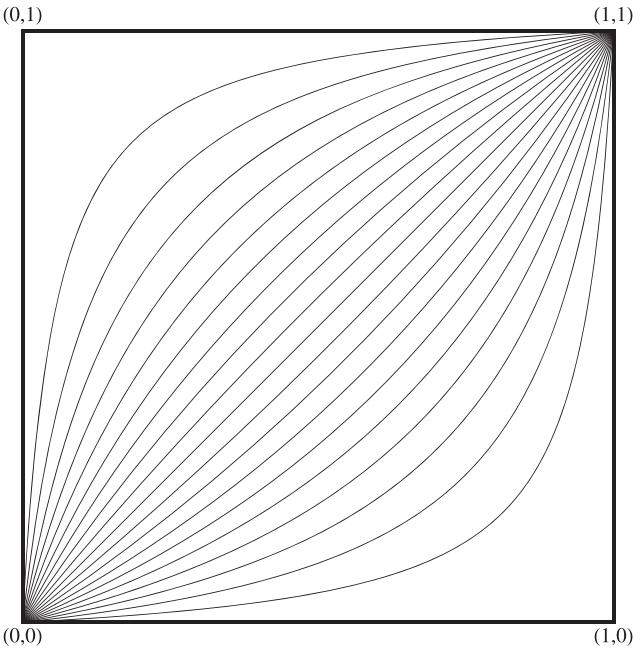



**Figure 1** The level curves for the James function *P(a, b)*

unchanged. Each level curve represents the path from $(0, 0)$ to $(1, 1)$ that such a pair would take in tandem. (See FIGURE 1.)

We conclude this section with one more observation relating to these level curves.

**Proposition 4.** *For any $0 < c < 1$, the corresponding level curve for the James function $P(a, b)$ is the unique solution to the differential equation*

$$\frac{db}{da} = \frac{b(1-b)}{a(1-a)}$$

*that passes through the point $(c, \frac{1}{2})$.*

Another way of stating this result is that, for two teams to maintain the same value of $P(a, b)$, they should increase (or decrease) their winning percentages according to the ratio $a(1-a) : b(1-b)$. One can either verify this assertion directly, by solving the differential equation to obtain (3.4), or by appealing to Proposition 2 and recalling that the gradient is always perpendicular to the level curve at a particular point.

## 4 Jamesian functions

We will now consider the question of whether there is a unique function satisfying the James conditions. We begin with the following observation, which is implicit in the construction of the James function.

**Proposition 5.** *The James function is the only function derived from the Bradley–Terry model that satisfies the James conditions.*

*Proof.* Suppose $\pi(A, B)$ satisfies the James conditions and is derived from the Bradley–Terry model. Let team $A$ have winning percentage $a$, with $0 < a < 1$, and

let team $C$ have winning percentage $\frac{1}{2}$. Condition (a) implies that

$$a = \pi(A, C) = \frac{w(A)}{w(A) + w(C)}.$$

Solving for $w(A)$, we obtain

$$w(A) = \frac{aw(C)}{1 - a} = q_c(a),$$

where $c = w(C)$. Thus $\pi(A, B)$ agrees with the James function $P(a, b)$ when both $a$ and $b$ belong to the interval $(0, 1)$. Since the James conditions uniquely determine the value of a function whenever $a$ or $b$ belongs to $\{0, 1\}$, the functions $\pi(A, B)$ and $P(a, b)$ must be identical. ■

Let $S$ denote the open unit square $(0, 1) \times (0, 1)$. We will say that any function $J(a, b)$, defined on the set $\overline{S} \setminus \{(0, 0) \cup (1, 1)\}$, that satisfies the James conditions is a *Jamesian function*. Our immediate objective is to disprove the James conjecture by identifying at least one example of a Jamesian function that is different from the James function $P(a, b)$. Proposition 5 guarantees that any such function, if it exists, cannot be derived from the Bradley–Terry model.

**Example 6.** We will reverse engineer our first example of a new Jamesian function by starting with its level curves. Consider the family of curves $\{j_c\}_{c\in(0,1)}$ defined as follows:

$$j_c(a) = \begin{cases} \dfrac{a}{2c}, & 0 < a \le \dfrac{2c}{1+2c} \\ 2ca + 1 - 2c, & \dfrac{2c}{1+2c} < a < 1 \end{cases}$$

for $0 < c \le \frac{1}{2}$ and

$$j_c(a) = \begin{cases} (2 - 2c)a, & 0 < a \le \dfrac{1}{3-2c} \\ \dfrac{a + 1 - 2c}{2 - 2c}, & \dfrac{1}{3-2c} < a < 1 \end{cases}$$

for $\frac{1}{2} < c < 1$. (See FIGURE 2.) These curves have been chosen to satisfy certain symmetry properties, which the reader can probably deduce but which we will not state explicitly. (Suffice it to say that $j_c(c) = \frac{1}{2}$ for all $c$.) We define the function $J(a, b)$ on $S$ by assigning to every point $(a, b)$ the value of $c$ associated with the particular curve $j_c$ that passes through that point. We assign the value 0 or 1 to points on the boundary of $S$, as dictated by the James conditions.

A bit more work yields an explicit formula for $J(a, b)$, from which one can verify directly that all of the James conditions are satisfied:

$$J(a, b) = \begin{cases} \dfrac{a}{2b}, & (a, b) \in \text{I} \\ \dfrac{2a - b}{2a}, & (a, b) \in \text{II} \\ \dfrac{1 - b}{2(1 - a)}, & (a, b) \in \text{III} \\ \dfrac{1 + a - 2b}{2(1 - b)}, & (a, b) \in \text{IV} \end{cases},$$

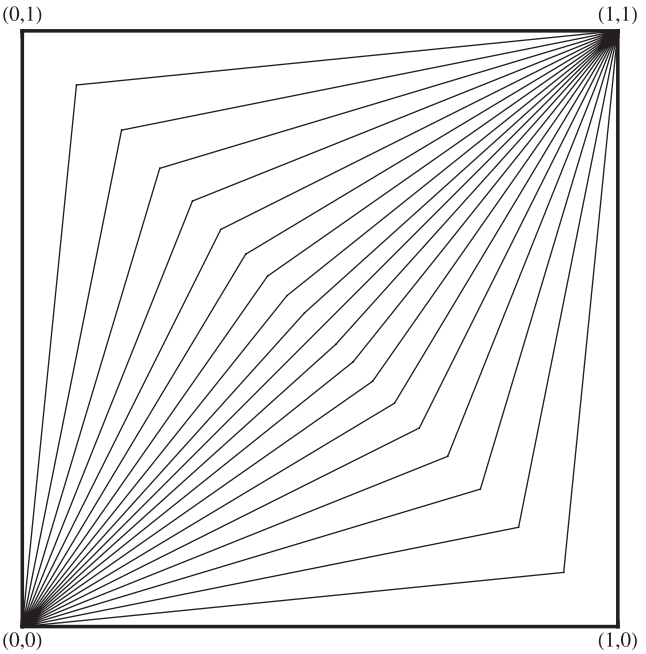



**Figure 2** The level curves for the function *J(a, b)* in Example 6

where I, II, III, and IV are subsets of $\overline{S} \setminus \{(0,0) \cup (1,1)\}$ that are defined according to FIGURE 3.

Observe that the appropriate definitions coincide on the boundaries between regions, from which it follows that $J(a,b)$ is continuous on $\overline{S} \setminus \{(0,0) \cup (1,1)\}$. On the other hand, it is not difficult to see that $J(a,b)$ fails to be differentiable at all points of the form $(a, 1-a)$ for $0 < a < \frac{1}{2}$ or $\frac{1}{2} < a < 1$. (With some effort, one can show that it is differentiable at the point $(\frac{1}{2}, \frac{1}{2})$.) In reference to Proposition 3, note that $J(\frac{1}{3}, \frac{1}{4}) = \frac{5}{8}$ and $J(\frac{1}{3}, \frac{5}{8}) = \frac{4}{15}$. In other words, the involutive property is not a necessary consequence of the James conditions.

In view of the preceding example, we need to refine our terminology somewhat. We will refer to any Jamesian function (such as the James function itself) that satisfies the condition

$$J\big(a, J(a,b)\big) = b$$

for $0 < a < 1$ as an *involutive Jamesian function*.

It turns out to be fairly easy to construct Jamesian functions with discontinuities in $S$ (see [**6**]). Proposition 8, which we will prove in the next section, guarantees that any such function is not involutive. Rather than considering such pathological examples, we will devote the next section to examining Jamesian functions that are involutive, continuous, and (in many cases) differentiable.

## 5 Involutive Jamesian functions

We now turn our attention to Jamesian functions that satisfy the involutive property

$$J\big(a, J(a,b)\big) = b,$$

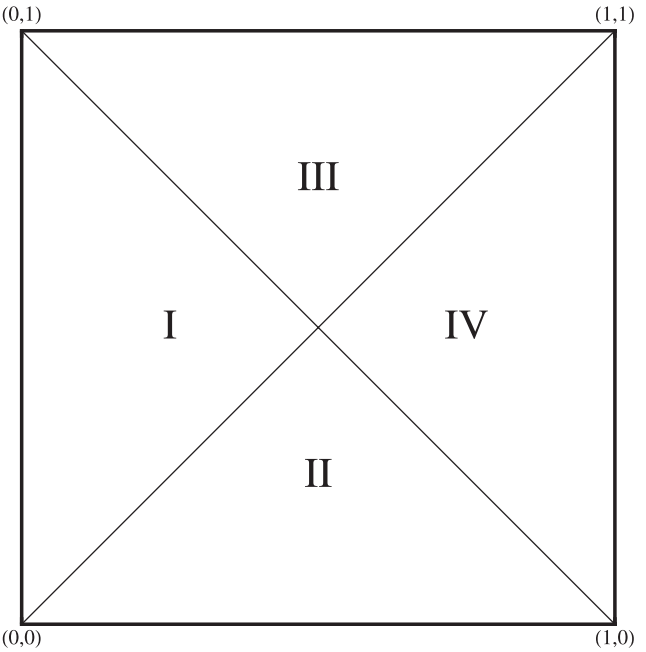



**Figure 3** The subsets of $\overline{S} \setminus \{(0,0) \cup (1,1)\}$ in Example 6

or equivalently

$$J(a,b) = c \text{ if and only if } J(a,c) = b,$$

whenever $0 < a < 1$. This property essentially subsumes three of the five James conditions (namely (a), (b), and (d)).

**Proposition 7.** *A function $J : \overline{S} \setminus \{(0,0) \cup (1,1)\} \to [0,1]$ is an involutive Jamesian function if and only if it satisfies the involutive property, James condition (c), and James condition (e).*

*Proof.* By definition, an involutive Jamesian function must satisfy the involutive property, as well as all five James conditions. Suppose then that $J(a,b)$ satisfies the involutive property, together with conditions (c) and (e).

To see that $J(a,b)$ satisfies condition (b), take $0 < a < 1$ and suppose that $J(a,0) = c$ for $0 \le c < 1$. The involutive property would then dictate that $J(a,c) = 0$, and thus condition (c) would imply that $J(c,a) = 1$. Hence $J(c',a) \le J(c,a)$ for $c < c' \le 1$, which would violate condition (e). Consequently $J(a,0) = 1$ for $0 < a < 1$. Since $J(a,b)$ is a nondecreasing function of $a$, we conclude that $J(1,0) = 1$ as well.

Next consider condition (d). Applying the involutive property three times and condition (c) twice, we see that

$$\begin{aligned} J(a,b) = c &\iff J(a,c) = b \\ &\iff J(c,a) = 1 - b \\ &\iff J(c,1-b) = a \\ &\iff J(1-b,c) = 1 - a \\ &\iff J(1-b,1-a) = c, \end{aligned}$$

as long as $a$, $b$, and $c$ all belong to the interval $(0, 1)$. The cases where $a$, $b$, or $c$ belongs to $\{0, 1\}$ can be dealt with by appealing to condition (b). In particular, we know that $J(a, 0) = 1$ for $0 < a \leq 1$, which implies that $J(1 - a, 0) = 1$ for $0 \leq a < 1$. The involutive property dictates that $J(1 - a, 1) = 0$ for $0 < a < 1$. Since $J(1, 0) = 1$, it follows from (c) that $J(1, 1 - a) = 1 = J(a, 0)$ for $0 < a \leq 1$. Hence condition (d) holds whenever $b = 0$. The remaining cases can be deduced from this observation.

Finally, consider condition (a). Taking $b = a$ in condition (c), we see that $J(a, a) = \frac{1}{2}$. Hence the involutive property dictates that $J(a, \frac{1}{2}) = a$ for $0 < a < 1$. For $a = 1$, simply note that conditions (d) and (b) imply that $J(1, \frac{1}{2}) = J(\frac{1}{2}, 0) = 1$. Similarly, condition (c) shows that $J(0, \frac{1}{2}) = 1 - J(\frac{1}{2}, 0) = 0$. ■

In other words, to identify an involutive Jamesian function, we can restrict our attention to the following set of conditions:

(i) $J\big(a, J(a, b)\big) = b$ for $0 < a < 1$.

(ii) $J(b, a) = 1 - J(a, b)$.

(iii) $J(a, b)$ is a nondecreasing function of $a$ for $0 \leq b \leq 1$ and a strictly increasing function of $a$ for $0 < b < 1$.

We will refer to this list as the *involutive James conditions*.

Condition (i) also guarantees that a Jamesian function possesses another important property.

**Proposition 8.** *Every involutive Jamesian function is continuous on* $\overline{S} \setminus \{(0, 0) \cup (1, 1)\}$.

*Proof.* Take a fixed value $0 < c < 1$ and consider the level curve $J(a, b) = c$, which can be rewritten $b = J(a, c)$ for $0 < a < 1$. Conditions (i) and (ii) imply that

$$J\big(1 - J(a, c), c\big) = 1 - a.$$

Thus $J(a, c)$, viewed as a function of $a$, is a bijection from the interval $(0, 1)$ onto itself. Hence it follows from (iii) that the curve $J(a, c)$ is a continuous, strictly increasing function of $a$ that connects the points $(0, 0)$ and $(1, 1)$.

Suppose, for the sake of contradiction, that $J(a, b)$ fails to be continuous at a point $(a_0, b_0)$ in $S$. In other words, there exists a positive number $\varepsilon_0$ such that, for any positive $\delta$, there is a point $(a, b)$ such that $\|(a, b) - (a_0, b_0)\| < \delta$ and $|J(a, b) - J(a_0, b_0)| \geq \varepsilon_0$. (If necessary, redefine $\varepsilon_0$ so it is less than $\min\{2J(a_0, b_0), 2 - 2J(a_0, b_0)\}$.) Let $c_1 = J(a_0, b_0) - \varepsilon_0/2$ and $c_2 = J(a_0, b_0) + \varepsilon_0/2$, and consider the level curves $J(a, c_1)$ and $J(a, c_2)$. Let $\delta_0$ denote the minimum of the distance between $(a_0, b_0)$ and $J(a, c_1)$ and the distance between $(a_0, b_0)$ and $J(a, c_2)$.

By assumption, there is a point $(a_3, b_3)$ such that $\|(a_3, b_3) - (a_0, b_0)\| < \delta_0$ and $c_3 = J(a_3, b_3)$ is either less than or equal to $J(a_0, b_0) - \varepsilon_0$ or greater than or equal to $J(a_0, b_0) + \varepsilon_0$. Since $J(a, c_i) = \frac{1}{2}$ at $a = c_i$, the level curve $J(a, c_3)$ intersects the line $b = \frac{1}{2}$ either to the left of the curve $J(a, c_1)$ or to the right of the curve $J(a, c_2)$. On the other hand, since $(a_3, b_3)$ lies within $\delta_0$ of $(a_0, b_0)$, the curve $J(a, c_3)$ must intersect the line $b = b_3$ between $J(a, c_1)$ and $J(a, c_2)$. Hence two of the level curves must intersect at a point in $S$, which is impossible. (See FIGURE 4 for a graphical illustration of this argument.)

Now consider a point $(a_0, b_0)$ on the boundary of $S$. The only difference in the proof is that, if $a = 0$ or $b = 1$, the level curve $J(a, c_1)$ does not exist. In this case, it is not difficult to see that $J(a, c_3)$ must intersect the curve $J(a, c_2)$. Similarly, if $a = 1$ or $b = 0$, there is no level curve $J(a, c_2)$, but one can show that $J(a, c_3)$ must intersect $J(a, c_1)$. ■

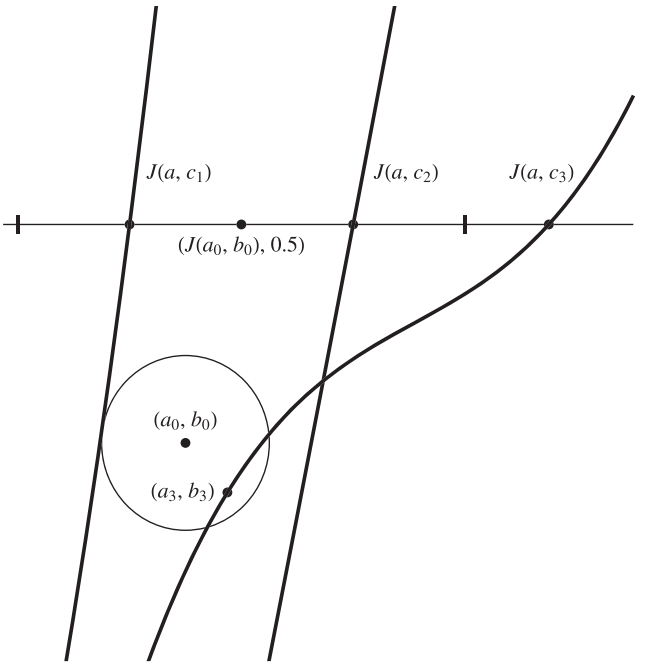



**Figure 4** An illustration of the proof of Proposition 8

Let $g\colon (0, 1) \to \mathbb{R}$ be a continuous, strictly increasing function that satisfies the conditions

- $g(1 - a) = -g(a)$.
- $\lim_{a \to 0^+} g(a) = -\infty$.

These conditions imply that $g(\frac{1}{2}) = 0$ and that

$$\lim_{a \to 1^-} g(a) = \infty.$$

Observe that $g^{-1}\colon \mathbb{R} \to (0, 1)$ is a continuous, strictly increasing function with $g^{-1}(-s) = 1 - g^{-1}(s)$. It makes sense to define $g(0) = -\infty$ and $g(1) = \infty$, so that $g^{-1}(-\infty) = 0$ and $g^{-1}(\infty) = 1$. We claim that any such function $g$ can be used to construct an involutive Jamesian function.

**Theorem 9.** *For any $g$ satisfying the conditions specified above, the function*

$$J(a, b) = g^{-1}\big(g(a) - g(b)\big) \tag{5.1}$$

*is an involutive Jamesian function.*

*Proof.* Consider each of the three involutive James conditions:
(i) Note that

$$\begin{aligned} J\big(a, J(a, b)\big) &= g^{-1}\big(g(a) - g\big(g^{-1}\big(g(a) - g(b)\big)\big)\big) \\ &= g^{-1}\big(g(a) - g(a) + g(b)\big) \\ &= g^{-1}\big(g(b)\big) = b, \end{aligned}$$

as long as $0 < a < 1$. (The cases where $a = 0$ and $a = 1$ yield the indeterminate forms $-\infty + \infty$ and $\infty - \infty$.)

(ii) Similarly,

$$J(b, a) = g^{-1}\big(g(b) - g(a)\big) = 1 - g^{-1}\big(g(a) - g(b)\big) = 1 - J(a, b).$$

(iii) Since both $g$ and $g^{-1}$ are strictly increasing, it follows that $J(a, b)$ is a strictly increasing function of $a$ when $0 < b < 1$. Moreover, $J(a, b)$ takes on the constant value 1 when $b = 0$ and the constant value 0 when $b = 1$. ■

While it is unnecessary to verify James conditions (a) and (d), it is worth noting that (a) corresponds to the property $g(\frac{1}{2}) = 0$ and (d) to the property $g(1 - a) = -g(a)$. In effect, we verified condition (b) in the process of considering (iii).

It is easy to use Theorem 9 to generate concrete examples.

**Example 10.** The function

$$g(a) = \frac{2a - 1}{a(1 - a)}$$

satisfies all the necessary conditions for Theorem 9, so (5.1) defines an involutive Jamesian function. Since

$$g^{-1}(s) = \frac{s - 2 + \sqrt{s^2 + 4}}{2s},$$

we obtain

$$J(a, b) = \frac{x + y - \sqrt{x^2 + y^2}}{2y} = \frac{x}{x + y + \sqrt{x^2 + y^2}},$$

where $x = 2ab(1 - a)(1 - b)$ and $y = (b - a)(2ab - a - b + 1)$.

**Example 11.** The function $g(a) = -\cot(\pi a)$ yields the involutive Jamesian function

$$J(a, b) = \frac{1}{\pi} \cot^{-1}\big(\cot(\pi a) - \cot(\pi b)\big),$$

where we are using the version of the inverse cotangent that attains values between 0 and $\pi$.

The construction described in Theorem 9 is closely related to what is known as a *linear model* for paired comparisons. In such a model,

$$\pi(A, B) = F\big(v(A) - v(B)\big),$$

where $v$ denotes a measure of worth and $F$ is the cumulative distribution function of a random variable that is symmetrically distributed about 0 (see [**2**, Section 1.3]). The Bradley–Terry model can be viewed as a linear model, where $F$ is the logistic function

$$F(s) = \frac{e^s}{e^s + 1} = \int_{-\infty}^{s} \frac{e^t}{(1 + e^t)^2} dt$$

and $v(A) = \log w(A)$. In particular, the James function can be constructed in the manner of Theorem 9, with $F = g^{-1}$ being the logistic function and $g$ being the so-called logit function

$$g(a) = \log\left(\frac{a}{1 - a}\right).$$

(This observation could charitably be construed as an *a posteriori* justification for the term "log5" originally used by James.)

What is distinctive about the James function in this context is that the construction is symmetric, with $v(A) = \log w(A)$ and $v(B) = \log w(B)$ replaced by $g(a) = \log(a/(1-a))$ and $g(b) = \log(b/(1-b))$, respectively. This symmetry corresponds to the twofold application of the Bradley–Terry model that was discussed in Section 1. Likewise, the fact that both $g$ and $g^{-1}$ appear in the general formulation of Theorem 9 can be interpreted as a consequence of the same model being used to define both worth and probability.

**Example 12.** Take

$$F(s) = g^{-1}(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{-\frac{t^2}{2}} dt,$$

so that $g$ is the so-called probit function. The involutive Jamesian function $J(a, b) = g^{-1}\big(g(a) - g(b)\big)$ can be considered the analogue of the James function relative to the Thurstone–Mosteller model (see [**2**]).

Theorem 9 allows us to identify a large class of functions that can be viewed as generalizations of the James function. Since

$$\log\left(\frac{a}{1-a}\right) = \int_{\frac{1}{2}}^{a} \left(\frac{1}{t} + \frac{1}{1-t}\right) dt = \int_{\frac{1}{2}}^{a} \frac{1}{t(1-t)} dt,$$

we define

$$g_n(a) = \int_{\frac{1}{2}}^{a} \frac{1}{(t(1-t))^n} dt$$

for any real number $n \geq 1$. It is not difficult to verify that $g_n$ satisfies all of the prescribed requirements for Theorem 9. (The stipulation that $g_n(0) = -\infty$ precludes the case where $0 < n < 1$.) Define

$$H_n(a, b) = g_n^{-1}\big(g_n(a) - g_n(b)\big). \tag{5.2}$$

For $n > 1$, we shall refer to $H_n(a, b)$ as a *hyper-James function*. Each of these functions is an involutive Jamesian function.

In some situations, it is possible to obtain a more concrete representation for $H_n(a, b)$. For example, one can show that

$$g_{\frac{3}{2}}(a) = \frac{2(2a-1)}{\sqrt{a(1-a)}}$$

and

$$g_{\frac{3}{2}}^{-1}(s) = \frac{s + \sqrt{s^2+16}}{2\sqrt{s^2+16}},$$

and hence

$$H_{\frac{3}{2}}(a, b) = \frac{1}{2} + \frac{v'\sqrt{u} - u'\sqrt{v}}{2\sqrt{u + v - 4uv - 2u'v'\sqrt{uv}}}$$

for $u = a(1-a)$, $v = b(1-b)$, $u' = 1-2a$, and $v' = 1-2b$ (see [**6**] for more details). In general, though, it seems unlikely that there is an explicit formula for $H_n(a, b)$ that is more useful than (5.2).

We will now examine the issue of differentiability. For any function defined according to Theorem 9, a routine calculation shows that

$$\frac{\partial J}{\partial a} = \frac{g'(a)}{g'\big(J(a,b)\big)} \tag{5.3}$$

and

$$\frac{\partial J}{\partial b} = \frac{-g'(b)}{g'\big(J(a,b)\big)} \tag{5.4}$$

at all points $(a, b)$ for which the above quotients are defined. Based on this observation, we are able to obtain the following result.

**Proposition 13.** *If $g$ is continuously differentiable on $(0, 1)$, with $g'$ never equal to 0, the corresponding Jamesian function $J(a, b)$ is differentiable on $S$. Conversely, if $J(a, b)$ is differentiable on $S$, the function $g$ must be differentiable on $(0, 1)$ with $g'$ never 0.*

*Proof.* Suppose that $g'$ is continuous and nonzero on $(0, 1)$. It follows from (5.3) and (5.4) that both $\frac{\partial J}{\partial a}$ and $\frac{\partial J}{\partial b}$ are defined and continuous at all points in the open set $S$, which guarantees that $J(a, b)$ is differentiable on $S$.

Now suppose that $J(a, b)$ is differentiable at every point in $S$. Let $a_0$ be an arbitrary element of $(0, 1)$. Since $g$ is strictly increasing, it could only fail to be differentiable on a set of measure 0 (see [**13**, p. 112]). In particular, there is at least one $c$ in $(0, 1)$ for which $g'(c)$ is defined. Since $J(a_0, b)$, viewed as a function of $b$, attains every value in the interval $(0, 1)$, there exists a $b_0$ in $(0, 1)$ such that $J(a_0, b_0) = c$. Note that

$$g(a) = g\big(J(a, b_0)\big) + g(b_0)$$

for all $a$ in $(0, 1)$, so the chain rule dictates that

$$g'(a_0) = g'(c) \cdot \frac{\partial J}{\partial a}(a_0, b_0).$$

Therefore, $g$ is differentiable on the entire interval $(0, 1)$. Suppose, for the sake of contradiction, that there were some $d$ in $(0, 1)$ for which $g'(d) = 0$. As before, there would exist a $b_1$ in $(0, 1)$ such that $J(a_0, b_1) = d$, which would imply that

$$g'(a_0) = g'(d) \cdot \frac{\partial J}{\partial a}(a_0, b_1) = 0.$$

Consequently, $g'$ would be identically 0 on $(0, 1)$, which is impossible. ∎

In other words, all the specific examples of Jamesian functions we have introduced in this section, including the hyper-James functions, are differentiable on $S$. We can now state a more general version of Proposition 2, which follows directly from (5.3) and (5.4).

**Proposition 14.** *For any differentiable Jamesian function $J(a, b)$ defined according to Theorem 9, the gradient at a point $(a, b)$ in $S$ is a positive multiple of the vector $\langle g'(a), -g'(b)\rangle$.*

If $g$ is differentiable on $(0, 1)$, the condition that $g(1 - a) = -g(a)$ implies that $g'(1 - a) = g'(a)$. Hence the gradient of $J(a, b)$ is a positive multiple of $\langle 1, -1\rangle$ whenever $b = a$ or $b = 1 - a$. This observation generalizes the fact that, whenever two teams have identical or complementary winning percentages, the optimal strategy for increasing $P(a, b)$ is to increase $a$ and decrease $b$ by equal amounts.

For any Jamesian function given by (5.1), the level curve $J(a, b) = c$ for $0 < c < 1$ can be rewritten

$$b = J(a, c) = g^{-1}\big(g(a) - g(c)\big),$$

or $g(a) = g(b) + g(c)$. Hence we have the following generalization of Proposition 4.

**Proposition 15.** *Let $J(a, b)$ be a differentiable Jamesian function defined according to Theorem 9. For any $0 < c < 1$, the corresponding level curve for $J(a, b)$ is the unique solution to the differential equation*

$$\frac{db}{da} = \frac{g'(a)}{g'(b)}$$

*that passes through the point $(c, \frac{1}{2})$.*

Thus the level curves for the Jamesian functions defined in Examples 10 and 11 are given by the differential equations

$$\frac{db}{da} = \frac{(2a^2 - 2a + 1)(b(1-b))^2}{(2b^2 - 2b + 1)(a(1-a))^2}$$

and

$$\frac{db}{da} = \left(\frac{\sin(\pi b)}{\sin(\pi a)}\right)^2,$$

respectively. Likewise, the level curves for any hyper-James function $H_n(a, b)$ are given by the differential equation

$$\frac{db}{da} = \left(\frac{b(1-b)}{a(1-a)}\right)^n.$$

FIGURE 5 shows the level curves for the hyper-James function $H_2(a, b)$.

## 6 Final thoughts

While it is possible to construct additional examples of noninvolutive Jamesian functions, it would be reasonable to focus any further investigation on the involutive case. Perhaps the most obvious question is whether one can assign any probabilistic significance to the involutive Jamesian functions we have just introduced, particularly the hyper-James functions. For instance, could one somehow alter the assumptions underlying Theorem 1 to obtain one of these functions in place of $P(a, b)$?

Within this context, several lines of inquiry seem especially worthwhile:

1. Does every involutive Jamesian function have the form described in Theorem 9, for some particular function $g$?
2. While it is clear how the involutive property arises mathematically, is there any *a priori* reason that it should hold, based on the probabilistic interpretation of the James function?
3. Are there any situations for which nondifferentiability would make sense in the setting of an athletic competition?

We would be delighted if this paper motivated other mathematicians (or sports enthusiasts) to consider any of these questions.

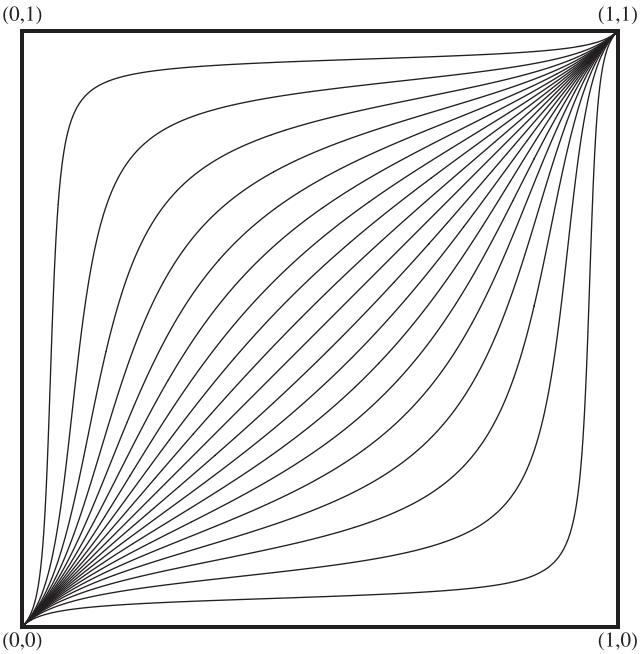



**Figure 5** The level curves for the hyper-James function $H_2(a, b)$

**Acknowledgments** We would never have written this paper if Caleb Garza, a recent alumnus of Connecticut College, had not decided to give a senior seminar talk on a topic from sabermetrics. We are sincerely grateful to him for prompting (or reviving) our interest in this material and for bringing the work of the third-named author to the attention of the first two. We would also like to thank the referees and editors of this paper for providing substantial assistance and guidance. The third-named author was partially supported by NSF grant DMS1265673.

## REFERENCES

1. R. A. Bradley, M. E. Terry, Rank analysis of incomplete block designs. I. The method of paired comparisons, *Biometrika* **39** (1952) 324–345, http://dx.doi.org/10.1093/biomet/39.3-4.324.
2. H. A. David, *The Method of Paired Comparisons*. Second edition. Oxford University Press, New York, 1988.
3. M. E. Glickman, Introductory note to 1928 (=1929), in *Ernst Zermelo: Collected Works*. Volume II, Edited by Heinz-Dieter Ebbinghaus and Akihiro Kanamori. Springer-Verlag, Berlin, 2013. 616–621, http://dx.doi.org/10.1007/978-3-540-70856-8_13.
4. M. E. Glickman, A. C. Jones, Rating the chess rating system, *Chance* **12** no. 2 (1999) 21–28.
5. I. J. Good, On the marking of chess-players, *Math. Gaz.* **39** (1955) 292–296.
6. C. N. B. Hammond, W. P. Johnson, S. J. Miller, Online supplement to "The James function," www.maa.org/mathmag/supplements.
7. B. James, *1980 Baseball Abstract*. self-published, Lawrence, KS, 1980.
8. B. James, *1981 Baseball Abstract*. self-published, Lawrence, KS, 1981.
9. T. Jech, A quantitative theory of preferences: some results on transition functions, *Soc. Choice Welf.* **6** no. 4 (1989) 301–314, http://dx.doi.org/10.1007/bf00446987.
10. R. D. Luce, *Individual Choice Behavior: A Theoretical Analysis*. John Wiley and Sons, New York, 1959.
11. S. J. Miller, A justification of the log5 rule for winning percentages, http://web.williams.edu/Mathematics/sjmiller/public_html/103/Log5WonLoss_Paper.pdf, 2008.
12. S. J. Miller, T. Corcoran, J. R. Gossels, V. Luo, J. Porfilio, Pythagoras at the bat, in *Social Networks and the Economics of Sports*. Edited by Panos M. Pardalos and Victor Zamaraev. Springer-Verlag, Berlin, 2014. 89–113, http://dx.doi.org/10.1007/978-3-319-08440-4_6.
13. H. L. Royden, P. M. Fitzpatrick, *Real Analysis*. Fourth edition. Prentice Hall, Boston, 2010.
14. M. Stob, A supplement to "A mathematician's guide to popular sports," *Amer. Math. Monthly* **91** (1984), no. 5, 277–282.
15. E. Zermelo, Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung, *Math. Z.* **29** no. 1 (1929) 436–460. The calculation of the results of a tournament as a maximum prob-

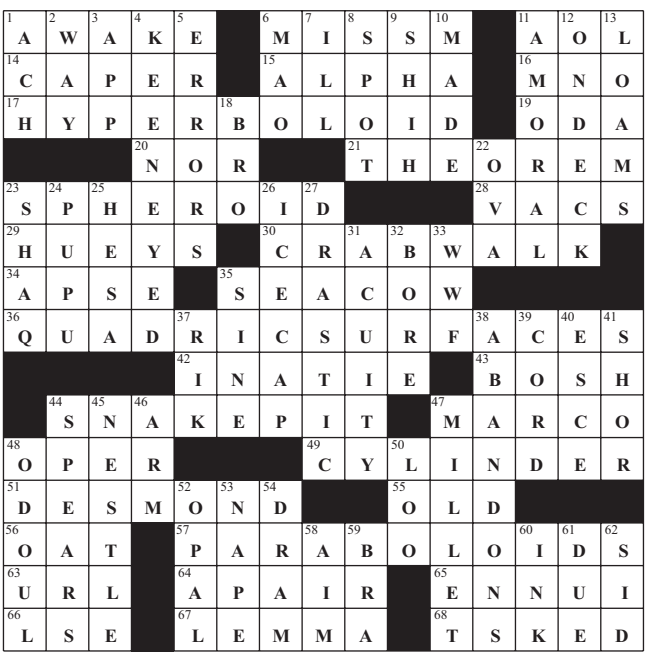
lem in the calculus of probabilities, in *Ernst Zermelo: Collected Works*. Volume II. Edited by Heinz-Dieter Ebbinghaus and Akihiro Kanamori. Springer-Verlag, Berlin, 2013. 622–671.

**Summary.** We investigate the properties of the James function, associated with Bill James's so-called "log5 method," which assigns a probability to the result of a game between two teams based on their respective winning percentages. We also introduce and study a class of functions, which we call *Jamesian*, that satisfy the same *a priori* conditions that were originally used to describe the James function.

**CHRISTOPHER HAMMOND** (MR Author ID: 728945) grew up in Durham, North Carolina, where he had the privilege of watching Duke basketball in its heyday. After graduating from the North Carolina School of Science and Mathematics, he earned his bachelor's degree at the University of the South in Sewanee, Tennessee. He received his doctorate at the University of Virginia, after which he had the good fortune to obtain his current position at Connecticut College. This paper is his first publication without the Mathematics Subject Classification 47B33.

**WARREN P. JOHNSON** (MR Author ID: 612409) grew up rooting for the Phillies, and became a Twins fan while attending the University of Minnesota. He received his Ph.D. from the University of Wisconsin despite failing to become a Brewers fan. This paper marks the first time that such knowledge of baseball as he possesses has been of any use in his career. He is an Associate Editor of Mathematics Magazine.

**STEVEN J. MILLER** (MR Author ID: 735278) received his bachelor's degree at Yale and his doctorate at Princeton. After teaching at Princeton, New York University, (The) Ohio State University, and Brown, he joined the faculty at Williams College in 2008. He is a lifelong Red Sox fan. As his great uncle promised him many years ago, he has lived long enough to see the Sox win it all. Not to be outdone, his seven-year-old son Cameron remarked in 2013 that he did it a lot earlier!



# PROBLEMS

BERNARDO M. ÁBREGO, *Editor*
California State University, Northridge

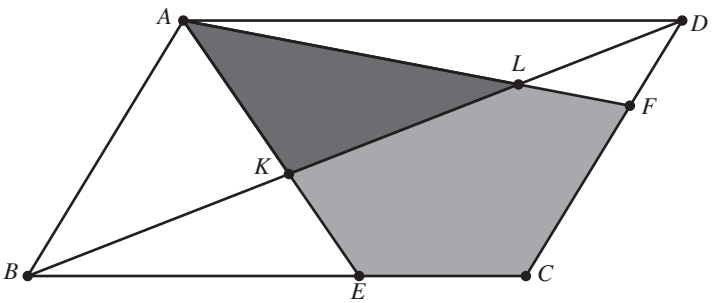
*Assistant Editors:* SILVIA FERNÁNDEZ-MERCHANT, California State University, Northridge; JOSÉ A. GÓMEZ, Facultad de Ciencias, UNAM, México; EUGEN J. IONASCU, Columbus State University; ROGELIO VALDEZ, Facultad de Ciencias, UAEM, México; WILLIAM WATKINS, California State University, Northridge

## PROPOSALS

*To be considered for publication, solutions should be received by July 1, 2015.*

**1961.** *Proposed by George Apostolopoulos, Messolonghi, Greece.*

Let $ABCD$ be a parallelogram and $E$ and $F$ points on the sides $\overline{BC}$ and $\overline{CD}$, respectively, such that $BE \cdot FD = EC \cdot CF$. The segments $\overline{AE}$ and $\overline{AF}$ meet the diagonal $\overline{BD}$ at the points $K$ and $L$, respectively. Prove that $\text{Area}(KECFL) \leq 2 \cdot \text{Area}(AKL)$.




**1962.** *Proposed by Timothy Hall, PQI Consulting, Cambridge, MA.*

Evaluate

$$\lim_{r\to 0^+} r^2 \sum_{k=1}^{\infty} k^2 \int_{(k-\frac{1}{2})r}^{(k+\frac{1}{2})r} e^{-x^2/2} dx.$$

*Math. Mag.* **88** (2015) 72–78. doi:10.4169/math.mag.88.1.72. 

We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals must, in general, be accompanied by solutions and by any 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.

Solutions should be written in a style appropriate for this MAGAZINE.

Solutions and new proposals should be mailed to Bernardo M. Ábrego, Problems Editor, Department of Mathematics, California State University, Northridge, 18111 Nordhoff St, Northridge, CA 91330-8313, or mailed electronically (ideally as a LaTeX or pdf file) to mathmagproblems@csun.edu. All communications, written or electronic, should include **on each page** the reader's name, full address, and an e-mail address and/or FAX number.

**1963.** *Proposed by D. M. Bătineţă–Giurgiu, Matei Basarab National College, Bucharest, Romania and Neculai Stanciu, George Emil Palade Secondary School, Buzău, Romania.*

Consider an arbitrary triangle. Let $R$, $r$, and $s$ denote the circumradius, the inradius, and the semiperimeter of the triangle, respectively. Prove that if $x$ and $y$ are positive real numbers, then

$$2s(s^2 - 6Rr - 3r^2)x + (s^2 + 4Rr + r^2)y \geq 8\sqrt[4]{12x(Rrsy)^3}.$$

**1964.** *Proposed by Branko Ćurgus, Western Washington University, Bellingham, WA.*

Let $n$ be a positive integer and $a_1, a_2, \ldots, a_n$ be distinct real numbers. Consider the $n \times n$ Vandermonde matrix defined as

$$V(a_1, a_2, \ldots, a_n) = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix}.$$

Prove that, for an arbitrary nonzero real number $x$, the eigenvalues of the matrix

$$V(xa_1, xa_2, \ldots, xa_n)(V(a_1, a_2, \ldots, a_n))^{-1}$$

are $1, x, x^2, \ldots, x^{n-1}$.

**1965.** *Proposed by Elias Lampakis, Kiparissia, Greece.*

Let $x$, $y$, and $z$ be real numbers such that $x > 1$, $y > 1$, $z > 1$, and $x + y + z = xyz$. Prove that,

$$\sum_{\text{cyc}} (3x^5 - 10x^3 + 3x)(1 - 3y^2) \leq xyz(3 - x^2)(3 - y^2)(3 - z^2),$$

where the sum runs cyclically over the variables $x$, $y$, and $z$.

## Quickies

*Answers to the Quickies are on page 78.*

**Q1047.** *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

Let $(X, d)$ be a complete metric space and let $f : X \to X$. Suppose there exists $r > 1$ such that $d(f(x_1), f(x_2)) \geq r \cdot d(x_1, x_2)$ for all $x_1$ and $x_2$ in $X$.

(a) Prove that if $f$ is onto, then $f$ has a unique fixed point in $X$.

(b) Is the conclusion in part (a) still true without the assumption that $f$ is onto?

**Q1048.** *Proposed by Steven R. Conrad, Manhasset, NY.*

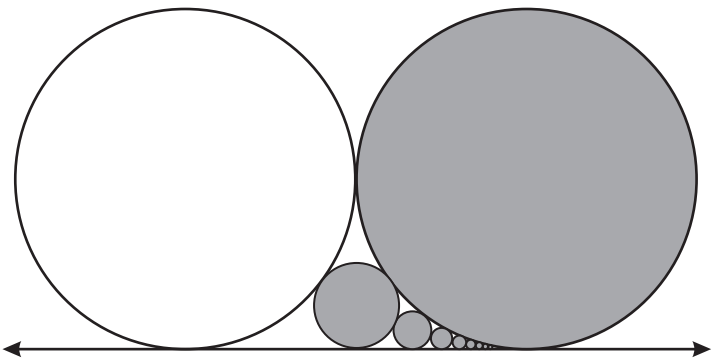
What is the slope of $y^5 + y^3 = x^3$ as it passes through the origin?

# Solutions

**A sequence of pairwise tangent disks** **February 2014**

**1936.** *Proposed by Allen Schwenk, Western Michigan University, Kalamazoo, MI.*

Two mutually tangent circles of radius 1 lie on a common tangent line as shown. The circle on the left is colored white and the circle on the right is colored gray. A third, smaller circle, is tangent to both of the larger circles and the line, and it is also colored gray. An infinite sequence of gray circles all tangent to the line are inserted as follows: Each subsequent circle is tangent to the preceding circle, and it is also tangent to the largest gray circle. What is the total area bounded by the gray circles?



*Solution by Northern University Math Problem Solving Group, Evanston, IL.*

We use Descartes' kissing circles theorem in the particular case in which one of the circles is degenerate with zero curvature (a straight line). Then the curvatures $k_1$, $k_2$, and $k_3$ of the other circles verify

$$(k_1 + k_2 + k_3)^2 = 2(k_1^2 + k_2^2 + k_3^2),$$

or equivalently,

$$k_3 = k_1 + k_2 \pm 2\sqrt{k_1 k_2} = (\sqrt{k_1} \pm \sqrt{k_2})^2,$$

with the plus and minus signs corresponding to the cases in which the circle with curvature $k_3$ inscribes or circumscribes the other two circles and line, respectively. If $k_n$, $n \geq 1$, are the curvatures of the gray circles when the circles are sorted in decreasing size (increasing curvature), then they verify the recurrence $k_1 = 1$, and $k_{n+1} = (1 + \sqrt{k_n})^2$ for $n \geq 1$. Rewriting the recurrence for the square roots of the curvatures $s_n = \sqrt{k_n}$, we get $s_1 = 1$, $s_{n+1} = 1 + s_n$, $n \geq 1$, hence $s_n = n$ and $k_n = n^2$. We can now write the area of the $n$th circle as $A_n = \pi/k_n^2 = \pi/n^4$. The gray area will then be given by the sum of the following series:

$$A_{\text{gray}} = \sum_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty} \pi/n^4 = \pi\zeta(4) = \frac{\pi^5}{90},$$

where $\zeta$ is the Riemann zeta-function.

*Also solved by Armstrong Problem Solvers; George Apostolopoulos (Greece); Herb Bailey; Dionne T. Bailey, Elsie M. Campbell and Charles R. Diminnie; The Barrie School Circle; Michel Bataille (France); Rich Bauer;*

*Gerald E. Bilodeau; Irina Boyadzhiev; Brian Bradie; Stan Byrd; Leah Buck and Richard Daquila; Robert Calcaterra; Tyler Chenhall and Rochelle Starrett; Byungmo Cho (Korea); John Christopher; Timothy V. Craine; Tim Cross (United Kingdom); Neil Curwen (United Kingdom); Prithwijit De (India); William Dickinson; Robert L. Doucette; Steve Edwards; Gino T. Fala; Dmitry Fleischman; Marty Getz and Dixon Jones; J. A. Grzesik; GW-stat Problem Solving Group; Marcus Harbol; Eugene A. Herman; John G. Heuver (Canada); Tom Jager; Stephen S. Jensen; Omran Kouba (Syria); Elias Lampakis (Greece); Abhay Malik; Luke Mannion; David R. Martin; Caleb G. Mcwhorter; Jerry Minkus; Somasundaram Muralidharan (India); Rituraj Nandan; Valerian M. Nita; Moubinool Omarjee (France); Herman Roelants (Belgium); Gregory S. Ronsse; Raul A. Simon (Chile); Gregory S. Ronsse; Carl M. Russell (Germany); Edward Schmeichel; Randy K. Schwartz; Seton Hall University Problem Solving Group; Shrey Shah (India); Brandon Sim; John H. Smith; Zoe Smith; Purva Thakre; Nora Thornber; Michael Vowe (Switzerland); Mazen Zarrouk; and the proposer. There were three incorrect or incomplete submissions.*

**The perimeter and inradius from a Pythagorean triple** **February 2014**

**1937.** *Proposed by Herb Bailey, Rose-Hulman Inst. Tech, Terre Haute, IN, Underwood Dudley, Tallahasse, FL, and W. C. Gosnell, Amherst, MA.*

Let $P$ and $r$, respectively, denote the perimeter and the inradius of a triangle. The value $P - r$ is a square for the right triangle with sides 20, 21, and 29. The value $P + r$ is a square for the right triangle with sides 51, 140, and 149. Show that there are no primitive integer right triangles such that both $P - r$ and $P + r$ are squares. (A right triangle with integer sides is primitive if the greatest common divisor of the side lengths is one.)

*Solution by Robert L. Doucette, McNeese State University, Lake Charles, LA.*

Let $a$, $b$, and $c$ be the lengths of the sides of a primitive integer right triangle, with $c$ the length of the hypotenuse. We may assume that $a$ is even and that $b$ is odd. Using the formula for primitive Pythagorean triples, $a = 2mn$, $b = m^2 - n^2$, and $c = m^2 + n^2$ for some positive integers $m$ and $n$ such that $m > n$, $\gcd(m, n) = 1$, and $m \not\equiv n \pmod 2$. Then the perimeter is $P = 2m(m + n)$ and twice the area is $rP = ab$. Thus $rm(m + n) = mn(m^2 - n^2)$ and $r = n(m - n)$.

Now suppose by contradiction that there exist integers $u$ and $v$ such that $P - r = u^2$ and $P + r = v^2$. Since $u^2 + v^2 = 2P$, then $u^2 + v^2 \equiv 0 \pmod 4$ and so both $u$ and $v$ are even. This implies that $r = v^2 - P = n(m - n)$ is even as well. Since $m - n$ is odd, $n$ must be even and $m$ must be odd. Because $m(m + n)$ is odd and equals the sum of two squares, $(u/2)^2 + (v/2)^2$, it follows that $m(m + n) \equiv 1 \pmod 4$. Note that $n \equiv 2 \pmod 4$ and $m \equiv 1$ or $3 \pmod 4$ imply $m(m + n) \equiv 3 \pmod 4$. Thus $n \equiv 0 \pmod 4$, and so $r \equiv 0 \pmod 4$. Since $P = r + u^2$, then $P \equiv 0 \pmod 4$. But this implies that $m(m + n)$ is even, which is impossible.

*Also solved by Michel Bataille (France), Brian D. Beasley, Robert Calcaterra, John Christopher, Bill Cowieson, Elias Lampakis (Greece), John P. Robertson, Seton Hall University Problem Solving Group, Neculai Stanciu and Titu Zvonaru (Romania), Timothy Woodcock, and the proposers.*

**An asymmetric triangle inequality** **February 2014**

**1938.** *Proposed by G. W. Indika Shameera Amarasinghe, The Open University of Sri Lanka (OUSL), Nawala, Sri Lanka.*

Let $ABC$ be a triangle with $a = BC$, $b = AC$, and $c = AB$. Suppose that there is a point $D$ on $\overline{BC}$ such that $AD = a$. Prove that

$$\frac{5a^2}{bc} + \frac{b^3}{a^2c} + \frac{c^3}{a^2b} \geq \frac{13}{2}.$$

*Solution by Michel Bataille, Rouen, France.*

Because $a$, $b$, and $c$ are positive, the required inequality is equivalent to

$$5a^4 + b^4 + c^4 - \frac{13}{2}a^2bc \geq 0. \tag{1}$$

First, it follows from the Arithmetic Mean–Geometric Mean inequality that

$$5a^4 + b^4 + c^4 - \frac{13}{2}a^2bc \geq 5a^4 + 2b^2c^2 - \frac{13}{2}a^2bc = \frac{1}{2}(5a^2 - 4bc)(2a^2 - bc). \quad (2)$$

Second, let $\theta = \angle ADB$, $R$ the circumradius of $\triangle ABC$, and $F$ its area. The Law of Sines for $\triangle ABD$ and $\triangle ABC$ implies that

$$\sin\theta = \frac{AB\sin B}{AD} = \frac{c\sin B}{a} = \frac{bc}{2aR}.$$

Because $F = abc/(4R)$, it follows that $4F = 2a^2\sin\theta$. By Heron's formula, $16F^2 = (a+b+c)(-a+b+c)(a-b+c)(a+b-c) = 2a^2b^2 + 2b^2c^2 + 2c^2a^2 - a^4 - b^4 - c^4$. From these observations and the Arithmetic Mean–Geometric Mean inequality, it follows that

$$\begin{aligned} 5a^4 + b^4 + c^4 - \frac{13}{2}a^2bc &\geq a^4(4\sin^2\theta + 1) + b^4 + c^4 - \frac{13}{2}a^2bc \\ &= 16F^2 + a^4 + b^4 + c^4 - \frac{13}{2}a^2bc \\ &= \frac{1}{2}\left(4b^2c^2 + 4a^2(b^2+c^2) - 13a^2bc\right) \\ &\geq \frac{1}{2}\left(4b^2c^2 + 8a^2bc - 13a^2bc\right) = \frac{1}{2}bc(4bc - 5a^2). \quad (3) \end{aligned}$$

The desired inequality (1) follows from (2) if $5a^2 - 4bc \geq 0$ and from (3) if $5a^2 - 4bc \leq 0$.

*Editor's Note.* It follows from the presented solution that equality occurs if and only if $b = c = (\sqrt{5}/2)a$; that is, if and only if $\triangle ABC$ is similar to the triangle with vertices $(-1, 0)$, $(1, 0)$, and $(0, 2)$.

*Also solved by Herb Bailey, Robert Calcaterra, Elias Lampakis (Greece), and the proposer.*

**A quadratic functional equation** **February 2014**

**1939.** *Proposed by Marcel Chirita, Bucarest, Romania.*

Find all the functions $f : \mathbb{R} \to \mathbb{R}$ such that $f(0) = 1$, $f$ is continuous at $x = 0$, and there is a fixed number $a$, $0 < a < 1$, such that

$$3f(x) - 5f(ax) + 2f(a^2x) = x^2 + x,$$

for all real numbers $x$.

*Solution by Robert L. Doucette, McNeese State University, Lake Charles, LA.*

Suppose that $0 < a < 1$. Let $g : \mathbb{R} \to \mathbb{R}$ be defined by

$$g(x) = \frac{x^2}{3 - 5a^2 + 2a^4} + \frac{x}{3 - 5a + 2a^2}.$$

The function $g$ satisfies the equation $3g(x) - 5g(ax) + 2g(a^2x) = x^2 + x$ for all real numbers $x$.

Suppose that the function $f$ is as described in the problem proposal. Then the function $h = f - g$ satisfies $3h(x) - 5h(ax) + 2h(a^2x) = 0$ for all real numbers $x$. The function $k$, defined by $k(x) = h(x) - h(ax)$, satisfies $3k(x) = 2k(ax)$ for all real $x$.

For $x \in \mathbb{R}$, we may show by induction that $k(x) = (2/3)^n k(a^n x)$ for $n \geq 0$. Because $\lim_{n\to\infty} a^n = 0$ and the continuity of $f$ at 0 insures the continuity of $k$ at 0, it follows that

$$k(x) = \lim_{n\to\infty} \left(\frac{2}{3}\right)^n k(a^n x) = 0 \cdot k(0) = 0,$$

for all real $x$. This means that $h(x) = h(ax)$, for all $x \in \mathbb{R}$. Again by induction, $h(x) = h(a^n x)$, for all $n \geq 0$ and $x \in \mathbb{R}$. Because the function $h$ is continuous at 0,

$$h(x) = \lim_{n\to\infty} h(a^n x) = h(0) = f(0) = 1,$$

for all real $x$. We conclude that if $f$ satisfies the conditions stated in the problem proposal then $f = g + 1$.

Since $g + 1$ satisfies these conditions, this problem has exactly one solution:

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

for all real $x$.

*Also solved by Arkady Alt, Armstrong Problem Solvers, Michel Bataille (France), Robert Calcaterra, Bill Cowieson, Costas Efthimiou, Eugene A. Herman, The Iowa State University Student Problem Solving Group, Tom Jager, Omran Kouba (Syria), Elias Lampakis (Greece), Charles Z. Martin, Moubinool Omarjee (France), Phil Schmidt, Connie Xu, and the proposer.*

**Fixed point on continuous nondilations** **February 2014**

**1940.** *Proposed by Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

(a) Let $(X, \|\cdot\|)$ be a real finite dimensional normed linear space and let $B = \{x \in X : \|x\| \leq 1\}$. Let $T : B \to X$ be a continuous function such that for each $x$, a unit vector in $B$, there is no $k > 1$ such that $T(x) = kx$. Is it necessary that $T$ have a fixed point in $B$?

(b) If $(X, \|\cdot\|)$ is not finite dimensional, is it necessary that $T$ have a fixed point in $B$?

*Solution by Gregory S. Ronsse, Eastern Illinois University, Charleston, IL.*

(a) The answer is yes. We will prove the contrapositive. Assume that $T$ has no fixed point in $B$, then for $x \in B$, the map

$$x \to \frac{T(x) - x}{\|T(x) - x\|}$$

is well defined and maps $B$ into the unit vectors in $B$. By the Brouwer or Schauder fixed point theorem, there exists a unit vector $x_0$ such that

$$x_0 = \frac{T(x_0) - x_0}{\|T(x_0) - x_0\|}.$$

That is, $T(x_0) = (\|T(x_0) - x_0\| + 1)x_0$. But $k = \|T(x_0) - x_0\| + 1 > 1$.

(b) The answer is no. Let $X$ be the space of polynomial functions on the interval $[0, 1]$. Equip $X$ with the sup norm. For $f \in X$ and $x \in [0, 1]$, define $T(f)(x) = xf(x) + 1$. The function $T : X \to X$ is easily seen to be continuous. Since the

degree of $T(f)$ is strictly greater than the degree of $f$, $T$ has no fixed points and there does not exist any constant $k$ such that $T(f) = kf$ for $f \in X$.

*Also solved by Robert Calcaterra, Bill Cowieson, Robert L. Doucette, Eugene A. Herman, Omran Kouba (Syria), Reiner Martin (Germany), Dave Trautman, and the proposer.*

## Answers

*Solutions to the Quickies from page 74.*

**A1047.**

(a) Clearly $f$ is one-to-one and since it is also onto, $f$ has an inverse function $f^{-1}$. Because $d(f(x_1), f(x_2)) \geq r \cdot d(x_1, x_2)$, it follows that

$$d(f^{-1}(y_1), f^{-1}(y_2)) \leq \frac{1}{r} \cdot d(y_1, y_2)$$

for all $y_1$ and $y_2$ in $X$. Since $0 < 1/r < 1$, the inverse function $f^{-1}$ has a unique fixed point by Banach's contraction theorem. If $x_0$ is this fixed point, then $f^{-1}(x_0) = x_0$ and thus $f(x_0) = x_0$. If $y_0$ were another fixed point of $f$, then $y_0$ would also be a fixed point of $f^{-1}$, and this would contradict the uniqueness. Thus $f$ has a unique fixed point in $X$.

(b) The answer is no. To see this, let $X = [1, \infty)$ with the usual metric, $r = 2$, and $f : X \to X$ be defined so that $f(x) = 2x$. Then $X$ is a complete metric space and $|f(x_1) - f(x_2)| = 2|x_1 - x_2| = r|x_1 - x_2|$ for all $x_1$ and $x_2$ in $X$. But clearly $f$ has no fixed points in $X$.

**A1048.** In the equation $y^5 + y^3 = x^3$, the higher degree term zeroes out before the two lower degree terms, so near the origin, the graph behaves like $y^3 = x^3$ or $y = x$, so its slope is 1. In other words, the slope at the origin $s$ satisfies that

$$s = \lim_{x,y \to 0} \frac{y}{x} = \lim_{y \to 0} \sqrt[3]{\frac{y^3}{x^3}} = \lim_{y \to 0} \sqrt[3]{\frac{y^3}{y^5 + y^3}} = \lim_{y \to 0} \sqrt[3]{\frac{1}{y^2 + 1}} = 1.$$

# 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.*

Shiflet, Angela B., and George W. Shiflet, *Introduction to Computational Science: Modeling and Simulation for the Sciences*, 2nd ed., Princeton University Press, 2014; xxxiii + 816 pp, $99.50. ISBN 978-0-691-16071-9.

What would an applied mathematics course look like if it left out proofs, de-emphasized skill in computation, but used mathematical concepts to model a wide range of real applications? It might look like this book. The first edition (2006) featured system dynamics models, empirical models, and cellular automaton simulations. This edition adds a few more modules of systems dynamics projects and two more sections (200 pages) on agent-based models and modeling with matrices. The authors take a generic approach but provide at the accompanying website (http://wofford-ecs.org/) computer files for several software tools, as well as tutorials, models, PDF files, and datasets. For each kind of model, at least one software tool is free. Computer programming is not a prerequisite; concepts and notation of calculus are, but no facility is required in "taking derivatives." So, is this book "computational science," or is it a new face for mathematics? Most of the modules are excellent. The project on colon cancer provides only one page of biological explanation (with intimidating terminology) and two pages of under-explained diagrams—in other words, it offers a real open-ended modeling problem, not "pre-digested," in which students first have to investigate and learn about the phenomenon.

Havil, Julian, *John Napier: Life, Logarithms and Legacy*, Princeton University Press, 2014; xv + 279 pp, $35. ISBN 978-0-691-15570-8.

Bellhouse, David, An analysis of errors in mathematical tables, in *Proceedings of the Canadian Society for History and Philosophy of Mathematics, 2013* **26** (2014) 18–30.

"John Napier deserves better than obscurity." When was the last time you heard natural logarithms called by their earlier name, "Napierian logarithms"? and do you remember what logarithms to base 10 used to be called?* Havil gives the lineage and a biography of Napier (1550–1617), an interpretation of his long book on the bible's Book of Revelation, the history of his original conception of logarithms for trigonometric calculations, and the details of the computation of his tables. A dozen appendices address everything from the politics and religion of Scotland in Napier's time to other work of Napier. Bellhouse carefully detects typographical and other errors in Napier's 1614 book of tables; he finds that the (relatively few) typographical errors are fit well by a Poisson distribution. *("Common"; if you're really old, "Briggsian.")

Special Issue: Trends in Undergraduate Research in the Mathematical Sciences (Chicago, 2012). *Involve: A Journal of Mathematics* **7** no. 3 (2014). http://msp.org/involve/2014/7-3/index.xhtml.

Special open-access issue devoted to 20 reports and ideas (promotion, funding, outlets, special-needs students, encouragement) on undergraduate research in mathematics, by authors from leading programs. (This is a journal, with an illustrious editorial board, that usually publishes papers by students co-authoring with faculty.)

*Math. Mag.* **88** (2015) 79–80. doi:10.4169/math.mag.88.1.79. 

Weber, Bruce, and Julie Rehmeyer. Alexander Grothendieck, math enigma, dies at 86. *New York Times* (14 November 2014). http://www.nytimes.com/2014/11/16/world/europe/alexander-grothendieck-math-enigma-dies-at-86.

Mumford, David, and John Tate. Alexander Grothendieck (1928–2014). *Nature* 517272 (15 January 2015) 272. http://www.nature.com/nature/journal/v517/n7534/full/517272a.html.

Mumford, David. Can one explain schemes to biologists? http://www.dam.brown.edu/people/mumford/blog/2014/Grothendieck.html.

Alexander Grothendieck invented the concept of scheme and revolutionized algebraic geometry. He won a Fields Medal in 1964 and is regarded by some as the greatest mathematician of the 20th century. He abandoned mathematics and spent the last 20 years as a recluse in the Pyrenees. It is a valuable lesson to compare Mumford and Tate's published obituary to the original text (rejected as too technical) at the blog by Mumford. The *Nature* editor had objected to "higher degree polynomials," "infinitesimal vectors," and "complex space." Mumford comments that "the central stumbling block for explaining schemes was the word 'ring.'" But the original text also featured the terms "variety," "affine scheme," "morphisms," "functors," "stacks," "third cohomology group," and "motives," the term and notation for a "Galois group," and the phrase "the sphere lives algebraically too." How are your math majors on all that? Mumford concludes that "the trick is to bootstrap the math on things scientists know, simplifying definitions . . . and getting to some core non-trivial motivating example if possible." That's indeed tough to do.

Spivak, David I., *Category Theory for the Sciences*, MIT Press, 2014; viii + 486 pp, $50. ISBN 978-0-26202813-4. Open-access version: http://category-theory.mitpress.mit.edu/.

Giesa, Tristan, David I. Spivak, and Markus J. Buehler. Reoccurring patterns in hierarchical protein materials and music: The power of analogies. *BioNanoScience* **1** no. 4 (December 2011) 153–161. http://web.mit.edu/mbuehler/www/papers/BioNanoScience_2011_3.pdf.

Spivak, David I., Tristan Giesa, Elizabeth Wood, and Markus J. Buehler. Category theoretic analysis of hierarchical protein materials and social networks. *PLoS ONE* **6** no. 9 (1 September 2011) e23911. http://web.mit.edu/mbuehler/www/papers/PLoS_ONE_2011.pdf.

Is the content of this book what Mumford and Tate would like every scientist to know about category theory, so as to understand their obituary of Grothendieck? Author Spivak introduces the concept and word *olog*, coined by him, which stands for "ontology log" and means a certain kind of analogy. Spivak calls it a "a category-theoretic model for knowledge representation," a kind of theory for analogy. Spivak endeavors to show that category theory can be a useful modeling language outside of mathematics, which "can be communicated to scientists with no background beyond linear algebra." (Well, that loses most nonphysical scientists.) There are coproducts, pullbacks, colimits, coequalizers, and simplicial complexes; there are also exercises. Unfortunately, the book does not include the motivating examples in the two papers cited above, which feature an intriguing analogy between spider silk and musical composition.

Cheng, Eugenia, *How to Bake $\pi$: An Edible Exploration of the Mathematics of Mathematics*, Basic Books, 2015; 288 pp, $27.99(P). ISBN 978-0-465-05171-7.

Here it is! We may have here the author who can explain schemes to biologists—especially since almost half of her book is devoted to explaining category theory to nonmathematicians! Each chapter begins with a simple food recipe: "Math, like recipes, has both ingredients and method." The analogy indeed works, and author Cheng's exposition captures much of the spirit of the "mathematical method." Chapters feature mayonnaise/hollandaise sauce (abstraction), olive oil plum cake (generalization), Jaffa cakes (axiomatization), and more: "[M]athematics works by abstraction, . . . seeks to study the principles and processes behind things, and . . . seeks to axiomatize and generalize those things." And what is mathematics? "[T]he study of anything that obeys the rules of logic, using the rules of logic. . . . [M]ath is defined by the techniques it uses to study things, and . . . the things it studies are determined by those techniques." And category theory? "The process of working out exactly which parts of math are easy, and the process of making as many parts of math easy as possible. . . . It serves to organize mathematics." However, though "math is easy [because it is logical], life is hard, and therefore math isn't life."

# 75th Annual William Lowell Putnam Mathematical Competition

*Editor's Note:* Additional solutions will be printed in the *Monthly* later in the year.

## PROBLEMS

**A1.** Prove that every nonzero coefficient of the Taylor series of

$$(1 - x + x^2)e^x$$

about $x = 0$ is a rational number whose numerator (in lowest terms) is either 1 or a prime number.

**A2.** Let $A$ be the $n \times n$ matrix whose entry in the $i$-th row and $j$-th column is

$$\frac{1}{\min(i, j)}$$

for $1 \le i, j \le n$. Compute $\det(A)$.

**A3.** Let $a_0 = 5/2$ and $a_k = a_{k-1}^2 - 2$ for $k \ge 1$. Compute

$$\prod_{k=0}^{\infty}\left(1 - \frac{1}{a_k}\right)$$

in closed form.

**A4.** Suppose $X$ is a random variable that takes on only nonnegative integer values, with $E[X] = 1$, $E[X^2] = 2$, and $E[X^3] = 5$. (Here $E[Y]$ denotes the expectation of the random variable $Y$.) Determine the smallest possible value of the probability of the event $X = 0$.

**A5.** Let

$$P_n(x) = 1 + 2x + 3x^2 + \cdots + nx^{n-1}.$$

Prove that the polynomials $P_j(x)$ and $P_k(x)$ are relatively prime for all positive integers $j$ and $k$ with $j \neq k$.

**A6.** Let $n$ be a positive integer. What is the largest $k$ for which there exist $n \times n$ matrices $M_1, \ldots, M_k$ and $N_1, \ldots, N_k$ with real entries such that for all $i$ and $j$, the matrix product $M_i N_j$ has a zero entry somewhere on its diagonal if and only if $i \neq j$?

**B1.** A *base* 10 *over-expansion* of a positive integer $N$ is an expression of the form

$$N = d_k 10^k + d_{k-1} 10^{k-1} + \cdots + d_0 10^0$$

with $d_k \neq 0$ and $d_i \in \{0, 1, 2, \ldots, 10\}$ for all $i$. For instance, the integer $N = 10$ has two base 10 over-expansions: $10 = 10 \cdot 10^0$ and the usual base 10 expansion $10 = 1 \cdot 10^1 + 0 \cdot 10^0$. Which positive integers have a unique base 10 over-expansion?

*Math. Mag.* **88** (2015) 81–88. doi:10.4169/math.mag.88.1.81. 

**B2.** Suppose that $f$ is a function on the interval $[1, 3]$ such that $-1 \leq f(x) \leq 1$ for all $x$ and $\int_1^3 f(x)\,dx = 0$. How large can $\int_1^3 f(x)/x\,dx$ be?

**B3.** Let $A$ be an $m \times n$ matrix with rational entries. Suppose that there are at least $m + n$ distinct prime numbers among the absolute values of the entries of $A$. Show that the rank of $A$ is at least 2.

**B4.** Show that for each positive integer $n$, all the roots of the polynomial

$$\sum_{k=0}^{n} 2^{k(n-k)} x^k$$

are real numbers.

**B5.** In the 75th Annual Putnam Games, participants compete at mathematical games. Patniss and Keeta play a game in which they take turns choosing an element from the group of invertible $n \times n$ matrices with entries in the field $\mathbb{Z}/p\mathbb{Z}$ of integers modulo $p$, where $n$ is a fixed positive integer and $p$ is a fixed prime number. The rules of the game are:

1. A player cannot choose an element that has been chosen by either player on any previous turn.
2. A player can only choose an element that commutes with all previously chosen elements.
3. A player who cannot choose an element on his/her turn loses the game.

Patniss takes the first turn. Which player has a winning strategy? (Your answer may depend on $n$ and $p$.)

**B6.** Let $f\colon [0, 1] \to \mathbb{R}$ be a function for which there exists a constant $K > 0$ such that $|f(x) - f(y)| \leq K|x - y|$ for all $x, y \in [0, 1]$. Suppose also that for each rational number $r \in [0, 1]$, there exist integers $a$ and $b$ such that $f(r) = a + br$. Prove that there exist finitely many intervals $I_1, \ldots, I_n$ such that $f$ is a linear function on each $I_i$ and $[0, 1] = \bigcup_{i=1}^{n} I_i$.

## SOLUTIONS

**Solution to A1.** From the Taylor series for $e^x$, we have

$$\begin{aligned}(1 - x + x^2)e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} - \sum_{n=1}^{\infty} \frac{x^n}{(n-1)!} + \sum_{n=2}^{\infty} \frac{x^n}{(n-2)!} \\ &= 1 + \sum_{n=2}^{\infty} \frac{1 - n + n(n-1)}{n!}\, x^n \\ &= 1 + \frac{1}{2}x^2 + \sum_{n=3}^{\infty} \frac{n-1}{(n-2)!\,n}\, x^n.\end{aligned}$$

Thus, it is enough to show that for $k = n - 1 \geq 2$, the numerator of the rational number $\dfrac{k}{(k-1)!}$ is either 1 or a prime number. If $k$ is not prime, write $k = pq$ where $p$ is prime. Then $q < k$ occurs as a factor in $(k - 1)!$, so $q$ can be canceled to simplify the fraction to one (perhaps not in lowest terms) of the form $p/m$. It follows that in lowest terms, the numerator is either 1 or $p$.

**Solution to A2.** The value is $\det(A) = \dfrac{(-1)^{n-1}}{n!(n-1)!}$. This is clear for $n = 1$, so we'll assume $n \geq 2$. Modify the notation so the given matrix is $A_n$. Note that the first $n-1$ rows and columns of $A_n$ form $A_{n-1}$. Subtracting the $(n-1)$st column of $A_n$ from the $n$th column, we get a matrix in which all entries in the last column are 0 except for the final entry, which is $1/n - 1/(n-1) = -1/n(n-1)$. Expanding the determinant along that column, we see that $\det(A_n) = -\det(A_{n-1})/n(n-1)$, and the answer follows by an easy induction on $n$.

**Solution to A3.** The answer is $3/7$. Note that $a_0 = 2 + \dfrac{1}{2}$ and $a_1 = 2^2 + \dfrac{1}{2^2}$; it is easy to see by induction that $a_k = 2^{2^k} + \dfrac{1}{2^{2^k}} = \dfrac{2^{2^{k+1}} + 1}{2^{2^k}}$. Thus,

$$\begin{aligned} 1 - \frac{1}{a_k} &= \frac{2^{2^{k+1}} - 2^{2^k} + 1}{2^{2^{k+1}} + 1} \\ &= \frac{1 - (1/2)^{2^k} + (1/2)^{2^{k+1}}}{1 + (1/2)^{2^{k+1}}} \\ &= \frac{1 - y_k + y_k^2}{1 + y_k^2} \\ &= \frac{1 + y_k^3}{(1 + y_k)(1 + y_k^2)}, \end{aligned}$$

where $y_k = (1/2)^{2^k}$.

Now for any $z$ with $|z| < 1$, we have a well-known convergent infinite product $P(z) = \prod_{k=0}^{\infty}(1 + z^{2^k})$ whose partial products are $\dfrac{1 - z^{2^{k+1}}}{1 - z}$ and whose value is therefore $P(z) = \dfrac{1}{1 - z}$. From the calculation above, the product we are looking for can be expressed as

$$\begin{aligned} \prod_{k=0}^{\infty}\left(1 - \frac{1}{a_k}\right) &= \frac{\prod_{k=0}^{\infty}(1 + y_k^3)}{\prod_{k=0}^{\infty}(1 + y_k)\prod_{k=0}^{\infty}(1 + y_k^2)} \\ &= \frac{P(1/8)}{P(1/2)P(1/4)} = \frac{8/7}{2 \cdot 4/3} = \frac{3}{7}. \end{aligned}$$

**Solution to A4.** The answer is $1/3$.

To show this, let $f(x) = (x-1)(x-2)(x-3) = x^3 - 6x^2 + 11x - 6$. Then

$$E[f(X)] = E[X^3] - 6E[X^2] + 11E[X] - 6 = -2.$$

On the other hand, $f(k) \geq 0$ for all *positive* integers $k$. Therefore, denoting the probability of the event $X = k$ by $p_k$, we have

$$E[f(X)] = \sum_{k=0}^{\infty} f(k)p_k \geq f(0)p_0 = -6p_0,$$

so $-2 \geq -6p_0$, that is, $p_0 \geq 1/3$. To see that the value $1/3$ is possible, look at the special case in which $p_k = 0$ for $k \geq 4$. From the given expectations, we then have that

$$\begin{aligned} p_0 + p_1 + p_2 + p_3 &= 1, \\ p_1 + 2p_2 + 3p_3 &= 1, \\ p_1 + 4p_2 + 9p_3 &= 2, \\ p_1 + 8p_2 + 27p_3 &= 5. \end{aligned}$$

These equations have the unique solution

$$p_0 = 1/3,\ p_1 = 1/2,\ p_2 = 0,\ p_3 = 1/6.$$

**Solution to A5.** Note that $P_n(x)$ is a partial sum of the Taylor series for $\dfrac{1}{(1-x)^2}$, which motivates considering the new polynomials

$$Q_n(x) = (1-x)^2 P_n(x) = 1 - (n+1)x^n + nx^{n+1}.$$

Now suppose that $P_j(x)$ and $P_k(x)$, with $j \neq k$, are not relatively prime, so they have a common factor $R(x)$ of positive degree. Then $R(x)$ has some root $z$ in $\mathbb{C}$, and because $P_j(0), P_j(1) \neq 0$, $z \neq 0, 1$. We have $Q_j(z) = Q_k(z) = 0$; that is, both for $n = j$ and for $n = k$ we have

$$\begin{aligned} 1 - (n+1)z^n + nz^{n+1} &= 0 \\ z^{-n} &= n + 1 - nz. \end{aligned}$$

Note that this equation also holds for $n = 0$ and for $n = -1$. Also, taking the square of the modulus of both sides yields

$$|z^{-2}|^n = |n(1-z) + 1|^2,$$

for the four values $-1, 0, j, k$ of $n$. If we put $a = |z^{-2}|$, then this equation can be written in the form $a^n - g(n) = 0$, where the right-hand side is a polynomial in $n$ of degree $\leq 2$ with real coefficients. Because this equation has four distinct solutions for $n$, by Rolle's theorem there must be three distinct solutions to $a^n \log a - g'(n) = 0$, and then two distinct solutions to $a^n (\log a)^2 - g''(n) = 0$. But $g''(n)$ is constant, so this is only possible if $a = 1$ and hence $g''(n) = 0$, meaning that $|n(1-z) + 1|^2$ is at most linear in $n$. But this is only true when $1 - z = 0$, so $z = 1$, a contradiction.

**Solution to A6.** We'll show that the largest value of $k$ is $n^n$. Given such matrices $M_1, \ldots, M_k$ and $N_1, \ldots, N_k$, let $A$ be the $k \times k$ matrix whose $i, j$ entry $A_{i,j}$ is the product of all the diagonal entries of $M_i N_j$. Then $A_{i,j} \neq 0$ if and only if $i = j$, so $A$ has rank $k$. On the other hand,

$$\begin{aligned} A_{i,j} &= \prod_{s=1}^{n} (M_i\, N_j)_{s,s} \\ &= \prod_{s=1}^{n} \sum_{m=1}^{n} (M_i)_{s,m} (N_j)_{m,s} \\ &= \sum_{m_1=1}^{n} \cdots \sum_{m_n=1}^{n} \prod_{s=1}^{n} (M_i)_{s,m_s} (N_j)_{m_s,s}. \end{aligned}$$

For each choice of $m_1, \ldots, m_n$, the matrix whose $i, j$ entry is

$$\prod_{s=1}^{n} (M_i)_{s,m_s} (N_j)_{m_s,s}$$

has rank at most 1, since this entry is the product of $\prod_{s=1}^{n}(M_i)_{s,m_s}$, which depends only on $i$, and $\prod_{s=1}^{n}(N_j)_{m_s,s}$, which depends only on $j$. Thus, $A$ is the sum of $n^n$ matrices of rank at most 1 and hence has rank at most $n^n$, showing that $k \le n^n$. To show that $k = n^n$ is possible, let $M_1, \ldots, M_k$ be all the matrices for which every row equals one of the $n$ rows of the identity matrix, and for every $1 \le j \le k$, let $N_j$ be the transpose of $M_j$.

**Solution to B1.** An integer has a unique base 10 over-expansion if and only if it has no digit zero in its usual base 10 expansion. To show this, let the usual base 10 expansion of $N$ be $N = \sum_{j=0}^{k} d_j 10^j$ with $0 \le d_j \le 9$ and $d_k \neq 0$. Suppose some digit is zero, and let $\ell$ be as large as possible so that $d_\ell = 0$. Then set $d'_\ell = 10$, $d'_{\ell+1} = d_{\ell+1} - 1$, and $d'_j = d_j$ if $j \neq \ell, \ell+1$. It is easy to see that $n = \sum_{j=0}^{k} d'_j 10^j$ is another base 10 over-expansion of $N$ (possibly with $d'_k = 0$, in which case $d'_{k-1} = 10$ is the highest "digit").

Conversely, suppose $N$ has two base 10 over-expansions. At least one of these over-expansions $N = \sum_{j=0}^{k} d_j 10^j$ must have some $d_j = 10$ since the usual base 10 expansion of $N$ is unique. We iterate the following procedure: choose $\ell$ as large as possible with $d_\ell = 10$, and put $d'_\ell = 0$, $d'_{\ell+1} = d'_{\ell+1} + 1$, and $d'_j = d_j$ if $j \neq \ell, \ell+1$. (If $\ell = k$, this causes us to replace $k$ with $k+1$.) Then $N = \sum d'_j 10^j$ is a new base 10 over-expansion of $N$. Note that $\sum d'_j = \sum d_j - 9$ but $\sum d'_j$ must still be a nonnegative integer, so this procedure must eventually terminate. At that point, we will have a base 10 over-expansion of $N$ with no "digit" 10, so we have the usual base 10 expansion of $N$. But we always have $d'_\ell = 0$ (for some $\ell$) at every step, so the usual base 10 expansion of $N$ contains a zero.

**Solution to B2.** We'll show that the largest possible value of the integral is $\log(4/3)$. Under the given conditions on $f$, we have

$$\int_1^3 \frac{f(x)}{x}\,dx = \int_1^3 f(x)\left(\frac{1}{x} - \frac{1}{2}\right)dx.$$

But $1/x - 1/2 \ge 0$ for $1 \le x \le 2$ and $1/x - 1/2 \le 0$ for $2 \le x \le 3$, so, using $|f(x)| \le 1$, we get the estimate

$$\begin{aligned}\int_1^3 \frac{f(x)}{x}\,dx &\le \int_1^2 \left(\frac{1}{x} - \frac{1}{2}\right)dx + \int_2^3 \left(\frac{1}{2} - \frac{1}{x}\right)dx \\ &= \log 2 - 1/2 + 1/2 - (\log 3 - \log 2) = \log(4/3).\end{aligned}$$

We get equality for the function defined by $f(x) = 1$ for $1 \le x \le 2$, $f(x) = -1$ for $2 < x \le 3$, which satisfies the conditions.

**Solution to B3** (based on a student paper). If there is a matrix for which the statement is false, that matrix must have rank 1. We prove by induction on $m + n$ that an $m \times n$ matrix $A$ of rank 1 with rational entries can have at most $m + n - 1$ distinct prime numbers among the absolute values of its entries. This is clearly true if $m = 1$ or $n = 1$ because then there are only $m + n - 1$ entries. Suppose $A$ has $P$ distinct prime numbers among the absolute values of its entries, where $P \ge m + n$. Choose $P$ entries of $A$ where these prime absolute values occur. Then if we omit any row or column of $A$, by the induction hypothesis at most $m + n - 2$ of these entries can still be in the resulting matrix, so the omitted row or column had at least $P - m - n + 2$ of the chosen entries. Adding the numbers of chosen entries over all $m + n$ rows and

columns, we get at least $(m+n)(P-m-n+2)$; on the other hand, we count each of the $P$ chosen entries exactly twice, so

$$\begin{aligned} 2P &\geq (m+n)(P-m-n+2) \\ (m+n-2)P &\leq (m+n-2)(m+n) \\ P &\leq m+n \end{aligned}$$

So under our assumption $P \geq m+n$, we have $P = m+n$, and all the inequalities above are equalities; in particular, every row and every column of $A$ has *exactly* 2 of the chosen entries. Now note that because $A$ has rank 1, there are rational numbers $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$ such that the $i, j$ entry of the matrix is $x_i y_j$ (for example, we can take $x_1, \ldots, x_m$ to be the entries of the first column, and $y_1 = 1, y_2, \ldots, y_n$ to be the ratios of the other columns to the first). Then the product of all the chosen entries is $x_1^2 x_2^2 \cdots x_m^2 y_1^2 y_2^2 \cdots y_n^2$, because each of the chosen entries contributes one factor $x_i$ and one factor $y_j$, corresponding to the row and column that that entry is in, and each row and each column comes up exactly twice. In particular, the product of the chosen entries is a perfect square; but its absolute value is a product of distinct primes, a contradiction.

**Solution to B4.** Let $f(x) = \sum_{k=0}^{n} 2^{k(n-k)} x^k$. For $n = 1$, $f(x) = x + 1$, and for $n = 2$, $f(x) = x^2 + 2x + 1$; in both cases, the only root is the real number $-1$. For $n \geq 3$ we consider the $n+1$ numbers

$$a_0 = -2^{-n}, a_1 = -2^{-n+2}, \ldots, a_n = -2^n$$

of the form $a_j = -2^{-n+2j}$. We claim that the values $f(a_0), f(a_1), \ldots, f(a_n)$ of the polynomial at these numbers alternate in sign; specifically, $f(a_j)$ has the sign $(-1)^j$. This will imply that $f$ has a root in each of the intervals $(a_0, a_1), (a_1, a_2), \ldots, (a_{n-1}, a_n)$; thus, $f$ will have $n$ distinct real roots, which is sufficient because $f$ has degree $n$. The signs of

$$f(a_0) = \sum_{k=0}^{n} 2^{k(n-k)} (-2^{-n})^k = \sum_{k=0}^{n} (-1)^k 2^{-k^2} \text{ and}$$

$$f(a_n) = \sum_{k=0}^{n} 2^{k(n-k)} (-2^n)^k = 2^{n^2} \sum_{k=0}^{n} (-1)^k 2^{-(n-k)^2}$$

are easily checked because the terms alternate in sign and their absolute values are strictly decreasing in the first case and increasing in the second.

For any $j$ with $0 < j < n$, we can first group the terms of $f(x)$ as

$$\begin{aligned} &\cdots \\ &+ 2^{(j-5)(n-j+5)} x^{j-5} + 2^{(j-4)(n-j+4)} x^{j-4} \\ &+ 2^{(j-3)(n-j+3)} x^{j-3} + 2^{(j-2)(n-j+2)} x^{j-2} \\ &+ 2^{(j-1)(n-j+1)} x^{j-1} + 2^{j(n-j)} x^j + 2^{(j+1)(n-j-1)} x^{j+1} \\ &+ 2^{(j+2)(n-j-2)} x^{j+2} + 2^{(j+3)(n-j-3)} x^{j+3} \\ &+ 2^{(j+4)(n-j-4)} x^{j+4} + 2^{(j+5)(n-j-5)} x^{j+5} \\ &+ \cdots \end{aligned}$$

Depending on the parity of $j$ and of $n-j$, there may be a single monomial left on each end. Now for $x = a_j$, the trinomial evaluates to 0. Still for $x = a_j$, in the binomials preceding the trinomial, the right-hand term, which has the sign $(-1)^j$, is greater in absolute value than the left-hand term, so the overall sign of the binomial is $(-1)^j$. Similarly, in the binomials following the trinomial, the absolute value of the left-hand term is larger, and the overall sign of the binomial is again $(-1)^j$. If there are monomials left over at the ends, their signs are also $(-1)^j$. Thus, $f(a_j)$ has that same sign, and the claim follows.

**Comment.** The values of the terms in $f(x)$ with $x^{j\pm r}$ at $x = a_j$ are equal (when both of them are present) and are given (when present) by $(-1)^{j\pm r}\, 2^{j^2-r^2}$.

**Solution to B5.** We show that Keeta has a winning strategy if $p$ is odd and Patniss has a winning strategy if $p = 2$. In any group $G$, any subset of commuting elements will generate a commutative subgroup; therefore, at the end of the game, the set of elements that have been chosen will form a subgroup. Patniss wishes to force the order of this subgroup to be odd, and Keeta wishes to force it to be even.

Suppose the finite group $G$ contains an element whose centralizer has odd order, then Patniss can choose such an element and is then guaranteed to win. Otherwise, Patniss must choose an element $g$ whose centralizer $Z_G(g)$ has even order, and Keeta can choose an element $h \in Z_G(g)$ of order 2 (unless the only such element is $g$ itself, in which case Keeta can choose $e$). Since any group that contains an element of order 2 must have even order, at this point Keeta is guaranteed to win. So the question comes down to: for which $n$ and $p$ does $G$ contain an element whose centralizer has odd order?

If $p$ is odd, then $\mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})$ contains the element $-I$ of order 2 in its center; thus, the centralizer of every element has even order, and Keeta will win.

On the other hand, if $p = 2$ we claim that $\mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})$ always contains an element whose centralizer has odd order, so that Patniss will win. Here is a construction of such an element. Let $\mathbb{F}_{2^n}$ be an extension of $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ of degree $n$, and choose $\alpha \in \mathbb{F}_{2^n}$ such that $1, \alpha, \ldots, \alpha^{n-1}$ form a basis of $\mathbb{F}_{2^n}$ over $\mathbb{F}_2$. Patniss can choose $g$ to be the matrix of the map from $\mathbb{F}_{2^n}$ to itself given by multiplication by $\alpha$ (with respect to the basis $1, \alpha, \ldots, \alpha^{n-1}$). Indeed, suppose that $h$ commutes with $g$, and let $T$ be the $\mathbb{F}_2$-linear transformation of $\mathbb{F}_{2^n}$ corresponding to $h$. If $T(1) = x$, the assumption that $T$ commutes with multiplication by $\alpha$ implies easily that $T(\alpha^i) = x\alpha^i$ for any $i \geq 0$. Then $\mathbb{F}_2$-linearity implies that $T$ is the map from $\mathbb{F}_{2^n}$ to itself given by multiplication by $x$. All such maps do commute with $g$, and the condition that $T$ must be invertible means precisely that $x \neq 0$; thus, the centralizer of $g$ has order $2^n - 1$, which is odd.

**Solution to B6.** We first recall a few basic properties of the Farey sequences from elementary number theory and set up notation for them. In this discussion, all rational numbers are written in lowest terms.

Define finite sequences $F_1, F_2, \ldots$ of rational numbers in $[0, 1]$ as follows. Start with $F_1 = (0, 1) = (0/1, 1/1)$. For $n > 1$, form $F_n$ from $F_{n-1}$ by inserting $(a + c)/(b + d)$ between $a/b$ and $c/d$ whenever $a/b, c/d$ are consecutive terms of $F_{n-1}$ with $b + d = n$.

By induction, two consecutive terms $a/b, c/d$ of $F_n$ always satisfy $bc - ad = 1$. If we have a rational number between those consecutive terms, say $a/b < r/s < c/d$, then

$$s = s(bc - ad) = b(cs - dr) + d(br - as) \geq b + d.$$

It follows by induction that $F_n$ consists of *all* the rational numbers $r/s \in [0, 1]$ with $s \leq n$; also, any two consecutive elements $a/b, c/d$ of $F_n$ satisfy $b + d \geq n + 1$ (else

$(a+c)/(b+d)$ would have been inserted in between at some stage). The sequence $F_n$ is commonly known as the $n$-th *Farey sequence*.

We now return to the problem at hand. By hypothesis, for each $r/s \in [0,1] \cap \mathbb{Q}$ we have $f(r/s) \in \frac{1}{s}\mathbb{Z}$. Choose $n \geq 4K$ and let $a/b, c/d$ be consecutive elements of $F_n$. Then $f(c/d) - f(a/b) \in \frac{1}{bd}\mathbb{Z}$; since $|f(c/d) - f(a/b)| \leq K|c/d - a/b| = K|1/(bd)| = K/(bd)$, we have $f(c/d) - f(a/b) = m/(bd)$ for some integer $m$ with $|m| \leq K$. Applying the same argument to the consecutive elements $a/b$, $(a+c)/(b+d)$ and $c/d$ in a later Farey sequence, we end up with the three equations

$$f\left(\frac{c}{d}\right) - f\left(\frac{a}{b}\right) = \frac{m}{bd}$$

$$f\left(\frac{a+c}{b+d}\right) - f\left(\frac{a}{b}\right) = \frac{m_1}{b(b+d)}$$

$$f\left(\frac{c}{d}\right) - f\left(\frac{a+c}{b+d}\right) = \frac{m_2}{d(b+d)}$$

for some $m, m_1, m_2 \in \mathbb{Z}$ with $|m|, |m_1|, |m_2| \leq K$. It follows that

$$\frac{m_1}{b(b+d)} + \frac{m_2}{d(b+d)} = \frac{m}{bd}, \qquad \text{so } m_1 d + m_2 b = m(b+d).$$

If $m_2 \neq m$, we can rewrite this equation as

$$\frac{m_1 - m}{m - m_2} = \frac{b}{d}.$$

Since $bc - ad = 1$, we know that $b/d$ is in lowest terms, so $m_1 - m$ is a multiple of $b$ and $m - m_2$ is a multiple of $d$. However, since $n \geq 4K$ and $b + d \geq n + 1$, we must have $\max(b,d) > 2K$, so either $|m_1 - m| > 2K$ or $|m - m_2| > 2K$, contradicting $|m|, |m_1|, |m_2| \leq K$, unless $m_1 = m_2 = m$. This means that we now have

$$f\left(\frac{c}{d}\right) - f\left(\frac{a}{b}\right) = \frac{m}{bd} = m\left(\frac{c}{d} - \frac{a}{b}\right) \text{ and}$$

$$f\left(\frac{a+c}{b+d}\right) - f\left(\frac{a}{b}\right) = \frac{m}{b(b+d)} = m\left(\frac{a+c}{b+d} - \frac{a}{b}\right).$$

Because we will get all rational numbers between $a/b$ and $c/d$ by repeatedly inserting new terms into the Farey sequence, it follows that

$$f(x) - f\left(\frac{a}{b}\right) = m\left(x - \frac{a}{b}\right), \text{ that is,}$$

$$f(x) = f\left(\frac{a}{b}\right) + m\left(x - \frac{a}{b}\right)$$

for all $x \in [a/b, c/d] \cap \mathbb{Q}$. Because $f$ is continuous, the same equality actually holds for all $x \in [a/b, c/d]$, so $f$ is a linear function on that interval. This proves the claim.

# MAA MATHFEST
August 5-8, 2015

## Join us in Washington, D.C., for our Centennial Celebration.

### 2015 MAA Centennial Lecturers

**Manjul Bhargava**, Princeton University
**Carlos Castillo-Chavez**, Arizona State University
**Jennifer Chayes**, Microsoft Research
**Ingrid Daubechies**, Duke University
**Erik Demaine**, Massachusetts Institute of Technology
**Karen Parshall**, University of Virginia

### Earle Raymond Hedrick Lecturer

**Karen Smith**, University of Michigan

### Other Invited Lecturers

**David Bressoud**, Macalester College
(James R. C. Leitzel Lecture)

**Noam Elkies**, Harvard University
(Pi Mu Epsilon J. Sutherland Frame Lecture)

**Joseph Gallian**, University of Minnesota Duluth
(The Jean Bee Chan and Peter Stanek Lecture for Students)

**Jeffrey Lagarias**, University of Michigan
(AMS-MAA Joint Invited Address)

**Erica Walker**, Columbia University
(AWM-MAA Etta Z. Falconer Lecture)

**Terrence Blackman**, The University of Denver
(NAM David Harold Blackwell Lecture)

**Mathematical sessions run Wednesday morning through Saturday evening (August 5–August 8)**

"0394, part of Contrast Series"
Artwork and Photograph by
Erik Demaine and Martin Demaine

### Register Today

maa.org/mathfest

Joint Meeting with the Canadian Society for History and Philosophy of Mathematics and the British Society for the History of Mathematics.

MAA 100
MATHEMATICAL ASSOCIATION OF AMERICA
CELEBRATING A CENTURY OF ADVANCING MATHEMATICS

MAA 100
MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth St., NW · Washington, DC 20036

# CONTENTS

## ARTICLES



## PROBLEMS AND SOLUTIONS



## REVIEWS



## NEWS AND LETTERS