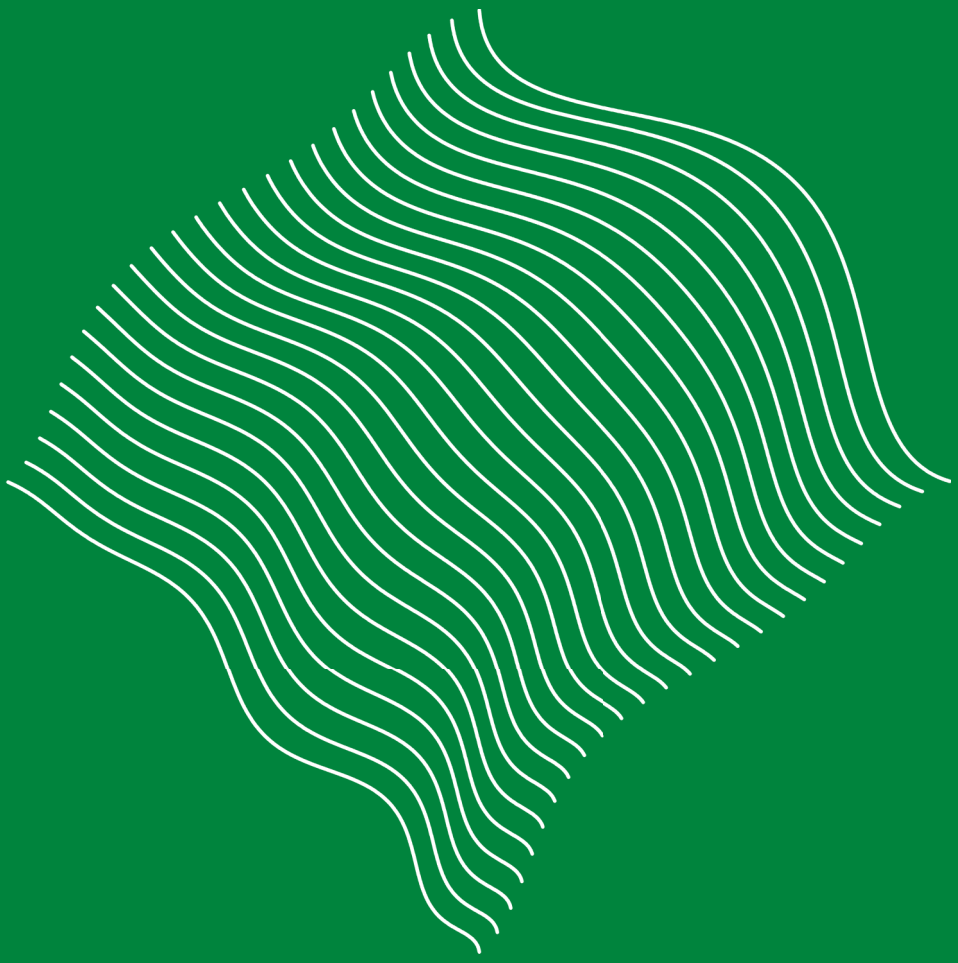
Vol. 90, No. 4, October 2017

MAA

MATHEMATICS MAGAZINE



- From dartboards to cities, different notions of center
- Viewing paintings using geometry, algebra, calculus, and differential geometry
- Paths in the lattice and the best location for a restaurant
- Shazam! Using wavelets to search signals from glucose monitors

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

Wave © 2017 David A. Feimann (*Albion College*). Used by permission.

The article, "Imitating the Shazam App with Wavelets" by Ed Aboufadel was the inspiration for this piece. A wave is formed as one polynomial is transformed into another in 25 steps while being offset by a third polynomial.

Vol. 90, No. 4, October 2017

MAA

MATHEMATICS MAGAZINE

LETTER FROM THE EDITOR

Nicholas R. Baeth, Rhonda McKee, and Loren Luther start off the issue with an examination of how the center of a downtown region changes when the notion of distance changes. The problem generalizes one about a dartboard from the Fiftieth William Lowell Putnam Mathematical Competition from 1989.

Where should one stand to view a painting? This problem is known as the Regiomontanus problem. In their article, Ben Letson and Mark Schwartz unite geometry- and calculus-based solutions to the problem, using parametric curves and results from differential geometry.

Continuing the theme of where it is best to stand to view a painting, Fumiko Futamura and Robert Lehr review geometric and algebraic techniques to determine where one should stand in front of an image in two-point perspective to view it correctly. They derive a simple algebraic formula and a technique that uses slopes on a perspective grid.

In the next article, Nathan Kaplan considers walks on the $\mathbb{Z}^2$ lattice. He counts how many minimal length walks between two points goes through another point to determine where he should open a hypothetical restaurant. He relates the answer to the hypergeometric function and the gamma function before suggesting the same question in higher dimensions.

Modern technology may amaze us, but it is often possible for us to understand how it works, especially when the technology is based on mathematics! Ed Aboufadel explains how the Shazam app works by considering a similar problem. He describes a wavelet-based method to search a database of signals to find other signals that are similar to it. However, the signals are those from continuous glucose monitors used in the management of type-1 diabetes.

Spliced in between the articles are three proofs without words. One by Charles Marion looks at a relationship between triangular sums and perfect quartics. Tom Edgar divides up a square into a series of perfect powers. The final proof without words is by Samuel Moreno. He visually evaluates the sum of the first k odd integers.

Maureen Carroll offers a crossword about geometry and Lai Van Duc Thinh provides another Pinemi puzzle. Problems and Reviews are followed by the announcement of the Allendoerfer award winners: Brian Conrey, James Gabbard, Katie Grant, Andrew Liu, and Kent Morrison for their article "Intransitive Dice" and Vladimir Pozdnyakov and Michael Steele for their article "Buses, Bullies, and Bijections." The issue concludes with the problems, solutions, and results from the 46th United States of America Mathematical Olympiad and 8th United States of America Junior Mathematical Olympiad.

Michael A. Jones, Editor

ARTICLES

The Downtown Problem: Variations on a Putnam Problem

NICHOLAS R. BAETH
LOREN LUTHER
RHONDA MCKEE
University of Central Missouri
Warrensburg, MO 64093
baeth@ucmo.edu
luther@ucmo.edu
mckee@ucmo.edu

Have you ever considered playing a game of darts where the dartboard is square but points are scored based on thrown darts being closer to the center of the board than to the edge? While standing somewhere in a city, have you ever wondered whether you might be closer to the city center than to the edge of town? Or, perhaps at some point you've wondered whether you should reposition a piece of furniture so that it's closer to the middle of the room than to a wall. If you have mathematical tendencies, then you might have wondered what does "closer" really mean. In a city, do we measure distances "as the crow flies," or do we measure distances along perpendicular paths given by streets and sidewalks? If you've pondered such questions, you're not alone. The square dartboard problem appeared in the afternoon session of the Fiftieth William Lowell Putnam Mathematical Competition in 1989 [**4**]. In this article we consider variations of this problem. The general theme will be to compute the relative sizes of regions inside squares and cubes consisting of points closer to the centers of these figures than to the boundaries. Throughout we will consider several different means of measuring distance and pose several open problems for the reader to explore. We begin now with the original problem.

Problem 1. (1989 Putnam Exam B-1) *A dart, thrown at random, hits a square target. Assuming that any two parts of the target of equal area are equally likely to be hit, find the probability that the point hit is nearer to the center than to any edge. Express your answer in the form* $\frac{a\sqrt{b}+c}{d}$*, where* a, b, c, d *are positive integers.*

The solution is given in [**4**], but we encourage the reader to find it themselves before reading any further. We now give a solution, though with a slightly different approach than the official solution. Let D denote the region inside the square C (Read on to see why we are naming a square "C".) consisting of all points that are closer to the center of C than to any of the four edges of C. Then the probability we are after is equal to the ratio $\frac{\mathbf{Area}(\mathsf{D})}{\mathbf{Area}(\mathsf{C})}$. Since rotations and translations are isometries, and since any scaling of C will correspond to a scaling of D by the same proportion, we may assume that C is the square of side length 1 centered at the origin with vertices at $\left(\frac{1}{2}, \frac{1}{2}\right)$, $\left(-\frac{1}{2}, \frac{1}{2}\right)$, $\left(-\frac{1}{2}, -\frac{1}{2}\right)$, and $\left(\frac{1}{2}, -\frac{1}{2}\right)$. Therefore the probability that a dart hits a point closer to the center of C than to any edge of C is simply $\mathbf{Area}(\mathsf{D})$.

Math. Mag. **90** (2017) 243–257. doi:10.4169/math.mag.90.4.243.
MSC: Primary 51K05

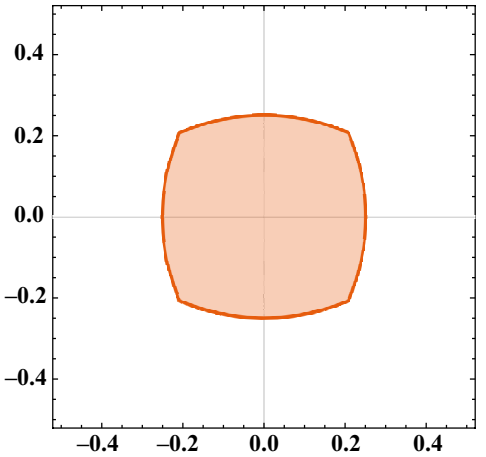


Figure 1 Putnam problem B-1 (1989).

A point (x, y) in the square is closer to the center of C than to the edge of C given by $y = \frac{1}{2}$ if and only if $\sqrt{x^2 + y^2} < \frac{1}{2} - y$, or equivalently, $y < \frac{1}{4} - x^2$. The inequality $y < \frac{1}{4} - x^2$ describes a region bounded by a parabola. Similar inequalities describe the regions within C consisting of all points closer to the center of C than to each of the remaining three sides. Then D is precisely the intersection of these four regions. Figure 1 shows the region D inside the unit square. It is easy to check that pairs of these parabolas intersect at the points $\left(\pm\frac{\sqrt{2}-1}{2}, \pm\frac{\sqrt{2}-1}{2}\right)$ and that D is comprised of a smaller square C′ with vertices at these four points and four parabolic regions appended to each of the four sides of C′. Then

$$\textbf{Area}(\mathsf{D}) = \textbf{Area}(\mathsf{C}') + 4\int_{\frac{1-\sqrt{2}}{2}}^{\frac{\sqrt{2}-1}{2}} \left(\left(\frac{1}{4} - x^2\right) - \frac{\sqrt{2}-1}{2}\right) dx,$$

providing the following solution:

Answer 1. The fraction of the square containing points closer to the center than to any side is $\frac{4\sqrt{2}-5}{3}$, or approximately 21.9%.

But why do we care? Well, suppose that the unit square C considered above represents a city with center at $(0, 0)$. Then perhaps the set of points D, being closer to the center than to the outskirts, represents the downtown region. Then Answer 1 gives the size of the downtown relative to the entire city; that is, the downtown region comprises approximately 21.9% of the city. Having this solution, and a little motivation, we now reframe and generalize the problem. First, it is clear that "closer" depends on the distance function being used. Distance functions are more properly known as metrics. For completeness we give the following well-known definition, which can be found, for example, in [**1**].

Definition. A *metric* on a set X is a function $\mathsf{d} : X \times X \to [0, \infty]$ such that:

1. $\mathsf{d}(x, y) = 0$ if and only if $x = y$,
2. $\mathsf{d}(x, y) = \mathsf{d}(y, x)$ for all $x, y \in X$, and
3. $\mathsf{d}(x, y) \leq \mathsf{d}(x, z) + \mathsf{d}(z, y)$ for all $x, y, z \in X$.

In this case (X, d) is called a *metric space* and d is referred to as a *distance* on X.

The usual Euclidean distance in $\mathbb{R}^2$ and $\mathbb{R}^3$ both satisfy the conditions of Definition and are thus metrics.

In Problem 1 we found the set of points x in the square C with center $(0, 0)$ such that $\mathsf{d}(x, (0, 0)) \leq \mathsf{d}(x, y)$ whenever y is any point on any side of C. To make this language less cumbersome, we introduce some additional terminology.

For an n-dimensional square region (hypercube) C in $\mathbb{R}^n$, we denote by $\partial(\mathsf{C})$, the *boundary* of C. That is, $\partial(\mathsf{C})$ is the union of the $(n-1)$-dimensional faces of C. For example, if C is a square in $\mathbb{R}^2$, then $\partial(\mathsf{C})$ denotes the union of the four line segments that constitute the sides of C, and if C is a cube, then $\partial(\mathsf{C})$ denotes the union of the six square faces of C. In later sections we will have need to compute a distance from a point to a set. We define that concept now. If (X, d) is a metric space and $x \in \mathsf{X}$ and $\mathsf{A} \subseteq \mathsf{X}$, then we define the distance from the point x to the subset A as $\mathsf{d}(x, \mathsf{A}) = \inf\{\mathsf{d}(x, y) : y \in \mathsf{A}\}$.

We now consider the following generalization of Problem 1.

Problem 2. (The Downtown Problem) *Let* d *be a metric on* $\mathbb{R}^n$ *and let* $\mathsf{C}(n, \mathsf{d})$ *(or just* C *if the dimension and metric are clear) denote a generalized square region (hypercube) in* $\mathbb{R}^n$*, which we shall think of as an* n*-dimensional city. We define the downtown region, denoted* $\mathsf{D}(n, \mathsf{d})$ *(or just* D*), of the city to be the set of all points in* $\mathsf{C}(n, \mathsf{d})$ *that are closer (using metric* d*) to the center of* $\mathsf{C}(n, \mathsf{d})$ *than to its boundary* $\partial(\mathsf{C}(n, \mathsf{d}))$*. We wish to find the relative size of the downtown region; that is, to find* $\mathsf{S}(n, \mathsf{d}) = \frac{size[\mathsf{D}(n,\mathsf{d})]}{size[\mathsf{C}(n,\mathsf{d})]}$*, where "size" will mean area in* $\mathbb{R}^2$ *and volume in* $\mathbb{R}^3$*.*

Note that if d is the usual Euclidean metric, denoted d_E, on $\mathbb{R}^2$, then Problem 2 is equivalent to Problem 1. That is, the relative size of the downtown region of our 2-dimensional square city is $\frac{4\sqrt{2}-5}{3} \approx 0.219$. In each of the remaining sections, we (1) give a variation of Problem 2, (2) solve that variation of the problem, and (3) pose one or more other interesting and tractable problems, extending that section's variation. There are, of course, many other variations one might consider, and we encourage the reader to do just that!

The Euclidean metric in three dimensions

In this section we consider Problem 2 for $n = 3$ and the usual Euclidian metric d_E. That is, we pose the following problem for a city similar to the Starbase Yorktown in *Star Trek Beyond* where one can travel in any direction in 3-dimensional space.

Problem 3. *What is the relative size of the downtown region in a futuristic cube-shaped city in* $\mathbb{R}^3$ *if distances are measured using the usual metric? That is, find* $\mathsf{S}(3, \mathsf{d}_E)$*.*

Using the notation of Problem 2, $\mathsf{C} = \mathsf{C}(3, \mathsf{d}_E)$ and $\mathsf{D} = \mathsf{D}(3, \mathsf{d}_E)$. We are looking for $\mathsf{S}(3, \mathsf{d}_E) = \frac{\mathbf{Vol}(\mathsf{D})}{\mathbf{Vol}(\mathsf{C})}$. Since rotations and translations are isometries and since any scaling of C will correspond to a scaling of D by the same proportion, we may assume that C is the unit cube centered at the origin $(0, 0, 0)$, in which case the relative size of the downtown region is $\mathsf{S}(3, \mathsf{d}_E) = \mathbf{Vol}(\mathsf{D})$.

A point (x, y, z) in C is closer to the center of C than to the plane defined by $z = \frac{1}{2}$ if and only if $\sqrt{x^2 + y^2 + z^2} < \frac{1}{2} - z$, or equivalently, $z < \frac{1}{4} - x^2 - y^2$. This inequality describes a solid bounded by a paraboloid. Similar inequalities describe regions within C consisting of all points closer to $(0, 0, 0)$ than to each of the remaining five faces comprising $\partial(\mathsf{C})$. Then D is the intersection of these six solids. This region D is shown in Figure 2.

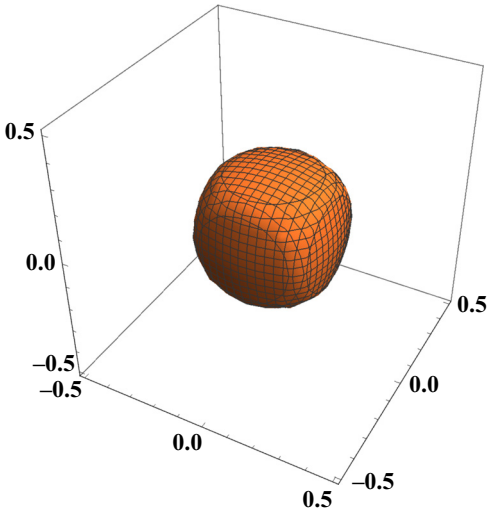


Figure 2 The downtown region for the Euclidean metric on $\mathbb{R}^3$.

To compute the volume of D, we follow an argument similar to that in the solution to Problem 1. Let $\mathbf{a} = (x_1, y_1, z_1)$ denote the point in the first octant where the paraboloids defined by $z = \frac{1}{4} - x^2 - y^2$, $x = \frac{1}{4} - y^2 - z^2$, and $y = \frac{1}{4} - x^2 - z^2$ intersect. The paraboloids defined by $z = \frac{1}{4} - x^2 - y^2$ and $y = \frac{1}{4} - x^2 - z^2$ intersect inside C only on the plane defined by $z = y$. Similarly the paraboloids defined by $z = \frac{1}{4} - x^2 - y^2$ and $x = \frac{1}{4} - y^2 - z^2$ intersect only on the plane defined by $z = x$. Since $\mathbf{a}$ is a point on each of these three paraboloids, $x_1 = y_1 = z_1 = \frac{\sqrt{3}-1}{4}$. Let A denote the solid bounded by $z = \frac{1}{4} - x^2 - y^2$, $z = \frac{\sqrt{3}-1}{4}$, $y = \pm\frac{\sqrt{3}-1}{4}$, and $x = \pm\frac{\sqrt{3}-1}{4}$. The parabolic region A is the portion of D lying above a smaller cube C′ centered at the origin with faces given by $x = \pm\frac{\sqrt{3}-1}{4}$, $y = \pm\frac{\sqrt{3}-1}{4}$, and $z = \pm\frac{\sqrt{3}-1}{4}$. Region A is shown in Figure 3. Thus

$$\mathbf{Vol}(\mathsf{A}) = \int_{\frac{1-\sqrt{3}}{4}}^{\frac{\sqrt{3}-1}{4}} \int_{\frac{1-\sqrt{3}}{4}}^{\frac{\sqrt{3}-1}{4}} \left(\frac{1}{4} - x^2 - y^2 - \frac{\sqrt{3}-1}{4}\right) dx\, dy = \frac{7-4\sqrt{3}}{12}.$$

A similar solid arises for each of the other five faces of the cube C′ and the volume of their union is $\frac{7-4\sqrt{3}}{2}$.

Finally we will consider the points that are in D but neither in the smaller cube C′ nor in any of the six parabolic regions just considered. These points constitute 12 "parabolic wedges." The solid in Figure 4 is half of one of the parabolic wedges that are attached to each of the 12 edges of C′. Since the sides of the larger parabolic region, shown in Figure 3, are flattened and since it has indentations in its four corners, it is not too difficult to imagine how these parabolic wedges from Figure 4 attach to the region from Figure 3.

We again consider the intersection of the paraboloids defined by $z = \frac{1}{4} - x^2 - y^2$ and $y = \frac{1}{4} - x^2 - z^2$. Since $z = y$ on this intersection, $x^2 + (y + \frac{1}{2})^2 = \frac{1}{2}$. That is, these two paraboloids intersect on the cylinder $x^2 + \left(y + \frac{1}{2}\right)^2 = \frac{1}{2}$. Let B be the region bounded by $y = \frac{\sqrt{3}-1}{4}$, $z = \frac{\sqrt{3}-1}{4}$, $z = \frac{1}{4} - x^2 - y^2$, $y = \frac{1}{4} - x^2 - z^2$, and $x = \pm\frac{\sqrt{3}-1}{4}$. Then

$$\mathbf{Vol}(\mathsf{B}) = 2\int_{\frac{\sqrt{3}-1}{4}}^{\frac{1-\sqrt{3}}{4}} \int_{\frac{\sqrt{3}-1}{4}}^{-\frac{1}{2}+\sqrt{\frac{1}{2}-x^2}} \left(\frac{1}{4} - x^2 - y^2 - y\right) dy\, dx$$

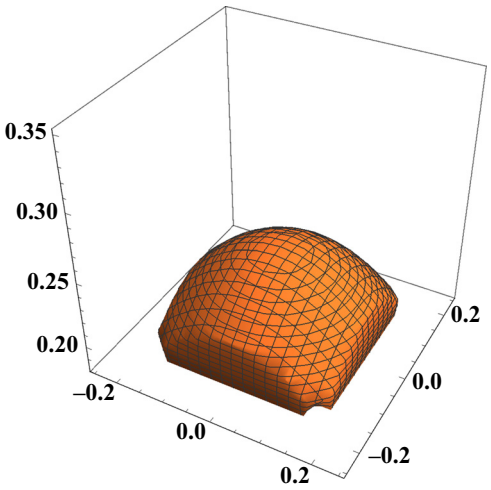


Figure 3 The parabolic region lying above C′.

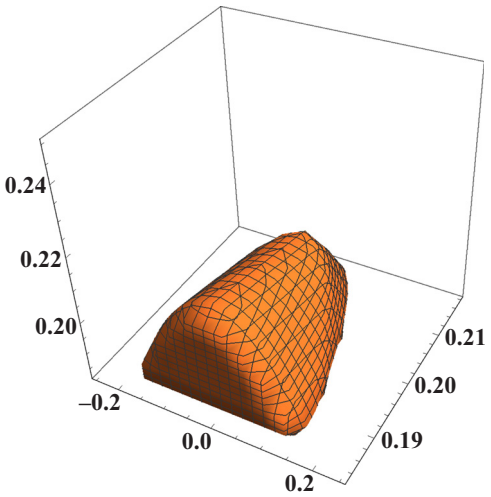


Figure 4 A parabolic wedge.

$$= \frac{\pi - 5 + \left(1 + \sqrt{3}\right)\sqrt{\frac{2+\sqrt{3}}{8}}}{48}.$$

A similar solid arises for each of the other 11 edges of the smaller cube C′ and the volume of their union is $\frac{\pi-5+\left(1+\sqrt{3}\right)\sqrt{\frac{2+\sqrt{3}}{8}}}{4}$. Thus the volume of D is the sum of the volume of the smaller cube, the six parabolic regions, and the 12 "parabolic wedges." In particular,

$$\mathbf{Vol}(\mathsf{D}) = \left(\frac{\sqrt{3}-1}{2}\right)^3 + \frac{7-4\sqrt{3}}{2} + \frac{\pi - 5 + \left(1 + \sqrt{3}\right)\sqrt{\frac{2+\sqrt{3}}{8}}}{4},$$

giving the following solution.

Answer 3. The relative size of the downtown region is

$$\mathsf{S}(3, \mathsf{d}_E) = \frac{\pi + 4 - 5\sqrt{3} + \left(1 + \sqrt{3}\right)\sqrt{\frac{2+\sqrt{3}}{8}}}{4}.$$

That is, the downtown region comprises approximately 8.6% of the 3-dimensional city.

Comparing to Answer 1, we see that, using the Euclidean metrics, the relative size of the downtown region in 3-dimensions is smaller than the relative size of the downtown region in 2-dimensions, that is, $\mathsf{S}(3, \mathsf{d}_E) < \mathsf{S}(2, \mathsf{d}_E)$. Also, $\mathsf{S}(3, \mathsf{d}_E)$ is, curiously unlike $\mathsf{S}(2, \mathsf{d}_E)$, transcendental and not algebraic. See, for example, [**8**], for a proof that π is transcendental.

Question 4. *It seems only natural to generalize even further. What is the relative size, using the Euclidean metric, of the downtown region in an n-dimensional city? As n grows larger, does* $\mathsf{S}(n, \mathsf{d}_E)$ *get smaller? Does* $\lim_{n\to\infty} \mathsf{S}(n, \mathsf{d}_E)$ *exist? If so, what is it? For which n is* $\mathsf{S}(n, \mathsf{d}_E)$ *algebraic?*

The taxicab metric

Dimension Two. Thus far we have used only the usual Euclidean metric to measure the distance between two points in our city. In a city, however, it is likely that we cannot walk or drive in a straight line to our destination. How do such considerations effect the relative size of the downtown? The taxicab metric d_T in $\mathbb{R}^2$ measures the distance between two points much the way a driver of a taxicab would; allowing only paths from point **a** to point **b** by way of "streets" parallel to either the x- or y-axis. More precisely, if $\mathbf{a} = (x_1, y_1)$ and $\mathbf{b} = (x_2, y_2)$, then the *taxicab distance* from **a** to **b** is $\mathsf{d}_T(\mathbf{a}, \mathbf{b}) = |x_2 - x_1| + |y_2 - y_1|$.

It's not difficult to show that d_T satisfies the definition of a metric. Early in the 20th century Herman Minkowski [**7**] defined a family of metrics which result in non-Euclidean geometries. The taxicab metric was among them, although Karl Menger [**6**] was the first, in 1952, to call it the "taxicab" metric. In 1975, Eugene Krause [**5**] further explored the geometry created by this metric. Much more information on this metric can be found at [**10**] and in [**2**].

One interesting difference between the Euclidean and taxicab metrics is the fact that rotations do not preserve taxicab distances. To see this note that $\mathsf{d}_T((0, 0), (1, 0)) = 1$ and that a rotation through an angle of $\pi/4$ radians takes the point (1,0) to the point $(\sqrt{2}/2, \sqrt{2}/2)$. Now, $\mathsf{d}_T((0, 0), (\sqrt{2}/2, \sqrt{2}/2)) = \sqrt{2} \neq 1$. Thus, rotations are not taxicab isometries. On the other hand, it is not difficult to see that both scaling and translations are taxicab isometries. In this section, we consider a square city whose boundaries are parallel to the coordinate axes. We leave the effect of rotation as a problem for the reader to explore, since rotating C will drastically change the shape of D. Also, in this city, area will be calculated using Euclidean geometry. It is only for measuring distances between points within the square that we use the taxicab metric. Thus we investigate the following problem.

Problem 5. *What is the the relative size of the downtown region in a square city in* $\mathbb{R}^2$*, with sides parallel to the coordinate axes, using the taxicab metric? That is, find* $\mathsf{S}(2, \mathsf{d}_T)$ *for this city.*

Since translations and scaling are taxicab isometries, we may assume that C is the square of side length 1 centered at the origin with vertices at $\left(\frac{1}{2}, \frac{1}{2}\right)$, $\left(-\frac{1}{2}, \frac{1}{2}\right)$, $\left(-\frac{1}{2}, -\frac{1}{2}\right)$, and $\left(\frac{1}{2}, -\frac{1}{2}\right)$. Then the relative size of the downtown region will be $\mathsf{S}(2, \mathsf{d}_T) = \mathbf{Area}(\mathsf{D})$.

A point (x, y) is closer to the center of C than the edge of C given by $y = \frac{1}{2}$ if and only if $|x| + |y| < \left|\frac{1}{2} - y\right|$. Equivalently, $y < \frac{1}{4} - \frac{|x|}{2}$, since if $y \leq 0$ the point will be closer to the center than to the edge defined by $y = \frac{1}{2}$. The inequality $y < \frac{1}{4} - \frac{|x|}{2}$ defines a region inside C bounded above by the lines $y = \frac{1}{4} + \frac{x}{2}$ and $y = \frac{1}{4} - \frac{x}{2}$. Similar inequalities describe the regions in C consisting of all the points closer to the center of C than to each of its remaining three sides. Then $\mathsf{D}(2, \mathsf{d}_T)$, the downtown region, is precisely the intersection of these four regions, as illustrated in Figure 5.

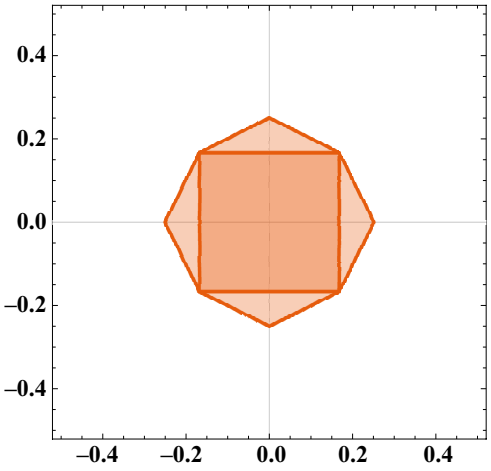


Figure 5 Downtown region for taxicab metric in $\mathbb{R}^2$.

We now compute the area of $\mathsf{D} = \mathsf{D}(2, \mathsf{d}_T)$. From Figure 5 it appears that D is a smaller square C' with triangles appended to each side (and is *not* a regular octagon). This is indeed the case. In the first quadrant, the lines defined by $y = \frac{1}{4} - \frac{|x|}{2}$ and $x = \frac{1}{4} - \frac{|y|}{2}$ intersect at $\left(\frac{1}{6}, \frac{1}{6}\right)$. Similarly, we find that the other pairs of lines intersect at the other four vertices of a smaller square C' with vertices at $\left(\pm\frac{1}{6}, \pm\frac{1}{6}\right)$. Let A denote the triangular region bounded by the three lines defined by $y = \frac{1}{4} - \frac{|x|}{2}$ and $y = \frac{1}{6}$. Then $\mathbf{Area}(A) = \frac{1}{72}$. A similar triangular region arises for each of the other three sides of C'. Thus $\mathbf{Area}(\mathsf{D}) = \frac{1}{9} + 4 \cdot \frac{1}{72}$ and we have our solution.

Answer 5. The relative size of the two-dimensional taxicab downtown is $\mathsf{S}(2, \mathsf{d}_T) = \frac{1}{6}$, or just over 16.6%.

We make two simple observations about this answer compared with the answer to Problem 1. First, since $\frac{1}{6} < \frac{4\sqrt{2}-5}{3}$, the downtown region in a square city is smaller if one is measuring using the taxicab metric than if using the Euclidean metric, i.e., $\mathsf{S}(2, \mathsf{d}_T) < \mathsf{S}(2, \mathsf{d}_E)$. Second, we note that $\mathsf{S}(2, \mathsf{d}_T)$ is not only algebraic, but rational, unlike the irrational $\mathsf{S}(2, \mathsf{d}_E)$.

Question 6. *Taxicab purists would likely balk at the way we are using one metric to compute distances, while using another metric to compute areas (see **[10]**). Consequently, we ask: What would the answer to Problem 5 be if areas are also calculated using the taxicab metric? As mentioned earlier, rotations are not taxicab isometries. How does rotation affect the answer to Problem 5?*

Dimension three We now consider the downtown problem in $\mathbb{R}^3$, using the taxicab metric. Imagine the scene from *Star Wars: Attack of the Clones* where a chase scene takes place on "streets" in space, each parallel to one of the x-, y-, or z-axes. Analogous

to the definition of the taxicab metric in $\mathbb{R}^2$, the taxicab distance between two points $\mathbf{a} = (x_1, y_1, z_1)$ and $\mathbf{b} = (x_2, y_2, z_2)$ in $\mathbb{R}^3$ is $\mathsf{d}_T(\mathbf{a}, \mathbf{b}) = |x_2 - x_1| + |y_2 - y_1| + |z_2 - z_1|$. For reasons similar to those outlined in the two-dimensional case, when we use the term "cube" we mean a cube from Euclidean geometry with faces parallel to the coordinate planes.

Problem 7. *What is the relative size of the downtown region in a cube-shaped city in $\mathbb{R}^3$, with faces parallel to the coordinate planes, using the taxicab metric? That is, find $\mathsf{S}(3, \mathsf{d}_T)$.*

Again, since translations and scalings are taxicab isometries, we may assume that C is the cube centered at the origin such that $\partial(\mathsf{C})$ is the union of the six unit squares lying in the planes $x = \pm\frac{1}{2}$, $y = \pm\frac{1}{2}$, and $z = \pm\frac{1}{2}$. Then $\mathsf{S}(3, \mathsf{d}_T) = \frac{\mathbf{Vol}(\mathsf{D})}{\mathbf{Vol}(\mathsf{C})} = \mathbf{Vol}(\mathsf{D})$ since $\mathbf{Vol}(\mathsf{C}) = 1$.

A point (x, y, z) in C is closer to the center of C than to the plane defined by $z = \frac{1}{2}$ if and only if $|x| + |y| + |z| < \left|\frac{1}{2} - z\right|$. Equivalently, $z < \frac{1}{4} - \frac{1}{2}(|x| + |y|)$ since if $z \leq 0$, the point will be closer to the center than to the edge defined by $z = \frac{1}{2}$. Similar inequalities describe the regions consisting of all points closer to the center of C than to each of the remaining five faces. Then D, the downtown region, is the intersection of these six solids. The solid D is illustrated in Figure 6. We will see that D is comprised of several geometric solids; a smaller cube C' centered at the origin with faces in the planes $x = \pm\frac{1}{8}$, $y = \pm\frac{1}{8}$, and $z = \pm\frac{1}{8}$, and a pyramid-like shape on each of the faces of C'.

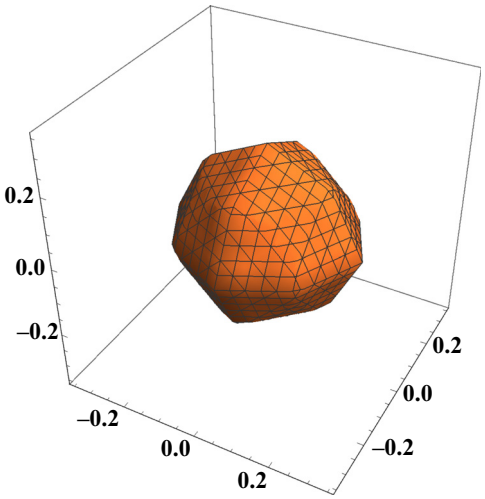


Figure 6 Downtown region for taxicab metric in $\mathbb{R}^3$.

Let $\mathbf{a} = (x_1, y_1, z_1)$ denote the point in the first octant where the planes defined by $z = \frac{1}{4} - \frac{1}{2}(|x| + |y|)$, $y = \frac{1}{4} - \frac{1}{2}(|x| + |z|)$, and $x = \frac{1}{4} - \frac{1}{2}(|y| + |z|)$ intersect. Since $\mathbf{a}$ lies in the first octant, $|x| = x$, $|y| = y$, and $|z| = z$. The planes defined by $z = \frac{1}{4} - \frac{1}{2}(x + y)$ and $y = \frac{1}{4} - \frac{1}{2}(x + z)$ intersect when $y = z$, so $y_1 = z_1$. Similarly, $x_1 = z_1$ and so $x_1 = y_1 = z_1 = \frac{1}{8}$ and $\mathbf{a} = \left(\frac{1}{8}, \frac{1}{8}, \frac{1}{8}\right)$.

Let A be the pyramid-like region bounded by $z = \frac{1}{4} - \frac{1}{2}(|x| + |y|)$, $z = \frac{1}{8}$, $z = \pm x$, and $z = \pm y$. That is, A is the pyramid A' bounded by $z = \frac{1}{4} - \frac{1}{2}(|x| + |y|)$ and $z = \frac{1}{8}$ but missing a smaller pyramid with a triangular base from each of the four corners of its square base. These smaller pyramids are the solids bounded by $z = \frac{1}{4} - \frac{1}{2}(|x| + |y|)$, $z = \frac{1}{8}$, and $z = y$; $z = \frac{1}{4} - \frac{1}{2}(|x| + |y|)$, $z = \frac{1}{8}$, and $z = -y$; $z = \frac{1}{4} - \frac{1}{2}(|x| + |y|)$,

$z = \frac{1}{8}$, and $z = x$; and $z = \frac{1}{4} - \frac{1}{2}(|x| + |y|)$, $z = \frac{1}{8}$, and $z = -x$. Figure 7 shows the region A, which can be seen to be missing a small pyramid from each corner.

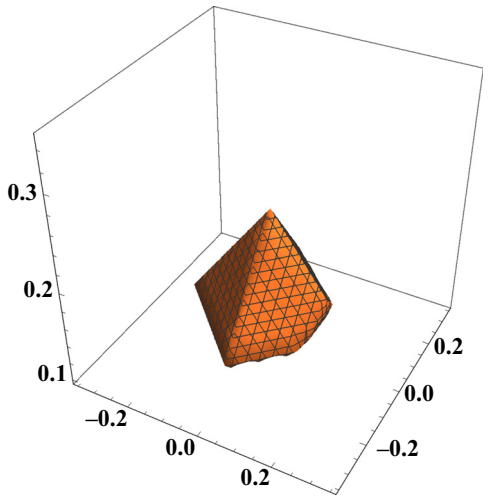


Figure 7 Region A.

To find the volume of A we will first consider the volume of the pyramid A'. The planes defined by $z = \frac{1}{4} - \frac{1}{2}(|x| + |y|)$ intersect the plane $z = \frac{1}{8}$ when $\frac{1}{8} = \frac{1}{4} - \frac{1}{2}(|x| + |y|)$, or equivalently, $|x| + |y| = \frac{1}{4}$. This equation describes the square base of the pyramid A', with vertices at $\left(0, \frac{1}{4}\right)$, $\left(0, -\frac{1}{4}\right)$, $\left(\frac{1}{4}, 0\right)$, and $\left(-\frac{1}{4}, 0\right)$. Thus A' is a pyramid with a square base of area $\frac{1}{8}$ and height $\frac{1}{8}$ and has volume $\frac{1}{192}$.

Each of the smaller pyramids to be removed from A', say the one bounded by $z = \frac{1}{4} - \frac{1}{2}(|x| + |y|)$, $z = \frac{1}{8}$, and $z = y$, has height $\frac{1}{24}$ and a triangular base with area $\frac{1}{64}$, and thus volume $\frac{1}{4608}$. Therefore $\mathbf{Vol}\,\mathsf{A} = \frac{1}{192} - 4\frac{1}{4608} = \frac{5}{1152}$. A similar solid is appended to each of the other five faces of the cube C'. The union of these six solids has volume $\frac{5}{192}$. Then $\mathbf{Vol}(\mathsf{D}) = \left(\frac{1}{4}\right)^3 + \frac{5}{192}$ giving us the following solution.

Answer 7. The relative size of the three-dimensional taxicab downtown is $\frac{1}{24}$, or just over 4.1%.

As with the Euclidean metric, the relative size of the downtown region in 3-dimensions is smaller than the relative size of the downtown region in 2-dimensions, that is, $\mathsf{S}(3, \mathsf{d}_T) < \mathsf{S}(2, \mathsf{d}_T)$. Also, as in the two-dimensional setting, $\mathsf{S}(3, \mathsf{d}_T) < \mathsf{S}(3, \mathsf{d}_E)$.

Question 8. *If* C *is an n-dimensional hypercube-shaped city, what is* $\mathsf{S}(n, \mathsf{d}_T)$*? What would* $\mathsf{S}(n, \mathsf{d}_T)$ *be if volumes are computed using only the taxicab metric? How does rotation affect the answer?*

An interesting bounded metric

We now consider Problem 2 with the following natural twist. Let δ be some fixed number and suppose that you have deemed any (Euclidean) distance larger than δ too far to walk. Consequently, when traveling between two points that are Euclidean distance more than δ apart, you will take some alternate transportation—car, bus, subway, etc. But, since you don't want your friends to label you as lazy, you have decided that

whenever taking one of these alternate modes of transportation, you will first take a walk of Euclidian distance exactly δ. In other words, every trip you take anywhere in the city will require walking a distance no more than δ, and that if you travel a total distance δ or more, you will walk exactly a distance of δ. We now consider the downtown region with respect to this new "walking metric." Let C be a square in $\mathbb{R}^2$ with side length s and let δ be a real number satisfying $0 < \delta < \frac{1}{2}s$. Define the delta (walking) metric on $\mathbb{R}^2$ by $\mathsf{d}_\delta(\mathbf{a}, \mathbf{b}) = \min\{\mathsf{d}_E(\mathbf{a}, \mathbf{b}), \delta\}$. Note that $\mathsf{d}_\delta(\mathbf{a}, \mathbf{b}) \leq \delta$ for all points $\mathbf{a}$ and $\mathbf{b}$. The restriction $\delta < \frac{1}{2}s$ is in place simply because any point in C is no more than $\frac{1}{2}s$ away from some side of C when using d_E, so if $\delta \geq \frac{1}{2}s$, the size of the downtown region can be found using the usual Euclidean metric as in Problem 1.

In mathematics it is sometimes convenient to work with metrics such as the one we use here, that are bounded by some fixed number. Of course all of the metrics we have considered are bounded when restricted to a bounded n-dimensional hypercube, but d_δ is bounded on all of $\mathbb{R}^2$. This particular metric is of interest to topologists, since it can be used to prove that every metric space is topologically equivalent (homeomorphic) to a bounded metric space (see [**1**]). We now consider the following problem.

Problem 9. *What is the relative size of the downtown region in a square city in* $\mathbb{R}^2$ *using the delta metric. That is, find* $\mathsf{S}(2, \mathsf{d}_\delta)$.

As before, considering isometries and scaling, we may assume C is the unit square centered at the origin with vertices $\left(\frac{1}{2}, \frac{1}{2}\right)$, $\left(-\frac{1}{2}, \frac{1}{2}\right)$, $\left(-\frac{1}{2}, -\frac{1}{2}\right)$, and $\left(\frac{1}{2}, -\frac{1}{2}\right)$, and that $0 < \delta < \frac{1}{2}$.

By definition of the delta metric, the distance $\mathsf{d}_\delta((0, 0), (x, y))$, from the origin to the point (x, y), is less than δ if and only if the Euclidean distance is less than δ, which is true if and only if $x^2 + y^2 < \delta^2$. Let A denote the circular region defined by $x^2 + y^2 < \delta^2$ so that $\mathsf{d}_\delta((0, 0), (x, y)) < \delta$ if and only if (x, y) is in A. Note that if (x, y) is outside A, $\mathsf{d}_\delta((0, 0), (x, y)) = \delta$.

Let B_1 be the set of boundary points of C defined by $y = \frac{1}{2}$, $\frac{-1}{2} \leq x \leq \frac{1}{2}$. By definition of the delta metric, $\mathsf{d}_\delta((x, y), \mathsf{B}_1) < \delta$ if and only if $\mathsf{d}_E((x, y), \mathsf{B}_1) < \delta$, which is true if and only if $\frac{1}{2} - y < \delta$, or equivalently $y > \frac{1}{2} - \delta$. Since (x, y) is in C, $y < \frac{1}{2}$. Thus, $\mathsf{d}_\delta((x, y), \mathsf{B}_1) < \delta$ if and only if (x, y) is in the strip defined by $\frac{-1}{2} \leq x \leq \frac{1}{2}$ and $\frac{1}{2} - \delta < y \leq \frac{1}{2}$. Similar strips along the other three sides of C describe regions in which points are closer to an edge of C than to the center of C. Let T denote the union of these strips. Then T is the frame shaped region between the square C and the smaller square C', centered at the origin with sides parallel to the axes and intersecting each axis at $\pm\left(\frac{1}{2} - \delta\right)$. Note that if (x, y) is inside C', then the delta distance from (x, y) to any edge is δ. The regions A, T, and C' are shown in Figure 8 with three different values of δ. So that this new metric can be compared with the usual Euclidean metric from the original Problem 1, each of the images in Figure 8 also shows the region from Figure 1.

We now consider three cases.

Case 1: $0 < \delta \leq \frac{1}{4}$. In this case $\delta \leq \frac{1}{4} \leq \frac{1}{2} - \delta$, and thus the circular region A does not intersect the frame-shaped region T (see the top image in Figure 8). Points in A are less than δ units away from the center of the square, but are exactly δ units from every edge. On the other hand, if (x, y) is outside A, then $\mathsf{d}_\delta((0, 0), (x, y)) = \delta$, while the distance from (x, y) to some edge is less than or equal to δ. Thus only the points in A are closer to the center than to any edge, and $\mathbf{Area}(\mathsf{D}) = \mathbf{Area}(\mathsf{A}) = \pi\delta^2$. It is worth noting that points that lie between the circular region A and the frame-shaped region T are equidistant from the center and every edge, that distance being δ.

Case 2: $\frac{1}{4} < \delta \leq \frac{2-\sqrt{2}}{2}$. The smaller square C' will be inscribed in the circle $x^2 + y^2 = \delta^2$ when half the diameter of C' is equal to the radius of A. This happens when $\sqrt{2}(\frac{1}{2} - \delta) = \delta$, that is, when $\delta = \frac{2-\sqrt{2}}{2}$. Therefore, in this case, the cir-

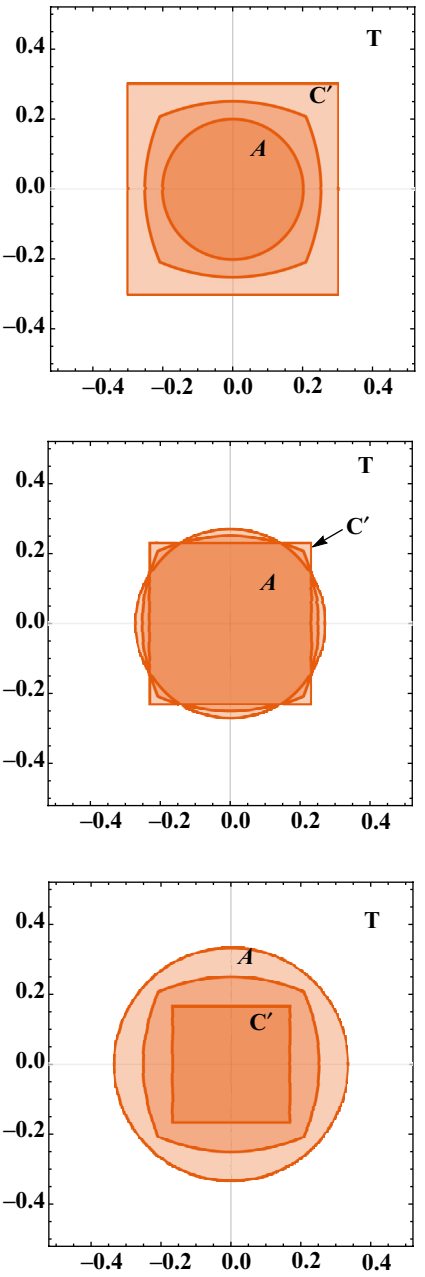


Figure 8 The downtown region using the delta metric with δ values of $\frac{1}{5}$, $\frac{27}{100}$, and $\frac{1}{3}$.

cular region A overlaps the frame-shaped region T, but A does not contain all of the smaller square C′. See the middle image in Figure 8. Note that the circle $x^2 + y^2 = \delta^2$ intersects the line $y = \frac{1}{2} - \delta$ when $x = \pm\sqrt{\delta - \frac{1}{4}}$.

Points that are in A\T are closer to the center of C than to any edge of C. Points that are in T\A are closer to some edge of C than to the center of C. Points that are in neither A nor T are equidistant from the center and every edge, that distance being δ.

For points in A ∩ T, the delta distance to either the center of C or the nearest side of C will be measured using the Euclidean distance. From the solution to Problem 1 we know that the region that consists of points that are closer, using the Euclidean dis-

tance, to the center of square C than to its upper edge is bounded above by the parabola $y = \frac{1}{4} - x^2$. This parabola intersects the line $y = \frac{1}{2} - \delta$, the upper edge of square C', when $x = \pm\sqrt{\delta - \frac{1}{4}}$, the same points at which the circular region A intersects that edge. Thus points in the parabolic region bounded below by the line $y = \frac{1}{2}$ and above by the parabola $y = \frac{1}{4} - x^2$ are closer to the center of C than to its top edge. Let R_1 denote this region and let R be the union of R_1 and three similar regions corresponding to the other three sides of C.

Then $\mathbf{Area}(\mathsf{D}) = \mathbf{Area}((\mathsf{A} \cap (\mathsf{C} \setminus \mathsf{T})) \cup \mathsf{R}) = \mathbf{Area}((\mathsf{A} \cap \mathsf{C}') \cup \mathsf{R})$. For the purposes of computation, note that this is simply the region A with the regions bounded by the circle and each of the four parabolas removed. Therefore

$$\begin{aligned}\mathbf{Area}(\mathsf{D}) &= \pi\delta^2 - 8\int_0^{\sqrt{\delta-\frac{1}{4}}} \left(\sqrt{\delta^2 - x^2} - \left(\frac{1}{4} - x^2\right)\right) dx \\ &= \pi\delta^2 - 4\delta^2 \sin^{-1}\left(\frac{\sqrt{\delta - \frac{1}{4}}}{\delta}\right) + \sqrt{\delta - \frac{1}{4}}\left(\frac{2}{3} + \frac{4}{3}\delta\right).\end{aligned}$$

Note that if $\delta = \frac{1}{4}$, this gives the same result as for $\delta = \frac{1}{4}$ in case 1.

Case 3: $\frac{2-\sqrt{2}}{2} < \delta \leq \frac{1}{2}$. Recall from case 2 that when $\delta = \frac{2-\sqrt{2}}{2}$, the smaller square C' is inscribed in the circle $x^2 + y^2 = \delta^2$. Therefore, when $\delta > \frac{2-\sqrt{2}}{2}$, the square C' is completely contained in the circular region A. From the solution to Problem 1, we know that the parabolas $y = \frac{1}{4} - x^2$ and $x = \frac{1}{4} - y^2$ intersect at the points $\left(\frac{\sqrt{2}-1}{2}, \frac{\sqrt{2}-1}{2}\right)$. The distance from this point to the origin is $\frac{2-\sqrt{2}}{2}$ and so when $\delta > \frac{2-\sqrt{2}}{2}$, the parabolic region shown in Figure 1 is also contained in the circular region A. See the bottom image in Figure 8.

Points that are in $\mathsf{A}\setminus\mathsf{T}$ are closer to the center of C than to any edge of C since the distance to the origin would be less than δ and the distance to any edge would be exactly δ. In this case, $\mathsf{A}\setminus\mathsf{T}$ consists of all points in the smaller square C'. By a similar argument, points that are in $\mathsf{T}\setminus\mathsf{A}$ are closer to some edge of C than to the origin. For points in $\mathsf{A} \cap \mathsf{T}$, the delta distance from the point to the center of C and the distance from the point to some edge of C are equal to the corresponding Euclidean distance. Therefore, from the solution to Problem 1, we know that points in $\mathsf{A} \cap \mathsf{T}$ will be closer to the center of A than to any side of A if and only if they lie in the region bounded by the four parabolas $y = \pm\left(\frac{1}{4} - x^2\right)$, $x = \pm\left(\frac{1}{4} - y^2\right)$. Since this region contains the smaller square C', the area we are looking for is the same as the area in Problem 1, namely $\mathbf{Area}(\mathsf{D}) = \frac{4\sqrt{2}-5}{3}$.

Summarizing, we have the following solution.

Answer 9. The relative size of the downtown region in the two-dimensional city C using the δ-metric depends greatly on the value of δ. In particular,

$$\mathsf{S}(2, \mathsf{d}_\delta) = \begin{cases} \pi\delta^2, & 0 < \delta \leq \frac{1}{4} \\ \pi\delta^2 - 4\delta^2 \sin^{-1}\left(\frac{\sqrt{\delta-\frac{1}{4}}}{\delta}\right) + \sqrt{\delta - \frac{1}{4}}\left(\frac{2}{3} + \frac{4}{3}\delta\right), & \frac{1}{4} \leq \delta \leq \frac{2-\sqrt{2}}{2} \\ \frac{4\sqrt{2}-5}{3}, & \frac{2-\sqrt{2}}{2} < \delta \leq \frac{1}{2} \end{cases}.$$

As with the Euclidean and taxicab metrics, one can generalize Problem 9 to three dimensions. The details can be worked out as in the two-dimensional solution, again with three cases determined by the value of δ. We state the problem and answer here, leaving the details to the reader.

Problem 10. *Let* C *be a cube in* $\mathbb{R}^3$ *with side length* 1 *and let* δ *be a real number satisfying* $0 < \delta < \frac{1}{2}$. *Define the delta metric on* $\mathbb{R}^3$ *by* $\mathsf{d}_\delta(\mathbf{a}, \mathbf{b}) = \min\{\mathsf{d}_E(\mathbf{a}, \mathbf{b}), \delta\}$. *What is the relative size of the downtown region in* C *using the delta metric. That is, find* $\mathsf{S}(3, \mathsf{d}_\delta)$.

Answer 10. The relative size of the downtown region in the three-dimensional city C using the δ-metric is given by,

$$\mathsf{S}(3, \mathsf{d}_\delta) = \begin{cases} \frac{4}{3}\pi\delta^3, & 0 < \delta \le \frac{1}{4} \\ \frac{4}{3}\pi\delta^3 + \pi\left[-\frac{1}{2}(4\delta^2 - 2\delta + 1)(4\delta - 1) + 3\delta\left(\delta - \frac{1}{4}\right)\right], & \frac{1}{4} < \delta \le \frac{3-\sqrt{3}}{4} \\ \frac{\pi + 4 - 5\sqrt{3} + (1+\sqrt{3})\sqrt{\frac{2+\sqrt{3}}{8}}}{4}, & \frac{3-\sqrt{3}}{4} < \delta \le \frac{1}{2} \end{cases}.$$

Not all metrics that are bounded by some $\delta < 1$ give new answers to Problem 2. For example, if a, b, c are positive real numbers, then $\mathsf{d}(\mathbf{x}, \mathbf{y}) = \frac{a\,\mathsf{d}_E(\mathbf{x},\mathbf{y})}{b + c\,\mathsf{d}_E(\mathbf{x},\mathbf{y})}$ is a metric on $\mathbb{R}^2$ that is bounded, on the unit square, by $\frac{a}{b+c}$. However, no matter how small $\frac{a}{b+c}$ is, using this metric gives the same answer to Problem 2 as when using the Euclidean metric. We conclude this section with a couple of questions.

Question 11. *Do other bounded metrics give interesting solutions to Problem 2? What is the relative size of the downtown region in a hypercube-shaped city in* $\mathbb{R}^n$ *using the delta metric? That is, find* $\mathsf{S}(n, \mathsf{d}_\delta)$.

A discrete variation

Suppose that our two-dimensional square city is overlaid with an $r \times r$ grid, as most cities are divided into blocks, and that a business is situated at each lattice point, that is, intersection. How many businesses are in the downtown region? This discrete version of Problem 2 is analogous to a famous problem of Guass (see, for example [**3**]) in which he asks how many integer lattice points are contained within a circle of radius n.

Problem 12. *Let* r *be a positive real number and let* C *be an* $r \times r$ *square in* $\mathbb{R}^2$ *centered at the origin with edges defined by* $y = \pm\frac{r}{2}$ *and* $x = \pm\frac{r}{2}$. *If a business is situated at every integer lattice point and distances are measured with the Euclidean metric, how many businesses are in the downtown region?*

Arguing as in Problem 1, one easily sees that the answer to Problem 12 is the number of lattice points that are in the region D bounded by the four parabolas defined by the inequalities $y < \frac{r}{4} - \frac{x^2}{r}$, $y > -\frac{r}{4} + \frac{x^2}{r}$, $x < \frac{r}{4} - \frac{y^2}{r}$, and $x > -\frac{r}{4} + \frac{y^2}{r}$. Also as in Problem 1, this region D comprises a smaller square C' with vertices at $\left(\pm\frac{r(\sqrt{2}-1)}{2}, \pm\frac{r(\sqrt{2}-1)}{2}\right)$ along with four parabolic regions, one appended to each side of C'. We now count the integer lattice points within D. For convenience, set $\alpha = \frac{r(\sqrt{2}-1)}{2}$.

We first consider the number of integer lattice points within C'. Fix an integer a with $-\alpha \le a \le \alpha$. On the portion of the line $x = a$ with y greater than 0 and less than α, there are precisely $\lfloor\alpha\rfloor$ points (a, y) with y an integer. There are the same number of points (a, y) with y an integer along the line $x = a$ with y-value between $-\alpha$ and 0. Thus, including the origin, there are $2\lfloor\alpha\rfloor + 1$ integer lattice points within the small square C' and on the line $x = a$. Since a was an arbitrarily chosen integer with $-\alpha \le a \le \alpha$, the number of integer lattice points inside C' is $(2\lfloor\alpha\rfloor + 1)^2$.

We now consider the integer lattice points in the region bounded by $y = \alpha$ and $y = \frac{r}{4} - \frac{x^2}{r}$, that is, in the portion of D above C'. Note that the line defined by $y =$

α intersects the parabola defined by $y = \frac{r}{4} - \frac{x^2}{r}$ at the points $(\pm\alpha, \pm\alpha)$. The set of integers in the interval $[-\alpha, \alpha]$ is

$$I = \{\lceil -\alpha \rceil, \lceil -\alpha \rceil + 1, \lceil -\alpha \rceil + 2, \ldots, 0, \ldots, \lfloor \alpha \rfloor - 2, \lfloor \alpha \rfloor - 1, \lfloor \alpha \rfloor\}.$$

For each $n \in I$, if $\frac{r}{4} - \frac{n^2}{r}$ is an integer, then $\left\lceil \frac{r}{4} - \frac{n^2}{r} \right\rceil = \frac{r}{4} - \frac{n^2}{r}$ and the point $\left(\frac{r}{4} - \frac{n^2}{r}, n\right)$ is equidistant from the center and the edge defined by $y = \frac{1}{2}$ and should not be included in our count (since we are only interested in points *closer* to the center than to any edge). Thus, for each $n \in I$, the number of integer lattice points between $y = \alpha$ and $y = \frac{r}{4} - \frac{x^2}{r}$ on the line defined by $x = n$ is $\left\lceil \frac{r}{4} - \frac{n^2}{r} \right\rceil - 1 - \lfloor \alpha \rfloor$. Thus the number of integer points in the region bounded by $y = \alpha$ and $y = \frac{r}{4} - \frac{x^2}{r}$ is $\sum_{n=\lceil -\alpha \rceil}^{\lfloor \alpha \rfloor} \left(\left\lceil \frac{r}{4} - \frac{n^2}{r} \right\rceil - 1 - \lfloor \alpha \rfloor \right)$. Identical counts are obtained for each of the remaining three parabolic regions. Now summing the number of integer lattice points within C' and within each of the four parabolic regions, we obtain the following solution.

Answer 12. In an $r \times r$ city divided into blocks, if there is a business at each intersection, there are precisely $(2\lfloor \alpha \rfloor + 1)^2 + 4 \sum_{n=\lceil -\alpha \rceil}^{\lfloor \alpha \rfloor} \left(\left\lceil \frac{r}{4} - \frac{n^2}{r} \right\rceil - 1 - \lfloor \alpha \rfloor \right)$ businesses within the downtown region.

We give a list of the first 25 terms in the sequence given by the number of points inside a square of integer side length r that are closer to the center than to any edge

1, 1, 1, 1, 9, 9, 9, 9, 21, 25, 25, 25, 37, 45, 49, 49, 69, 69, 77, 81, 101, 109, 117, 117, 141.

As with Gauss's circle problem, we cannot obtain a closed form for the number of integer lattice points closer to the center of a square than to any side. In fact, finding the order of the error term to a polynomial approximation to Gauss's problem is of great interest (cf. [**3**]). However, we can easily obtain a closed form asymptotically. By placing a unit square (with sides parallel to the coordinate axes) around each integer lattice point in D, we see that the number of integer lattice points closer to the center of C than to any side is approximated by the area of D. By appropriately scaling the calculations from Problem 1, we see that the area of region D is $r^2 \left(\frac{4\sqrt{2}-5}{3} \right)$. Since the area of the square C is r^2, as r grows larger, the ratio of the number of integer lattice points in D to the number in C approaches $\frac{4\sqrt{2}-5}{3}$. Already at $r = 25$, the ratio is $\frac{141}{25^2} = 0.22564$ while $\frac{4\sqrt{2}-5}{3} \approx 0.21895$.

Question 13. *Gauss's circle problem can be generalized to higher dimensions, as can Problem 12. If* C *is an n-dimensional hypercube-shaped city with side length r and downtown region* D*, how many businesses are in* D*? If distances are measured with a different metric, how many businesses are in* D*?*

REFERENCES

1. S. Davis, *Topology*. McGraw-Hill Science/Engineering/Math, New York, 2004.
2. L. Golland, Karl Menger and taxicab geometry, *Math. Mag.* **63** no. 5 (1990) 326–327.
3. E. Grosswald, *Representations of Integers as Sums of Squares*. Springer-Verlag, Berlin, 1985.
4. L. F. Klosinski, G. L. Alexanderson, L. C. Larson, The Fiftieth William Lowell Putnam Mathematical Competition, *Amer. Math. Monthly* **98** no. 4 (1991) 319–327.
5. E. F. Krause, *Taxicab Geometry: An Adventure in Non-Euclidean Geometry*. Dover, New York, 1986.
6. K. Menger, *You Will Like Geometry, Guidebook of the Illinois Institute of Technology Geometry Exhibit*. Museum of Science and Industry, Chicago, IL, 1952.

7. H. Minkowski, *Gesammelte Abhandlungen*. Chelsea Publishing, New York, 1967.
8. I. Niven, The Transcendence of π, *Amer. Math. Monthly* **46** no. 8 (1939) 469–471.
9. B. E. Reynolds, Taxicab geometry, *The Pi Mu Epsilon J.* **7** no. 2 (1980) 77–88.
10. K. P. Thompson, `TaxicabGeometry.net`.

Summary. The following problem appeared in the afternoon session of the Fiftieth William Lowell Putnam Mathematical Competition in 1989. *A dart, thrown at random, hits a square target. Assuming that any two parts of the target of equal area are equally likely to be hit, find the probability that the point hit is nearer to the center than to any edge.* In general, which points are closer to the center of a square (or cube) than to any of the four edges (six faces)? And what do we mean by "closer"? Using a variety of distance functions (metrics), we will investigate these questions and pose several additional open problems for the reader to explore.

NICHOLAS R. BAETH (MR Author ID: 802867) received a B.S. in mathematics and a B.A. in computer science in 2000 from Pacific Lutheran University. He earned his Ph.D. in mathematics from the University of Nebraska-Lincoln in 2005. Since then he has been a member of the mathematics faculty at the University of Central Missouri where he teaches a myriad of courses from freshman-level nonmajor courses through graduate-level courses in algebra. In 2013 he held a post in Graz Austria as the NAWI-Graz Fulbright Visiting Professor in the Natural Sciences. His mathematical interests are algebraic in nature, and much of his research is in the area of factorization theory. His nonmathematical interests are many, and include chasing after his toddler son.

LOREN LUTHER (MR Author ID: 118871) received his M.S. in Mathematics in 2016. His thesis, directed by Nicholas Baeth and Rhonda McKee, included much of the content of the current article. He is currently teaching at the University of Central Missouri and at State Fair Community College.

RHONDA MCKEE (MR Author ID: 294871) earned a Ph.D. in topology from Missouri University of Science and Technology in 1989 and is a professor of mathematics at the University of Central Missouri. She is currently serving as the president of Kappa Mu Epsilon, national mathematics honor society. During summer breaks, she enjoys bicycling on the back roads of Missouri.

PINEMI PUZZLE

	4		8	7	8		6	6	
2	6	8				10			7
3									8
			6	8		8	9		
		11		8	8			7	6
9	11			9			8	6	
		13			7	8			5
8			10			7		8	
	10		8				9		
		2	5		8				4

How to play. Place one jamb (|), two jambs (||), or three jambs (|||) in each empty cell. The numbers indicate how many jambs there are in the surrounding cells—including diagonally adjacent cells. Each row and each column has 10 jambs. Note that no jambs can be placed in any cell that contains a number.

The solution is on page 266.

—contributed by Lai Van Duc Thinh
Vietnam; fibona2cis@gmail.com

Proof Without Words: Triangular Sums and Perfect Quartics

CHARLES F. MARION
Yorktown Heights, NY 10598
charliemath@optionline.net

$$1 + 15 = 4^2 = 2^4$$

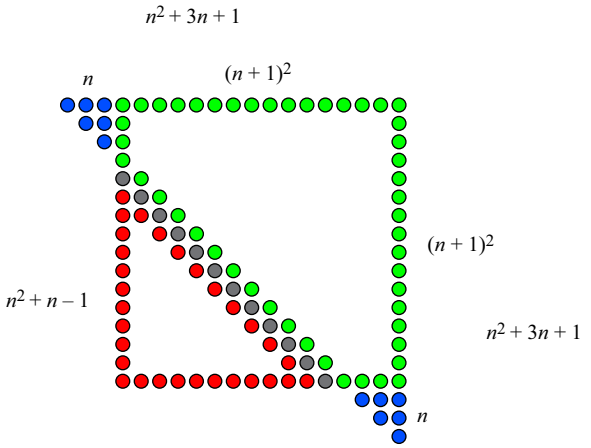
$$15 + 66 = 9^2 = 3^4$$

$$66 + 190 = 16^2 = 4^4$$

$$\vdots$$

$$T(n) = 1 + 2 + \ldots + n \Rightarrow T(n^2 + n - 1) + T(n^2 + 3n + 1) = \left((n+1)^2\right)^2 = (n+1)^4$$

For example, for n = 3,



Summary. It is well known that the sum of two consecutive terms from the sequence of triangular numbers is a perfect square. We show that the sum of two consecutive terms from a subsequence of that sequence is a perfect quartic.

CHARLES F. MARION (MR Author ID: 1211614) taught mathematics at Lakeland High School (Shrub Oak, NY) for 32 years. In retirement, he has enjoyed searching for and writing about undiscovered patterns in recreational number theory and sharing his enthusiasm for mathematics with his grandchildren, Allison, Jack, Joseph, Courtney and Sam.

Math. Mag. **90** (2017) 258. doi:10.4169/math.mag.90.4.258.
MSC: Primary 11A99, Secondary 11H99

The Regiomontanus Problem

BENJAMIN LETSON
University of Pittsburgh
Pittsburgh, PA 15260
bgl14@pitt.edu

MARK SCHWARTZ
Ohio Wesleyan University
Delaware, OH 43015
mdschwar@owu.edu

Do you know the name "Regiomontanus"? You undoubtedly know his problem. What is the optimal place to stand to view a painting? The usual assumptions are in effect, namely, the painting hangs on a wall above eye level, and optimal refers to maximizing the angle from the observer's eye. The problem is ubiquitous in modern calculus books, but Regiomontanus, born Johann Müller (1436-76) in Königsberg, formulated the problem in 1471, some 200 years before the discovery of calculus. Not surprisingly then, the first known solution is geometric and is shown in Figure 1. A circle is drawn passing through the top and bottom of the painting, tangent to the horizontal line at eye level. The point of tangency is the optimal point. That the so-constructed angle is maximal follows from the intersecting secants theorem in geometry. It isn't known if Regiomontanus found the solution. A careful construction is a nontrivial Appolonian problem. See [**3**], [**5**], and [**6**] for additional background. Modern calculus students (and mathematicians) are routinely asked to solve the Regiomontanus problem. Whatever their solution, if it involves a derivative, critical points, etc., it reveals nothing of the original geometry. What follows is an approach that uses ideas from differential geometry to connect the calculus-based solution to the geometric.

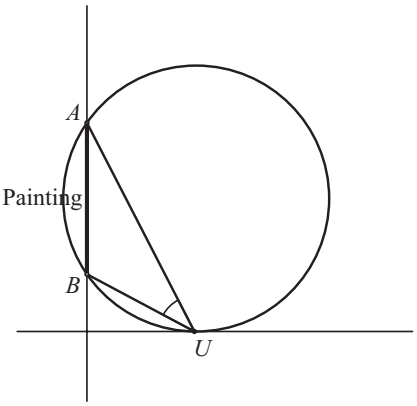


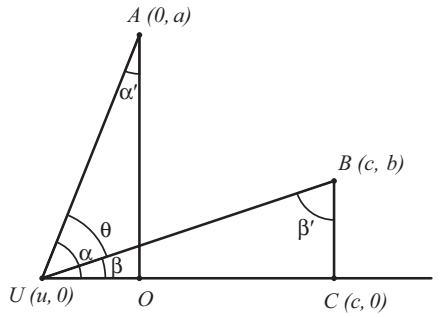
Figure 1 Geometric solution of the Regiomontanus problem.

A more general problem The Regiomontanous problem is a special case of a more general angle-maximization problem. Consider two points A and B, where $A = (0, a)$ is on the positive y-axis and $B = (c, b)$ where $c \geq 0$ and $b \leq a$. The latter condition is

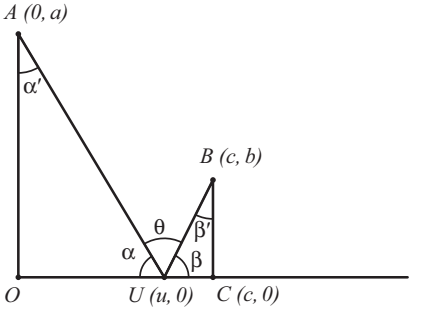
Math. Mag. **90** (2017) 259–266. doi:10.4169/math.mag.90.4.259.
MSC: Primary 26A06, Secondary 51M16; 53A04.

without loss of generality, for we could translate B to the y-axis and reflect the figure through the y-axis to place the points in the desired configuration. The problem is to find the point $U = (u, 0)$ on the x-axis such that $\theta = \angle AUB$ is maximal. A specific case of the angle-maximization problem appears in [**8**], with values $a = 5$, $b = 2$, $c = 3$, and u restricted to $[0, 3]$. Unlike that problem, we seek the global maximum of θ over the entire real line.

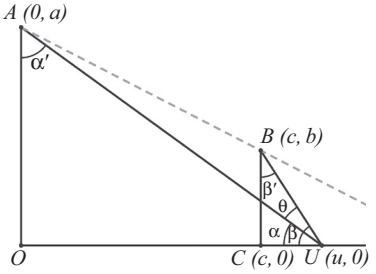
To start, we need an expression for θ. There are four cases depending on the location of U, though a single calculation covers them all. Throughout, the points A and B are fixed. Figure 2 illustrates the four possibilities.



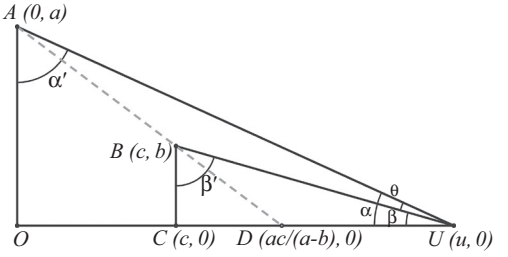
Case 1 Angle θ when U lies to the left of O (origin).



Case 2 Angle θ when U lies between O and C.



Case 3 Angle θ when U lies between C and the x-intercept, D , of the line AB.



Case 4 Angle θ when U lies to the right of D.

Figure 2 The four cases.

With the angles labeled as shown in the figures, we seek to maximize θ where, proceeding sequentially through the cases,

$$\theta = \begin{cases} \alpha - \beta = \left(\frac{\pi}{2} - \alpha'\right) - \left(\frac{\pi}{2} - \beta'\right) = \beta' - \alpha' = \tan^{-1}\left(\frac{c-u}{b}\right) - \tan^{-1}\left(\frac{-u}{a}\right) & \text{if } u \le 0, \\ \pi - \alpha - \beta = \left(\frac{\pi}{2} - \alpha\right) + \left(\frac{\pi}{2} - \beta\right) = \alpha' + \beta' = \tan^{-1}\left(\frac{u}{a}\right) + \tan^{-1}\left(\frac{c-u}{b}\right) & \text{if } 0 \le u \le c, \\ \beta - \alpha = \left(\frac{\pi}{2} - \beta'\right) - \left(\frac{\pi}{2} - \alpha'\right) = \alpha' - \beta' = \tan^{-1}\left(\frac{u}{a}\right) - \tan^{-1}\left(\frac{u-c}{b}\right) & \text{if } c \le u \le \frac{ac}{a-b}, \\ \alpha - \beta = \left(\frac{\pi}{2} - \alpha'\right) - \left(\frac{\pi}{2} - \beta'\right) = \beta' - \alpha' = \tan^{-1}\left(\frac{u-c}{b}\right) - \tan^{-1}\left(\frac{u}{a}\right) & \text{if } \frac{ac}{a-b} \le u < \infty. \end{cases}$$

Note, in every case, the expression for θ is the same, with the exception of Case 4, where θ is the negative of the others. As we are interested only in the magnitude of θ, this minor detail has no bearing on the calculations and solution that follow.

For definiteness, we consider Case 2. Computing the derivative of θ with respect to u we get

$$\theta' = \frac{1}{1 + \left(\frac{u}{a}\right)^2}\left(\frac{1}{a}\right) + \frac{1}{1 + \left(\frac{c-u}{b}\right)^2}\left(-\frac{1}{b}\right) = \frac{a}{a^2 + u^2} - \frac{b}{b^2 + (c-u)^2}.$$

Setting this equal to zero and simplifying, we get

$$b(a^2 + u^2) = a(b^2 + (c - u)^2). \tag{1}$$

Solving, we get

$$u = \frac{ac \pm \sqrt{ab\left((a-b)^2 + c^2\right)}}{a - b}.$$

Let

$$u_- = \frac{ac - \sqrt{ab\left((a-b)^2 + c^2\right)}}{a - b} \quad \text{and} \quad u_+ = \frac{ac + \sqrt{ab\left((a-b)^2 + c^2\right)}}{a - b}.$$

Clearly, $u_+ \geq \frac{ac}{a-b}$, and it is straightforward to show that $u_- \leq c$. Hence, $(u_+, 0)$ lies to the right of D (not shown in Cases 1, 2), while $(u_-, 0)$ lies to the left of C. Referring to any of the figures, the graphical evidence suggests that θ obtains a local maximum value when $u = u_+$ but obtains a global maximum value when $u = u_-$. We prove this below but in a geometric way. In the Regiomontanous problem, $c = 0$, in which case $u = \pm\sqrt{ab}$. Owing to the symmetry in this case, both points provide θ with its maximum value. While the calculus-based solution is complete, it reveals nothing of the underlying geometry of the problem. It is here that we make a departure.

Calling on differential geometry Up to this point, we kept A and B fixed and allowed U to vary. Now fix A and U, and allow B to vary. To emphasize that B is a variable point, relabel its coordinates from (c, b) to (x, y). With this change, equation (1) becomes

$$y(a^2 + u^2) = a(y^2 + (x - u)^2). \tag{2}$$

Now we reason "backwards." If u is fixed, (2) defines the circle $(x - u)^2 + \left(y - \frac{a^2+u^2}{2a}\right)^2 = \left(\frac{a^2+u^2}{2a}\right)^2$. This circle represents the set of points B such that $U = (u, 0)$ is the optimizing point. The situation is shown in Figure 3.

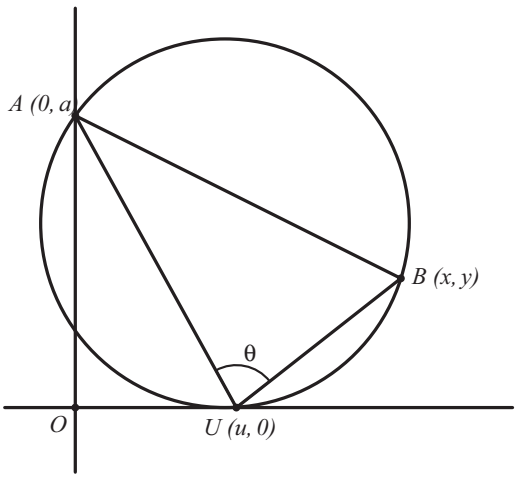


Figure 3 With A and U fixed, the circle is the set of points B for which U is the optimizing point.

Thus, the circle has center $\left(u, \frac{a^2+u^2}{2a}\right)$, radius $\frac{a^2+u^2}{2a}$, and is tangent to the x-axis at $U = (u, 0)$. Also, note that the circle passes through point A. As u varies through all real values, we get a family of such circles, as shown in Figure 4.

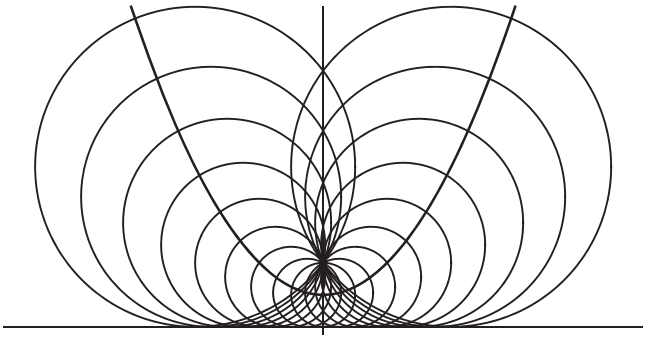


Figure 4 Family of circles through common point A (not labeled for clarity), tangent to the x-axis.

In Figure 4, every circle passes through A and is tangent to the x-axis. Furthermore, regarding the coordinates of the center of a circle in the family as a set of parametric equations, the centers lie on the parabola $y = \frac{a^2+x^2}{2a}$. It is clear that the family has an *envelope*, namely, the x-axis and point A. We digress briefly to discuss envelopes and describe the special relationship between our family of circles, their envelope, and the parabola of centers.

Envelopes Under mild conditions, a family of curves will have an envelope that, in general, is a visible feature of a plot of the family. Often (but not always) the envelope is a curve that, at each point, is tangent to a (different) member of the family. The reference [**1**] has a particularly complete description of envelopes. We give a definition here that indicates how to calculate them.

Definition. Let $F(t, x, y) = 0$ define a family of smooth curves (F is a smooth real-valued function). The *envelope* E of the family is given by the simultaneous solution of the equations $F(t, x, y) = 0$ and $\frac{\partial}{\partial t}F(t, x, y) = 0$. That is,

$$E = \{(x, y) : \text{there exists } t \text{ such that } F(t, x, y) = \tfrac{\partial}{\partial t}F(t, x, y) = 0\}.$$

In the present case, the family of circles is given by the equation $F(t, x, y) = 0$ where

$$F(t, x, y) = (x - t)^2 + \left(y - \frac{a^2 + t^2}{2a}\right)^2 - \left(\frac{a^2 + t^2}{2a}\right)^2.$$

Finding the envelope requires the simultaneous solution of the equations $F = \frac{\partial}{\partial t}F = 0$. That is,

$$\begin{aligned} &(x - t)^2 + \left(y - \frac{a^2 + t^2}{2a}\right)^2 - \left(\frac{a^2 + t^2}{2a}\right)^2 = 0, \\ &x + \frac{t}{a}y - t = 0. \end{aligned}$$

Solving for x and y, we get $x = t$, $y = 0$, which is the x-axis, and the single point $(0, a)$, which is point A.

In fact, the result here is well known in the study of planar curves. The envelope of a family of circles centered on a curve $\gamma(t)$ and passing through a common point P consists of the singleton $\{P\}$ along with a curve, known as the *orthotomic* of γ, relative to P. Thus, the x-axis is the orthotomic of the parabola $y = \frac{a^2+x^2}{2a}$, relative to A. As described in [**1**], the orthotomic of a curve γ, relative to a point P, can also be realized as the reflection of P through the tangents of γ. This allows us to compute a closed formula for the orthotomic, namely $\phi(t) = \mathbf{P} + 2((\gamma - \mathbf{P}) \cdot \mathbf{N})\ \mathbf{N}$ where $\mathbf{N}$ is the unit normal vector of γ and $\mathbf{P}$ is the position vector of P. Substituting directly into this formula confirms the result above, namely, $\phi(t) = \langle t, 0\rangle$. For completeness, we mention that since the x-axis is the orthotomic of the parabola $y = \frac{a^2+x^2}{2a}$, relative to A, then the parabola is the *antiorthotomic* of the x-axis, relative to A. We return to this idea later.

Summarizing the developments, after setting up a formula for θ and differentiating with respect to u, we arrive at the condition $(x-u)^2 + \left(y - \frac{a^2+u^2}{2a}\right)^2 = \left(\frac{a^2+u^2}{2a}\right)^2$ where x and y are in place of c and b, respectively. For fixed u, this is a circle centered at $\left(u, \frac{a^2+u^2}{2a}\right)$ on the parabola $y = \frac{a^2+x^2}{2a}$ and passing through A. In addition, it is tangent to the x-axis at $U = (u, 0)$. This circle represents the set of all points B for which U is the optimizing point. As u varies, we get a family of circles whose envelope is the orthotomic of the parabola, namely, the x-axis, along with point A. We apply these results to the maximum angle problem.

The solution Start from the beginning, where $A = (0, a)$ and $B = (c, b)$ are fixed. As before, assume $c \geq 0$ and $b \leq a$. We claim there are exactly two members of the family of circles described above that pass through A and B, and are tangent to the x-axis. To show this, plot the antiorthotomic of the x-axis relative to A, which is the parabola $y = \frac{a^2+x^2}{2a}$. Construct the perpendicular bisector of the segment AB. The intersection of this bisector with the antiorthotomic gives the center of a circle from the family. In particular, the circle with this center, passing through A and B, will be tangent to the x-axis, thanks to the way in which the antiorthotomic was constructed. Figure 5 illustrates the result. That there are two such circles comes from the fact that B lies above the x-axis. As a result, the perpendicular bisector of AB intersects the parabola in exactly two points. The only instances in which the bisector fails to intersect the parabola twice is when B lies below the x-axis (no intersection) or when B lies on the x-axis (the bisector is tangent to the parabola). Figure 6 illustrates the full result (two circles). Call the two circles C_1 and C_2, where C_1 is the leftmost (the smaller of the two). The critical points of θ are u_1 and u_2 where $U_1 = (u_1, 0)$ and $U_2 = (u_2, 0)$ are the points of tangency of C_1 and C_2 with the x-axis, respectively. Note, u_1 and u_2 are the former u_- and u_+, respectively, calculated from equation (1).

The last remaining detail is to determine which of the two critical points provides the global maximum for θ. The answer is u_1. Since the perpendicular bisector of AB has positive slope (except when $c = 0$, in which case both u_1 and u_2 provide the global maximum), the y-coordinate of the center of C_1 is less than that of the center of C_2. Consequently, the radius of C_1 is less than that of C_2. Both critical points subtend arcs from A to B on their respective circles. Refer to Figure 6. From standard Euclidean geometry, e.g., see [**3**] for results about circles and inscribed angles, $m\angle AU_1B = m\angle AGB$, and $m\angle AU_2B = m\angle AHB$. Then, comparing the isosceles triangles $\triangle AGB$ and $\triangle AHB$, the latter has legs whose (equal) length is greater than that of the former (since C_2 has larger radius), yet they share the same base. Hence, $m\angle AGB > m\angle AHB$. That is, $\theta_1 > \theta_2$.

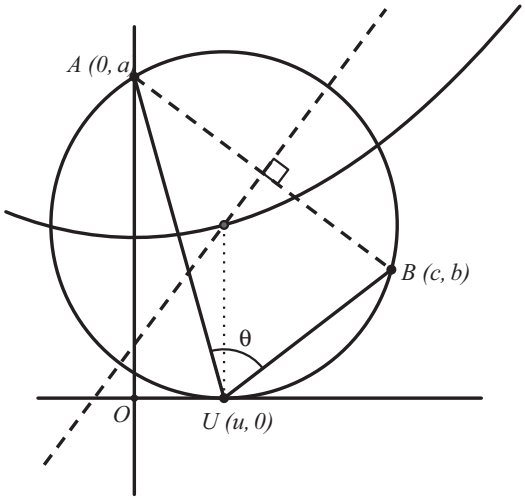


Figure 5 The intersection of the perpendicular bisector of AB with the parabola locates the center of one member of the family of circles through A and B and tangent to the x-axis. In turn, this locates a critical point u of θ.

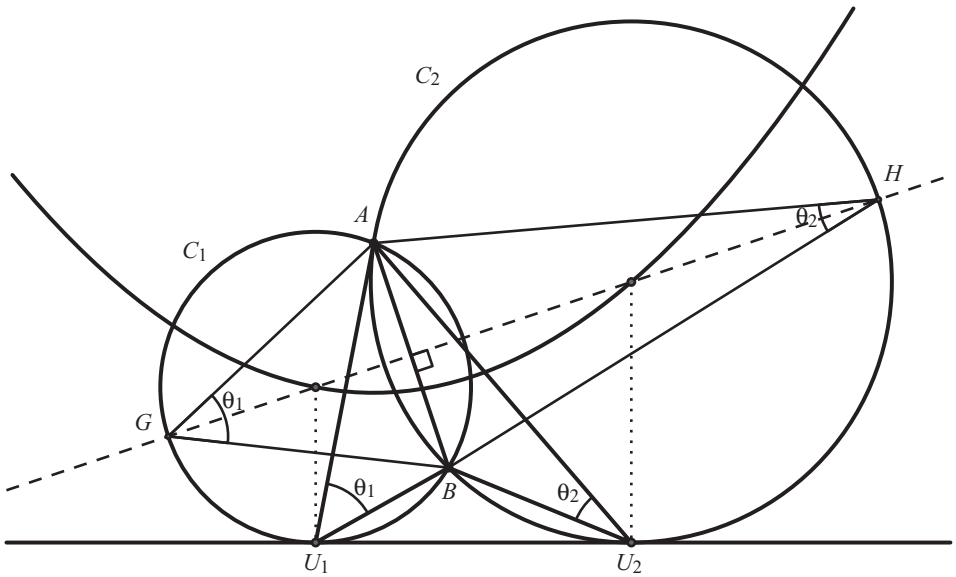


Figure 6 The perpendicular bisector of AB intersects the parabola twice, resulting in identification of the two circles from the family of circles through A and B and tangent to the x-axis. In turn, this locates the two critical points, u_1 and u_2 of θ.

Return to the Regiomontanus problem The geometric solution of the Regiomontanus problem described at the start of the paper is contained within our general result. Letting $c = 0$, points $A = (0, a)$ and $B = (0, b)$ are on the y-axis and represent the upper and lower edge of a painting, respectively. Construct the perpendicular bisector $y = \frac{a+b}{2}$ of AB and intersect it with the antiorthotomic (parabola). As above, the contending points U_1 and U_2 lie on the x-axis directly below the intersections. By the symmetry in this case, both points yield the same maximal angle θ, as in Figure 7. For the given parameters, $U = (\pm\sqrt{ab}, 0)$.

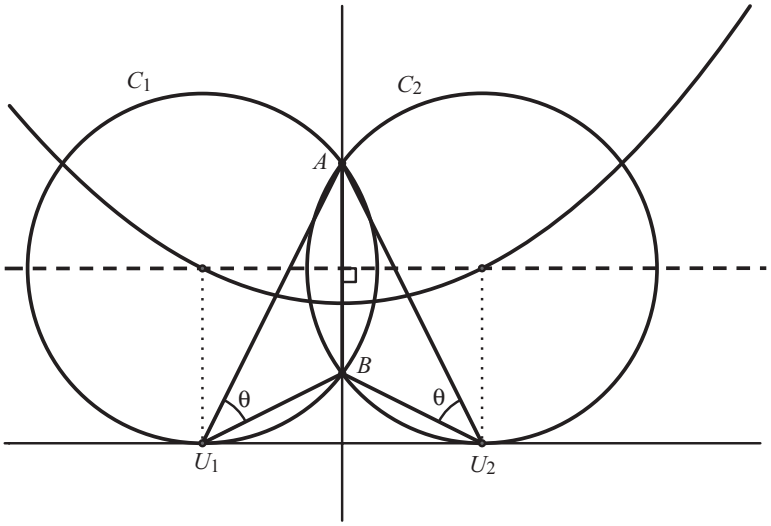


Figure 7 The solution of the Regiomontanus problem is a special case of a more general angle-maximization problem.

The method described here connects the analytic (calculus) solution of the Regiomontanus problem (as a special case of a more general angle-maximization problem) with the geometric. Neither solution, alone, suggests the other. The methods of differential geometry, namely, families of curves and envelopes, provide a link between the two. Readers may be interested in the variants of the Regiomontanus problem in [**2**] and [**7**], though [**7**] is a restatement of [**2**], and both are restatements of the problem in [**4**].

Acknowledgments The authors wish to thank the referees for their careful reading of the drafts. Their suggestions, and those of the editor, made for a much improved paper. This work was initiated in the 2012 Summer Science Research Program at Ohio Wesleyan University, where the first author was an undergraduate student. The authors thank OWU for its support.

REFERENCES

1. J. W. Bruce, P. J. Giblin, *Curves and Singularities*. Second ed. Cambridge Univ. Press, Cambridge, 1992.
2. M. Chamberlain, M. D. Myerson, Problems and solutions: Problem #532, *College Math. J.* **25** (1994) 334.
3. I. M. Isaacs, *Geometry for College Students*. The Sally Series, Pure and Applied Undergraduate Texts No. 8, American Mathematical Society, Providence, RI, 2001.
4. L. H. Lange, Constructing an old extreme viewpoint, *Math. Mag.* **5** (1987) 301-304.
5. P. J. Nahin, *When Least Is Best*. Princeton Univ. Press, Princeton, NJ, 2004.
6. I. M. Niven, *Maxima and Minima Without Calculus*. Dolciani Mathematical Expositions, No. 6, The Mathematical Association of America, Washington, DC, 1981.
7. H. Sedinger, Problems and solutions: Problem #624, *College Math. J.* **29** (1998) 153.
8. J. Stewart, *Calculus*: *Early Transcendentals (Single Variable)*. Sixth Ed. Cengage Learning, Boston, MA, 2007.

Summary. The Regiomontanus problem is that of determining the optimal place to stand to view a painting. The problem, which predates calculus by some 200 years, appears in most modern calculus books, though isn't known by that name. The original solution was purely geometric, the optimality following from the well-known intersecting secants theorem. Modern treatments are analytic, using calculus, but none of the underlying geometry (or its original solution) is apparent. In this paper, we generalize the problem and unite the two solutions using parametric curves and results from differential geometry.

BENJAMIN LETSON (MR Author ID: 1150971) is a Ph.D. student at the University of Pittsburgh. His primary research interests are multiple-time-scale dynamical systems, nonlinear contraction in dynamical systems, and theoretical neuroscience. In his spare time, he likes to read and make simple problems harder than they ought to be.

MARK SCHWARTZ (MR Author ID: 227853) earned his Ph.D. from Northwestern University and recently completed his 30th year in the Department of Mathematics and Computer Science at Ohio Wesleyan University. One day, he happened across an optimization problem admitting a differential geometric solution. He has been interested in such things ever since.

SOLUTION TO PINEMI PUZZLE

|	4	|||	8	7	8	|||	6	6	|||
2	6	8	||	|||	||	10	||	|	7
3	|	|	|	|	|	|	|	|||	8
|	|	|	6	8	|||	8	9	||	||
|||	|||	11	||	8	8	|	|	7	6
9	11	||	||	9	|	|||	8	6	||
||	|	13	|||	|	7	8	||	|	5
8	|||	||	10	|	|	7	|	8	||
||	10	|	8	||	||	|	9	|	|
|	|	2	5	||	8	|	|||	||	4

1 S	2 L	3 I	4 P	■	5 E	6 R	7 R	8 S	■	9 F	10 A	11 N	12 O
13 L	I	N	E	■	14 S	Q	U	A	T	■	15 I	R	O	N
16 A	M	E	N	■	17 A	U	N	T	Y	■	18 A	C	N	E
19 P	O	S	T	20 U	L	A	T	E	■	21 A	S	H	E	S
■	■	■	22 A	P	T	L	Y	■	23 A	S	C	I	■	■
24 L	25 A	26 R	G	O	S	■	■	27 A	X	I	O	M	28 S	■
29 O	D	E	O	N	■	30 M	31 A	R	I	A	■	32 E	E	33 C
34 T	H	A	N	■	35 P	A	L	E	S	■	36 E	D	D	Y
37 S	O	N	■	38 A	O	R	T	A	■	39 C	L	E	A	T
■	40 C	I	41 R	C	L	E	■	■	42 F	R	E	S	N	O
■	■	43 M	A	T	E	■	44 I	45 D	I	O	M	■	■	■
46 H	47 E	A	D	S	■	48 I	S	O	S	C	E	49 L	50 E	51 S
52 A	N	T	I	■	53 A	S	S	E	T	■	54 N	O	D	E
55 L	I	E	U	■	56 S	L	U	R	S	■	57 T	O	G	A
58 O	D	D	S	■	59 S	E	E	S	■	■	60 S	T	E	M

A New Perspective on Finding the Viewpoint

FUMIKO FUTAMURA
Southwestern University
Georgetown, TX 78626
futamurf@southwestern.edu

ROBERT LEHR
University of Texas at Austin
Austin, Texas 78701
robert.z.lehr@gmail.com

Figure 1 Hendrick van Vliet, Interior of the Oude Kerk, Delft, 1660.

You are in the Metropolitan Museum of Art in New York, and you come across Hendrick van Vliet's *Interior of the Oude Kerk, Delft*, 1660 [**14**]. Through his skilled use of perspective, van Vliet seems eager to make you feel as though you were actually there in the Oude Kerk, the oldest building still standing in Amsterdam. A perspective painting done with mathematical accuracy will act like a window to the 3-dimensional world, but you need to be standing where the artist stood to get this window effect.

As Leonardo da Vinci wrote in his notebooks, the spectator will see "every false relation and disagreement of proportion that can be imagined in a wretched work,

Math. Mag. **90** (2017) 267–277. doi:10.4169/math.mag.90.4.267.
MSC: Primary 51M16.

unless the spectator, when he looks at it, has his eye at the very distance and height and direction where the eye or the point of sight was placed in doing this perspective" (543, [**9**]). So how do we find this "point of sight"? Your instinct may be to stand directly in front of the center of the van Vliet painting. However, according to our measurements, you will need to stand with your eye about 2/3 of the way down far over to the left, nearly at the edge of the painting, back about the length of the height of the painting.

Although the painting is still realistic and impressive when viewed from other locations, without knowledge of the "point of sight" more commonly known as the *viewpoint*, it would be difficult to experience the spacious depth of the centuries-old church. Thus, we offer this article as a proposal to curators and museums to consider identifying the viewpoints of several perspective paintings in their collections and present the information in a format that helps visitors appreciate the art from the intended angle. We review several known geometric methods, simplify a known algebraic method, and use the insights gained from the simplification to introduce a new method which we call the perspective slope method, a satisfying blend of geometric and algebraic techniques.

Background

To determine the odd viewpoint for van Vliet's painting, we first notice that two main sets of lines parallel in the Oude Kerk appear to go to two different vanishing points, to the left and right of the painting as in Figure 2. This tells us that the painting is done in *two-point perspective*. Compared to a painting done in one-point perspective, this makes our job considerably harder (for more on finding the viewpoint of one-point perspective drawings along with the mathematics of perspective drawing, we recommend Frantz and Crannell's Viewpoints [**5**]).

Figure 2 Two-point perspective.

Over the centuries, several mathematicians have solved the problem of finding the viewpoint for two-point perspective for special cases, Simon Stevin (1605), Johann Heinrich Lambert (1759), and most notably, Brook Taylor of Taylor series fame, who gave several solutions in his two books on linear perspective (1715, 1719). A readable account of the history of what is known as the inverse problem of perspective is found in Andersen's book on the history of mathematical perspective [**1**]. A more modern,

related problem in computational projective geometry involves determining the exact location of a camera using on-site measurements and clues given in a photograph, called "camera calibration" or "camera resectioning." Devotees of THIS MAGAZINE may remember two articles written on this topic, "Where the Camera Was" by Byers and Henle [**2**] and "Where the Camera Was, Take Two" by Crannell [**3**], which respectively discussed an algebraic and a geometric approach to the problem.

The geometric methods

The standard geometric method described in [**5**] uses semicircles to find the viewpoint, as shown in Figure 3. There is no need for measurement or computation, but as we shall see, this method can sometimes be impractical.

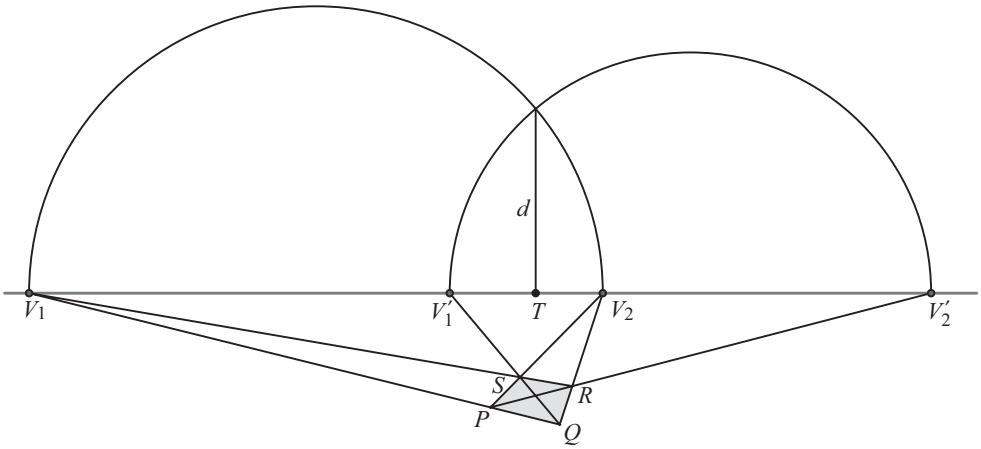


Figure 3 Finding the viewpoint for two-point perspective.

The first step is to identify a four-sided figure in the painting, such as $PQRS$ in Figure 3, which we know or can reasonably assume to be a square on the ground or in a plane parallel to the ground drawn in perspective. The two pairs of opposite sides of the square are parallel in the real world and not parallel to the canvas "window," which we call the *picture plane*, so their perspective images intersect at the *principal vanishing points* V_1 and V_2 along the *horizon line*. The diagonals of our square are also not parallel to the picture plane, and so their perspective images have vanishing points V_1' and V_2' along the horizon line.

The next step is to draw two semicircles, one with diameter $v = \overline{V_1V_2}$ and another with diameter $v' = \overline{V_1'V_2'}$, and find their intersection point. Through that intersection point, we draw a line orthogonal to the horizon line. The viewpoint is then determined by placing your eye directly in front of T as shown in Figure 3 at a distance d from the painting.

To understand why this works, let's float above the scene, to get the bird's eye view. We would see the viewer (or the viewer's eye) at O, the picture plane seen as a horizontal line, and the undistorted square, as in Figure 4. We can now see the distance from the viewer to the picture plane d and the point T on the picture plane directly in front of the viewer's eye.

V_1 is the vanishing point of the image of the line PQ. To understand how to find V_1 on this top view, consider a point X on line PQ. Its image X' is located on the

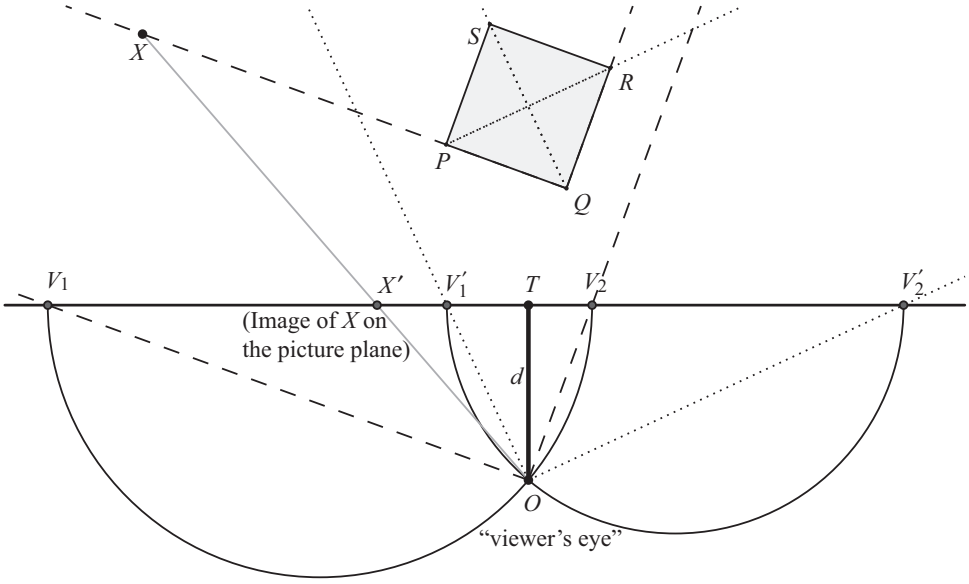


Figure 4 Bird's eye view.

picture plane where the sight line from the viewer's eye O to the point X intersects the picture plane. As we pull X off towards the left, X' is also pulled along the line of the picture plane towards the left. Notice that X' will converge to the point where the viewer can no longer see the line PQ, i.e., where the line parallel to PQ intersects the picture plane. So this must the the vanishing point V_1. Similarly, we can find the other vanishing points.

Since the sight lines to the principal vanishing points (indicated as the dashed lines in Figure 4) form a right angle as do the sight lines to the diagonal vanishing points (indicated as the dotted lines in Figure 4), we can use Thales' theorem for triangles inscribed in a semicircle to draw two semicircles between the vanishing points. We find the viewer's eye at the intersection of the semicircles, which determines the viewing distance d and viewing target T.

So in summary, the standard geometric method is as follows:

1. Find the vanishing points V_1, V_2, V_1', and V_2' along the horizon line.
2. Draw the semicircles with diameters $v = \overline{V_1V_2}$ and $v' = \overline{V_1'V_2'}$.
3. Find the intersection of the semicircles, and drop a line down perpendicular to the horizon line to find T, the point that should be directly in front of your eye.
4. The distance between the intersection and the horizon line is d, how far back from the painting you should stand.

Generally, this is a good method. Finding a square in the painting can sometimes be easy, for example, if there is a tiled floor. In van Vliet's painting, it's more difficult. The floor is tiled, but we know that the tiles in the Oude Kerk are not square (it should be remarked that all of these techniques generalize to a more general parallelogram situation, but you will need to know the angles and ratios of lengths). We decided to use the base of the front column, which we can reasonably assume is a square with the corners cut off for the following reason. The walls of the Oude Kerk are at 90 degree angles, and the two lines along the base vanish at the same vanishing points as the perpendicular lines along the walls as seen in Figure 2, hence are perpendicular themselves. We can reasonably assume that the columns have a circular rather than an

oval cross section, so we can conclude that the base is square with the corners cut off. If we apply the geometric method to the van Vliet painting using this square, we find that V_2' is located to the left of the painting, and it is quite a distance away from the other vanishing points as you can see in Figure 5.

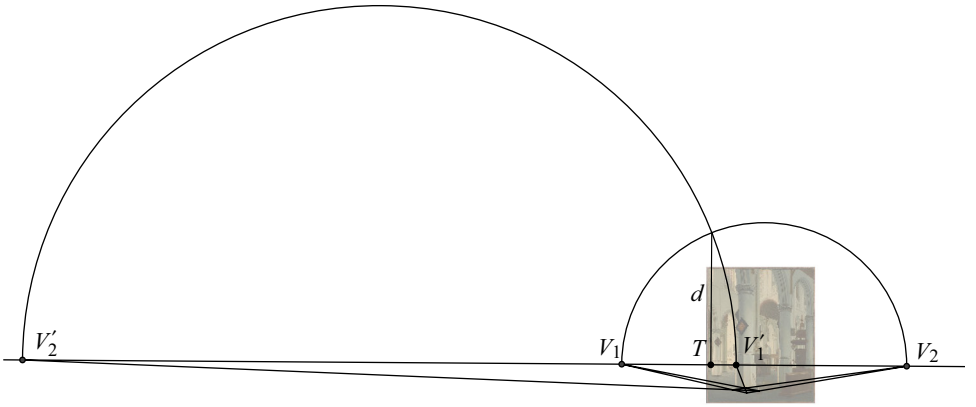


Figure 5 The standard geometric method for van Vliet's painting.

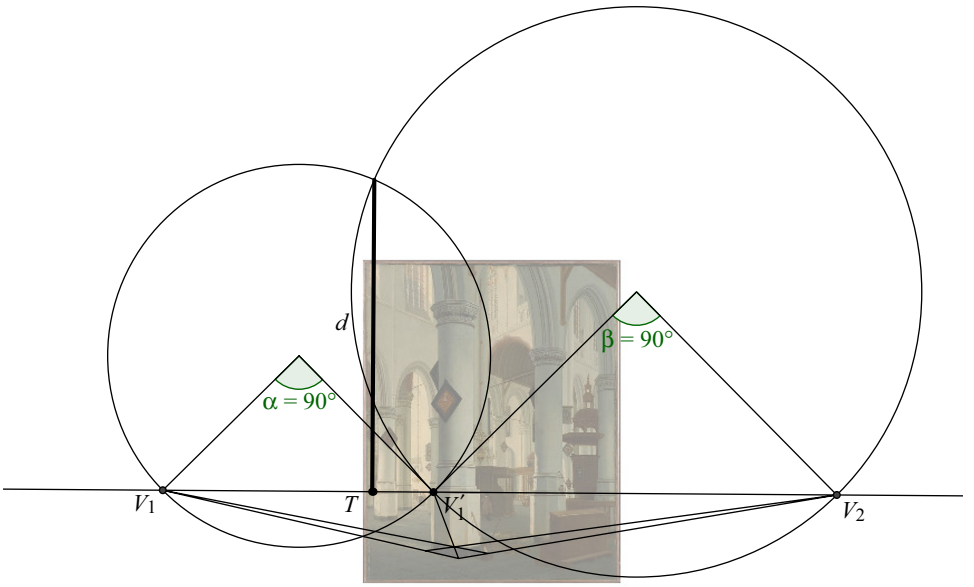


Figure 6 Taylor's method.

Taylor [**10, 1**] and Lambert [**8, 1**] provide two alternatives, which do not require the distant V_2' vanishing point. As shown in Figure 6, Taylor suggests we draw two right isosceles triangles with hypotenuses $\overline{V_1V_1'}$ and $\overline{V_1'V_2}$, respectively, then draw circles using the apexes as the centers and the legs of the triangles as the radii. The intersection of these two circles will be exactly where the intersection of the semicircles was in the previous method, thus giving us T and d. If you are able to draw right isosceles triangles and circles, this is not a bad alternative. Lambert gives us yet another alternate method, as shown in Figure 7. Again, it only requires three vanishing points, however, you have to be able to determine where the lines form a 45 degree angle.

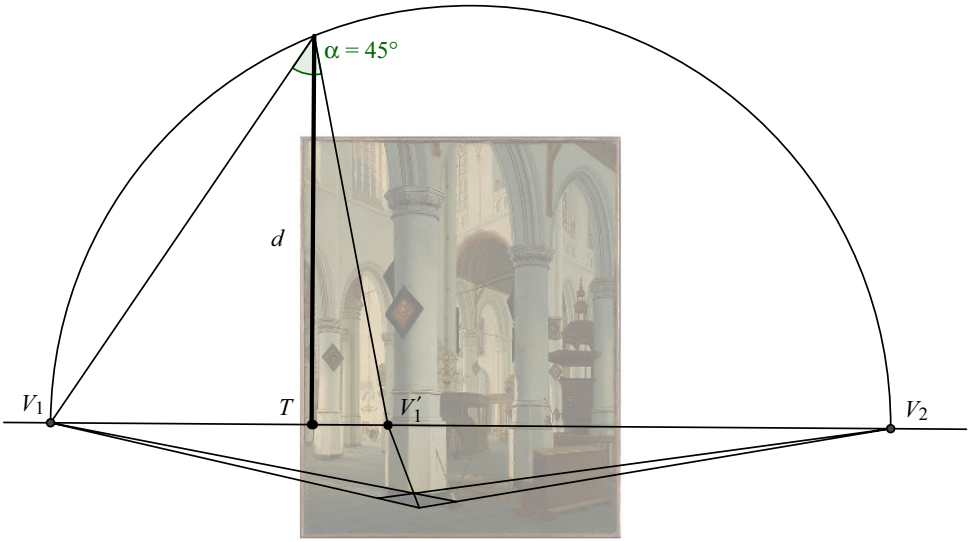


Figure 7 Lambert's method.

The algebraic method

Since we can find T and d by finding the intersection of two circles, we can surely find an algebraic formula for T and d. Indeed, this was done by Greene [**6**] who came up with a rather complicated formula which we will describe at the end of this section. We provide here a simplified version of Greene's formula whose setup and derivation provide insight into our new perspective slope method. For details on the calculations, we refer readers to the supplementary document at http://www.maa.org/sites/default/files/pdf/pubs/mm_supplements/Futamura-Lehr-appendix.pdf

In Figure 8 we superimpose a coordinate system with the horizon line as the x-axis in order to describe the semicircles algebraically. The ratio of the distances between the principal vanishing points and the diagonal vanishing point between them becomes quite important, so we shall denote this ratio as $\rho = \frac{v_L}{v_R}$, where v_L is the distance on the left (between V_1 and V_1') and v_R is the distance on the right (between V_1' and V_2).

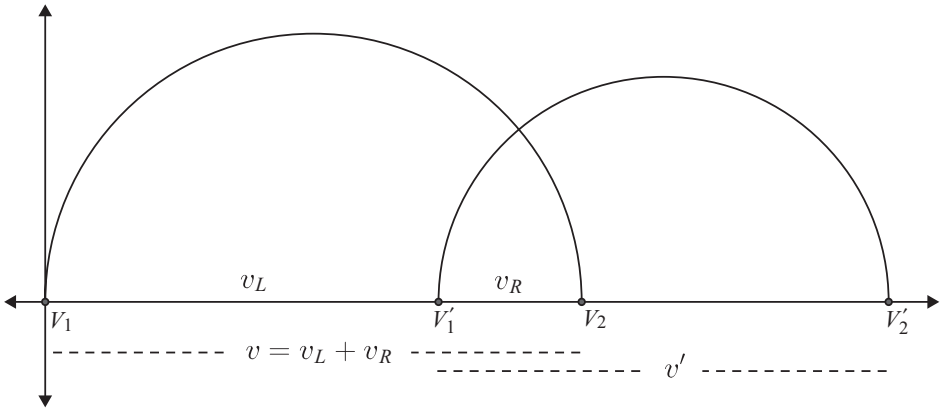


Figure 8 The algebraic method.

Theorem 1. *Let t be the distance from the leftmost principal vanishing point to the viewing target and d the viewing distance. Then*

$$t = \frac{\rho^2 v}{\rho^2 + 1} \quad \text{and} \quad d = \frac{\rho v}{\rho^2 + 1} = \frac{t}{\rho},$$

where $\rho = \frac{v_L}{v_R}$ and $v = v_L + v_R$.

Proof. Consider Figure 8. The equations for the semicircles are

$$\left(x - \frac{v}{2}\right)^2 + y^2 = \left(\frac{v}{2}\right)^2 \quad \text{and} \quad \left(x - \left(v_L + \frac{v'}{2}\right)\right)^2 + y^2 = \left(\frac{v'}{2}\right)^2.$$

To find t, the distance to the viewing target, we need to find the x-value of the circles' intersection point, so we subtract one equation from the other and solve for x. We find that

$$x = \frac{v_L(v_L + v')}{v_L - v_R + v'}. \tag{1}$$

In order to write (1) entirely in terms of v_L and v_R, we use what is known as the cross ratio. A *cross ratio* of four points along a line, A, B, C and D, is the following product of ratios of directed distances:

$$\times(ABCD) = \frac{|AB|}{|BC|}\frac{|CD|}{|DA|},$$

where a distance becomes directed by choosing a direction for a line, for example, a positive direction from A to B, so that the directed distance $|AB|$ is positive and $|BA|$ is negative. There are some very nice properties of the cross ratio, most notably that it is invariant under projections. For more on cross ratios, see [**4**]. The property that is most important for us here is that $\times(V_1V_1'V_2V_2') = -1$.

To see this, notice that the four-sided figure *PQRS* along with its two pairs of opposite sides and its diagonals (the third pair of opposite sides) form what is known as a *complete quadrangle*, with three *diagonal points* V_1, V_2, and the intersection of the diagonals. By definition, a set of four collinear points is a *harmonic set* if there exists a complete quadrangle such that two of the points are diagonal points and the other two points are on the opposite sides determined by the third diagonal point. Hence the two principal vanishing points and the two diagonal vanishing points form a harmonic set, denoted $H(V_1, V_2; V_1', V_2')$. It is well known that the cross ratio of a harmonic set equals -1. Thus, we have

$$\times\,(V_1V_1'V_2V_2') = \frac{|V_1V_1'|}{|V_1'V_2|}\frac{|V_2V_2'|}{|V_2'V_1|} = \frac{v_L}{v_R}\cdot\frac{v' - v_R}{-v' - v_L} = -1. \tag{2}$$

Solving for v' in (2), we find

$$v' = \frac{2v_Lv_R}{v_L - v_R}. \tag{3}$$

Going back to our equation (1) and substituting in (3) gives

$$x = \frac{v_L\left(v_L + \frac{2v_Lv_R}{v_L - v_R}\right)}{v_L - v_R + \frac{2v_Lv_R}{v_L - v_R}} = \frac{v_L^2(v_L + v_R)}{v_L^2 + v_R^2} = \frac{(\frac{v_L}{v_R})^2(v_L + v_R)}{(\frac{v_L}{v_R})^2 + 1} = \frac{\rho^2 v}{\rho^2 + 1} = t.$$

Solving for the y-value of the intersection point gives d from

$$y^2 = \left(\frac{v}{2}\right)^2 - \left(\frac{\rho^2 v}{\rho^2+1} - \frac{v}{2}\right)^2 = \frac{v^2}{4}\left(\frac{2\rho^2}{\rho^2+1}\right)\left(\frac{2}{\rho^2+1}\right) = \frac{v^2\rho^2}{(\rho^2+1)^2}.$$

Taking the positive solution, we get

$$d = \frac{v\rho}{\rho^2+1} = \frac{t}{\rho}.$$

The algebraic method is as follows:

1. Find the vanishing points V_1, V_2, and V_1' along the horizon line.
2. Measure v_L and v_R, and calculate $\rho = \frac{v_L}{v_R}$ and $v = v_L + v_R$.
3. Calculate $t = \frac{\rho^2 v}{\rho^2+1} = |V_1 T|$ to find the viewing target T.
4. Divide t by ρ to find the viewing distance d.

Applying the algebraic method to the van Vliet painting, $\rho = \frac{69.44}{103.44} \approx 0.67$, $v = 69.44 + 103.44 = 172.88$. Thus $t \approx 53.71$ cm and $d \approx 80.00$ cm. In Figure 9 we see where this T is located on the picture plane, and we see that it nearly coincides with the T found using the geometric method. The small difference of approximately 0.34 cm is due to round-off error. We can make the calculation easier, by approximating $\rho \approx 2/3$ and $v \approx 173$, and the simple calculation $173 \cdot 4/13 \approx 53.23$ is still less than a centimeter off.

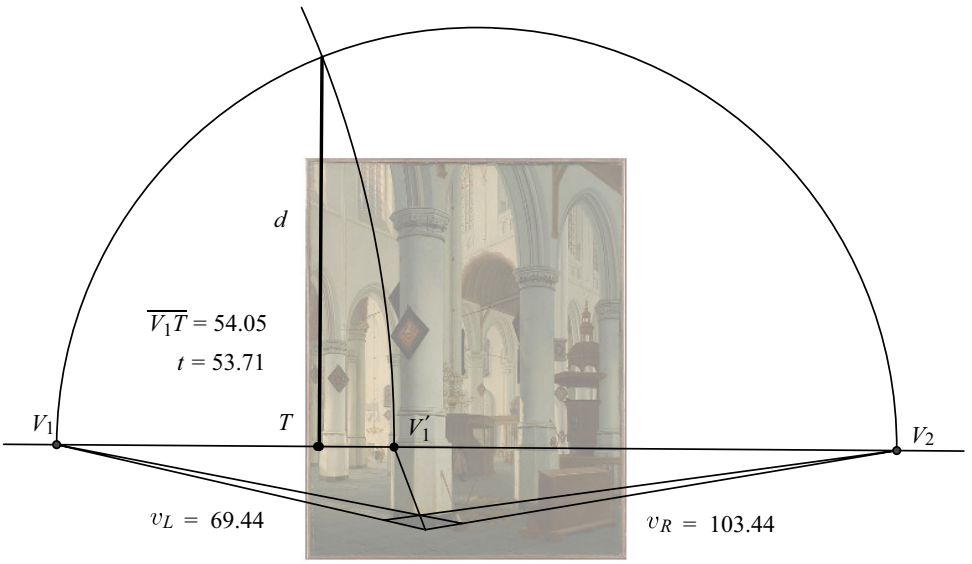


Figure 9 The algebraic method applied to Van Vliet's painting.

We compare this with Greene's formula given in [**6**],

$$t = \frac{vv_L^2}{v^2 - 2vv_L + 2v_L^2}\,,\; d = \frac{[vv_L^2(v^3 - 2v^2v_L + vv_L^2)]^{1/2}}{v^2 - 2vv_L + 2v_L^2}.$$

Instead of ρ and v, he used v_L and v (denoted s and D in [**6**]). This version of the formula is considerably harder to remember and harder to use.

The perspective slope method

We come now to our new method, which we call the perspective slope method. Earlier we mentioned that the cross ratio of a harmonic set equals -1. Let's see what happens when we replace the diagonal vanishing point V_1' with T.

$$\times(V_1TV_2V_2') = \frac{|V_1T|}{|TV_2|}\frac{|V_2V_2'|}{|V_2'V_1|} = \frac{t}{t-v_L}\cdot\frac{v'-v_R}{v'+v_L}. \tag{4}$$

By (2), $\frac{v'-v_R}{v'+v_L} = -\frac{1}{\rho}$. So using this substitution along with the formula for t in (4), we find

$$\times(V_1TV_2V_2') = \frac{\frac{\rho^2 v}{\rho^2+1}}{v-\frac{\rho^2 v}{\rho^2+1}}\frac{-1}{\rho} = \frac{\rho^2 v}{\rho^2 v + v - \rho^2 v}\frac{-1}{\rho} = -\rho.$$

This is rather nice, so how might we use it?

To answer this question, we establish a relationship between cross ratios and the slopes of lines on a coordinate grid drawn in perspective. Assuming $PQRS$ is the perspective image of a square, we can use this to create a perspective image of a coordinate grid with $PQRS$ as the image of one square of the grid. Assuming V_1 is the vanishing point of the x-axis and V_2 the vanishing point of the y-axis, we give a sketch of the proof below that a line with slope m in the perspective coordinate grid will vanish at a point M such that $\times(V_1MV_2V_2') = m$.

Theorem 2. *Let $PQRS$ be a complete quadrangle with associated harmonic set $H(V_1, V_2; V_1', V_2')$. Assuming that $PQRS$ is the perspective image of a square with the perspective coordinate grid set up as described above, a line with slope m in the perspective coordinate grid will vanish at a point M such that $\times(V_1MV_2V_2') = m$.*

Sketch of Proof.

We first look at the one-point perspective situation, as in Figure 10. ℓ will vanish at M located d/m away from V_2. This is clear from considering the top view, but details are found in [**5**]. Hence, the cross ratio will be

$$\times(V_1MV_2V_2') = \frac{|V_1M|}{|MV_2|}\frac{|V_2V_2'|}{|V_2'V_1|} = \frac{\infty}{-d/m}\frac{d}{-\infty} = m.$$

Since the cross ratio is a projective invariant, this will also hold in two-point perspective.

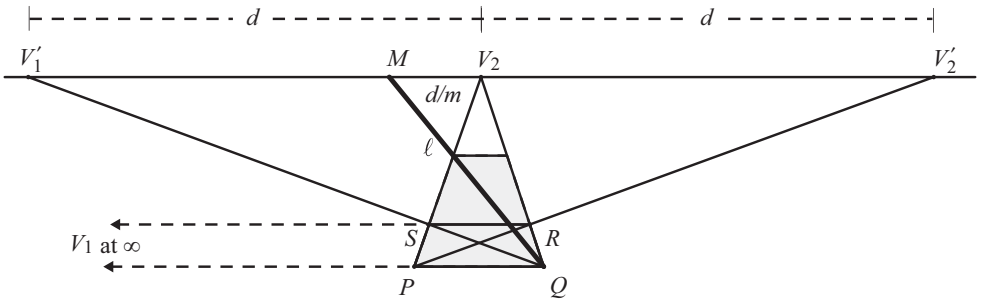


Figure 10 One-point perspective situation.

So in order to find T, we need only calculate $\rho = \frac{v_L}{v_R}$, draw the line with slope $-\rho$, and its vanishing point will be T! Once you determine the location of T and measure $t = |V_1T|$, $d = t/\rho$.

So our final method is as follows:

1. Find the vanishing points V_1, V_2, and V_1' along the horizon line.
2. Measure v_L and v_R, and calculate $\rho = \frac{v_L}{v_R}$.
3. Use your perspective drawing skills to draw a line with slope $-\rho$ in perspective, and find its vanishing point. This is T.
4. Divide $t = |V_1T|$ by ρ to find the viewing distance d.

In van Vliet's painting, we approximate $\rho = 0.67$ by $2/3$, which makes finding the slope a bit easier. One way is to draw a 2×3 grid, as shown in Figure 11. Despite our rough estimation, it does a pretty good job! The viewing distance is then roughly $3/2$ of t. In a two-point perspective painting with a tiled floor, this method is very easy to use.

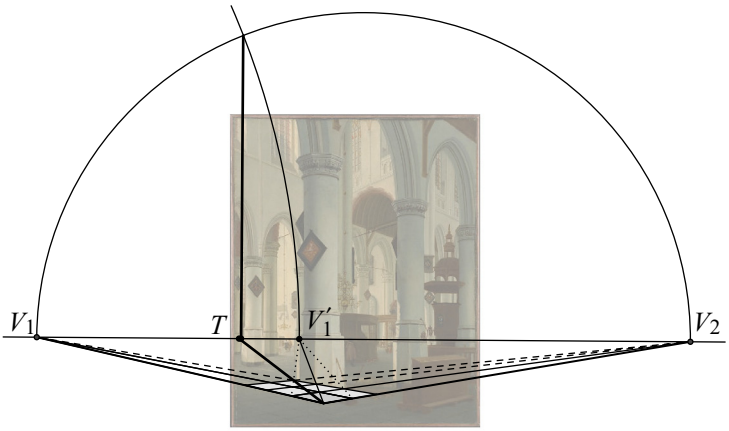


Figure 11 The perspective slope method applied to the van Vliet drawing.

Conclusion

We have presented a variety of methods for finding the viewpoint of a two-point perspective painting, each with their advantages and disadvantages. We have also found that in many of our calculations, we used approximations that shifted our viewpoint by a small amount. Do we have to stand with our eye "at the very distance and height and direction" as Leonardo da Vinci suggests or do we have a little room for error? Since the Renaissance, much research has been done on the psychology of human vision and picture perception and the answer is more nuanced than previously thought. Regarding being too close or too far from the picture plane, a simple geometrical analysis of the situation tells us that the distortion is proportional to the viewing distance so we have more room for error if the viewing distance is large. It was found in one experiment of 12 college students in [**12**] that on average, they perceived a 1:1 ratio of sides of a rectangle in perspective at 21.4 cm, although the actual viewing distance was 28 cm. This was found to not be significantly different from the mean with a standard deviation of 16.7 cm. And although the math seems to indicate that we would perceive a 2:1 ratio of sides at double the viewing distance, the same study showed that actually this was perceived at an average of 422.1 cm with standard deviation 92.3 cm. In another

study [**7**], it was shown that we compensate for a difference in viewing distance better if the painting does not have a low eye height and does not depict a wide-angle view. In fact, even at the correct viewpoint, we will see distortions around the periphery in a mathematically accurate, wide-angle perspective drawing. And finally, if we are to the left or right of the viewpoint or if we view the painting at an angle, having both eyes open allows us to compensate for a larger difference [**13**].

If our eye is at the viewpoint, the research generally supports the idea that we will have the feeling of being immersed in a 3D environment. Indeed, if you try this with the van Vliet painting, you see the interior from the height of an average person standing in the church, and you become suddenly aware of the people and chandelier in the far room. The arches soar overhead and you can feel the spaciousness of the old church. The effect is magical. Our great hope is that you will share this article with your local museums or at the very least feel empowered to use these techniques yourself on digital images to determine viewpoints prior to a museum visit.

Acknowledgments The authors would like to thank Annalisa Crannell, Marc Frantz, Michael A. Jones, and the referees for their helpful comments and suggestions, which greatly improved this manuscript. Futamura would also like to acknowledge support from NSF-DUE grant #1140113.

REFERENCES

1. K. Andersen, *The Geometry of an Art: The History of the Mathematical Theory of Perspective from Alberti to Monge*. Springer, New York, 2006, http://dx.doi.org/10.1080/00033790902730636.
2. K. Byers, J. Henle, Where the camera was, *Math. Mag.* **77** (2004) 251–259.
3. A. Crannell, Where the camera was, take two, *Math. Mag.* **79** (2006) 306–308.
4. H. Eves, *A Survey of Geometry: Revised Edition*. Allyn and Bacon, Boston, MA, 1972.
5. M. Frantz, A. Crannell, *Viewpoints: Mathematical Perspective and Fractal Geometry in Art*. Princeton Univ. Press, Princeton, NJ, 2011.
6. R. Greene, Determining the preferred viewpoint in linear perspective, *Leonardo* **16** no. 2 (1983) 97–102.
7. I. Juricevic, J. Kennedy, Looking at perspective pictures from too far, too close, and just right, *J. Exp. Psychol. General* **135** no. 3 (2006) 448–461.
8. J. H. Lambert, *La Perspective Affranchie de l'Embaras du Plan Géometral*. Chez Heidegguer et Comp., Paris, 1759.
9. Leonardo da Vinci, J. P. Richter, *The Notebooks of Leonardo Da Vinci: Volume 1*. Courier Dover Publications, Mineola, NY, 1970.
10. B. Taylor, *Linear Perspective*. R. Knaplock, London, 1715.
11. ———, *New Principles of Linear Perspective*. R. Knaplock, London, 1719.
12. B. Todorović, The effect of the observer vantage point on perceived distortions in linear perspective images, *Atten. Percept. Psychophys.* **71** no. 1 (2009) 183–193.
13. D. Vishwanath, A. Girshick, M. Banks, Why pictures look right when viewed from the wrong place, *Nat. Neurosci.* **8** (2005) 1401–1410.
14. H. van Vliet, *Interior of the Oude Kerk, Delft*. 1660. Metropolitan Museum of Art, New York, http://www.metmuseum.org/toah/works-of-art/1976.23.2.

Summary. We take a fresh perspective on an old idea and create an alternate way to answer the question: where should we stand in front of an image in two-point perspective to view it correctly? We review known geometric and algebraic techniques, then use the cross ratio to derive a simple algebraic formula and a technique that makes use of slopes on a perspective grid.

FUMIKO FUTAMURA (MR Author ID: 801853) received her Ph.D. from Vanderbilt University and is an associate professor at Southwestern University in Georgetown, TX. She enjoys discovering mathematics in works of art with her students and colleagues and also enjoys making art in a variety of media including oil paint, charcoal, and crochet.

ROBERT LEHR (MR Author ID: 1201123) has a B.A. in Mathematics from Southwestern University and is currently researching transportation patterns for the Center for Transportation Research: Network Modeling Center at UT. His passion is creativity and innovation, and he enjoys solving brain teasers and proofs.

Where Should I Open My Restaurant?

NATHAN KAPLAN
University of California, Irvine
Irvine, CA 92697
nckaplan@math.uci.edu

> There are three things that matter in property: *location, location, location*.
>
> *Unknown. Possibly Lord Samuel of Britain [**8**]*

There is a lot of truth in this opening quote, especially applied to commercial real estate. If you want your restaurant to succeed, customers have to walk through the door.

The problem

A walk on the $\mathbb{Z}^2$ lattice using only the steps $(0, 1)$, which we call "North," and $(1, 0)$, which we call East," is a *lattice path*.

Question 1. 1. *How many lattice paths begin at* $(0, 0)$ *and end at* (m, n)*?*

2. *What proportion of these paths pass through the point* (a, b)*?*

Versions of these elementary exercises appear in many introductory combinatorics textbooks, as lattice paths provide a nice setting to explore identities involving binomial coefficients. See for example [**3**, Chapter 4], [**2**, Section 3.5], [**4**, Section 8.5], and [**5**, Section 2.6].

Solution. 1. A path from $(0, 0)$ to (m, n) consists of m steps East, and n steps North, and can be uniquely identified with a word of length $m + n$ consisting of m E's and n N's. Therefore, there are $\binom{m+n}{m}$ total lattice paths.

2. Every lattice path passing through (a, b) consists of two parts: a lattice path from $(0, 0)$ to (a, b), followed by a lattice path from (a, b) to (m, n). The number of paths of the first type is $\binom{a+b}{a}$ and paths of the second type are in bijection with lattice paths from $(0, 0)$ to $(m - a, n - b)$. Therefore, the proportion of paths passing through (a, b) is

$$F_{m,n}(a, b) := \frac{\binom{a+b}{a}\binom{m-a+n-b}{m-a}}{\binom{m+n}{m}}. \tag{1}$$

This function is our main object of study. We think of the integers m and n being fixed and therefore focus on the numerator, which we denote $f_{m,n}(a, b)$.

Question 2. *In your town, everyone lives in one giant apartment building at* $(0, 0)$ *and works in the local factory at* (m, n)*. For some strange reason, locals choose their walks to work randomly among all minimal lattice paths.*

Math. Mag. **90** (2017) 278–285. doi:10.4169/math.mag.90.4.278.
MSC: Primary 60C05, Secondary 05A15

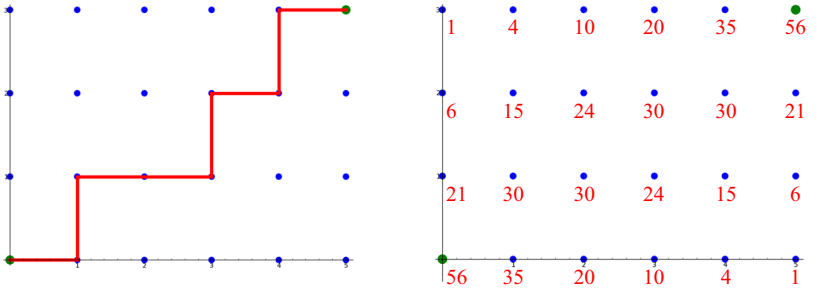


Figure 1 (Left): A path of minimal length from $(0, 0)$ to $(5, 3)$. (Right): The number of such paths passing through each point.

Nathan has plans to open a restaurant at some point (a, b)*. If you happen to pass by on your way to work, you will buy a coffee and muffin, arriving at work caffeinated, satisfied, and productive. If you do not pass Nathan's, you arrive at work miserable and tired.*

Zoning regulations forbid building on the occupied vertices $(0, 0)$ *and* (m, n)*, but otherwise Nathan is free to choose where to locate his establishment. Where should he open his restaurant in order to maximize the chance that people will visit?*

We rephrase this question in terms of the function defined above.

Question 3. *Given* $m, n > 0$ *what is the maximum value of* $f_{m,n}(a, b)$ *subject to* $0 \leq a \leq m$, $0 \leq b \leq n$, *and* $0 < a + b < m + n$*?*

Basic symmetries of binomial coefficients imply that $f_{m,n}(a, b) = f_{n,m}(b, a)$ and $f_{m,n}(a, b) = f_{m,n}(m - a, n - b)$. Therefore, answering this question for $m \geq n$ and subject to the additional constraint $0 < a + b \leq \frac{m+n}{2}$, is enough to solve it in all instances.

Theorem. Let m and n be positive integers with $m \geq n$. The maximum value of $f_{m,n}(a, b)$ subject to $0 \leq a \leq m$, $0 \leq b \leq n$, and $0 < a + b \leq \frac{m+n}{2}$ is given by $f_{m,n}(1, 0) = \binom{m+n-1}{m-1}$ if $m > n$ and $f_{m,n}(1, 1) = 2\binom{m+n-2}{m-1}$ if $m = n$.

We say that a point (a', b') *gives a maximum* if $f_{m,n}(a', b')$ solves the optimization problem in Question 3.

The solution we give is elementary, but not obvious. Instead of solving a two variable optimization problem we restrict to certain single variable refinements, finding maximum values by analyzing ratios of consecutive terms. We first focus on the "square case" $m = n$, and use it in the proof of the general case. In the final section we discuss connections to the Gamma function, the hypergeometric distribution, and higher-dimensional lattice paths.

The proof

Before giving a formal proof of the theorem we give some intuition for why it is true. Since all of our paths start at $(0, 0)$ and end at (m, n) it seems reasonable that points (a, b) that are close to $(0, 0)$ or to (m, n) will have many paths passing through them. Since the most direct path from $(0, 0)$ to (m, n) in $\mathbb{R}^2$ is given by the straight line $y = \frac{n}{m}x$, we expect that most lattice paths do not stray too far from this line.

In order to test this intuition we consider $f_{m,n}(a, b)$ for some points close to the origin, those with $a + b \in \{1, 2\}$. We compare these points by taking the ratio of the corresponding function values. Suppose that $n \geq 2$. Since $m \geq n$, we have

$$\frac{f_{m,n}(1,0)}{f_{m,n}(0,1)} = \frac{\binom{m+n-1}{m-1}}{\binom{m+n-1}{n-1}} = \frac{m}{n} \geq 1.$$

Similarly, $f_{m,n}(2,0) \geq f_{m,n}(0,2)$. We also have

$$\frac{f_{m,n}(2,0)}{f_{m,n}(1,1)} = \frac{m-1}{2n}.$$

Which of $(2,0)$ and $(1,1)$ has more paths passing through it depends on the ratio $\frac{m-1}{n}$, which is closely related to the slope of the line $y = \frac{n}{m}x$.

It is easy to check that $f_{m,n}(1,0) \geq f_{m,n}(2,0)$, so the only other comparison to make is

$$\frac{f_{m,n}(1,0)}{f_{m,n}(1,1)} = \frac{m+n-1}{2n},$$

which is at least 1 except in the square case, $m = n$. This square case is the only situation in which the point $(1,1)$ actually lies on the line $y = \frac{n}{m}x$, making up for the fact that it is two steps from the origin instead of one.

Our goal is to make this reasoning more precise. We first prove the theorem in the square case, which involves optimizing $f_{n,n}(a,b)$ restricted to two different kinds of diagonal lines. We then use this result to prove the theorem for general rectangular grids, inducting on the size of $m-n$.

The square case

Proof of the theorem (for $m = n$). The idea of the proof is to turn this two variable optimization problem into a single variable one. For each $k \in [1, 2n-1]$, we find the maximum value of $f_{n,n}(a,b)$ restricted to the diagonal line $a+b=k$. By symmetry, we need only consider $k \in [1,n]$.

Diagonals: $a+b=k$

Consider the set of all $(a, k-a)$ where $a \in [0,k]$. We claim that

$$\frac{f_{n,n}(a-1,k-a+1)}{f_{n,n}(a,k-a)} \geq 1, \text{ if and only if } a \geq \frac{k+1}{2}.$$

As we move along the diagonal consisting of points $(k,0),(k-1,1),\ldots,(0,k)$ the value of this function increases as we take steps "Northwest" until we cross the vertical line $a = \frac{k+1}{2}$. When k is odd the two points closest to the central diagonal $(\frac{k+1}{2}, \frac{k-1}{2})$ and $(\frac{k-1}{2}, \frac{k+1}{2})$ give the same maximum value. When k is even the maximum value occurs at $(\frac{k}{2}, \frac{k}{2})$.

Consider the ratio of consecutive terms

$$\frac{f_{n,n}(a-1,k-a+1)}{f_{n,n}(a,k-a)} = \frac{a}{k-a+1} \cdot \frac{n-k+a}{n-a+1}. \tag{2}$$

This is at least 1 if and only if

$$a(n-k+a) - (k-a+1)(n-a+1) = (n+1)(2a-(k+1)) \geq 0. \tag{3}$$

Since $n+1$ is positive, this holds for $a \in \left[\frac{k+1}{2}, k\right]$.

Diagonals: $a=b$ **and** $a=b+1$

Since $f_{n,n}(a,b)$ restricted to each line $a+b=k$ reaches a maximum either at $f_{n,n}\left(\frac{k}{2}, \frac{k}{2}\right)$ or at $f_{n,n}(\frac{k+1}{2}, \frac{k-1}{2})$, the symmetry of $f_{n,n}(a,b)$ implies that we need

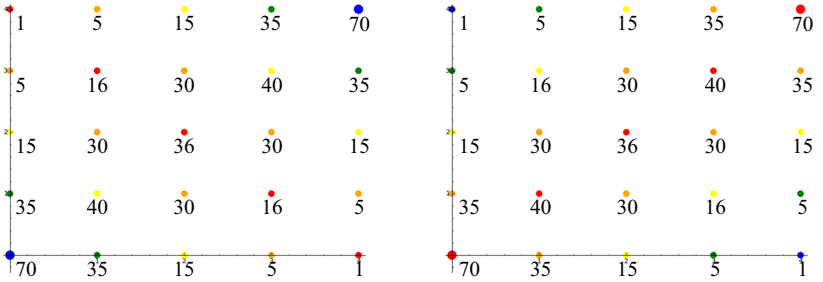


Figure 2 Values of $f_{4,4}(a, b)$. In the first plot vertices are grouped along diagonal lines $a + b = k$. In the second plot vertices are group along diagonal lines $b = a + k$.

only show that $f_{n,n}(1, 1)$ is a maximum among all $f_{n,n}(a, a)$ for $a \in \left[1, \frac{n}{2}\right]$ and that $f_{n,n}(1, 0)$ is a maximum among all $f_{n,n}(a + 1, a)$ for $a \in \left[0, \frac{n-1}{2}\right]$.

We consider the ratio

$$\frac{f_{n,n}(a+k, a)}{f_{n,n}(a+k+1, a+1)} = \frac{(a+1)(a+k+1)}{(2a+k+2)(2a+k+1)} \cdot \frac{(2n-2a-k)(2n-2a-k-1)}{(n-a)(n-a-k)}$$

in the special cases $k = 0$ and $k = 1$. When $k = 0$, this simplifies to

$$\frac{(a+1)(2n-2a-1)}{(2a+1)(n-a)},$$

which is greater than 1 if and only if

$$(a+1)(2n-2a-1) - (2a+1)(n-a) = n-1-2a \geq 0.$$

This holds for $a \in \left[0, \frac{n-1}{2}\right]$. When $k = 1$, our ratio simplifies to

$$\frac{(a+2)(2n-2a-1)}{(2a+3)(n-a)},$$

which is greater than 1 if and only if

$$(a+2)(2n-2a-1) - (2a+3)(n-a) = n-2-2a \geq 0.$$

This holds for $a \in \left[0, \frac{n-2}{2}\right]$.

We see that the maximum value of $f_{n,n}(a, a)$ subject to $a \in \left[1, \frac{n}{2}\right]$ is given by $f_{n,n}(1, 1)$ and that the maximum value of $f_{n,n}(a + 1, a)$ subject to $a \in \left[0, \frac{n-1}{2}\right]$ is given by $f_{n,n}(1, 0)$. Noting that $f_{n,n}(1, 1) > f_{n,n}(1, 0)$, completes the proof of the theorem in the square case.

In our argument for the general rectangular case we require a slight refinement of the square case.

Lemma. For $n \geq 5$ the largest value of $f_{n,n}(a, b)$ where

$$(a, b) \notin \{(0, 0), (n, n), (1, 1), (n-1, n-1)\}$$

is given by $f_{n,n}(1, 0)$.

The $n \geq 5$ assumption is necessary. For $n = 4$ the largest value of $f_{4,4}(a, b)$ where $(a, b) \notin \{(0, 0), (4, 4), (1, 1), (3, 3)\}$ is given by $f_{4,4}(2, 2) = 36 > f_{4,4}(1, 0) = 35$.

Proof. A point (a', b') with $0 < a' + b' = k \leq n$ maximizing $f_{n,n}(a, b)$ subject to the conditions of the lemma either gives the maximum value of $f_{n,n}(a, b)$ among all points on the diagonal line $a + b = k$, or has $a' + b' = 2$, in which case we note that $f_{n,n}(1, 0) \geq f_{n,n}(2, 0)$. Therefore, we need only consider points on the diagonal lines $a = b$ and $a = b + 1$. By the analysis in the proof above, we need only consider the ratio

$$\frac{f_{n,n}(1,0)}{f_{n,n}(2,2)} = \frac{\binom{2n-1}{n-1}}{6\binom{2n-4}{n-2}} = \frac{(2n-1)(2n-2)(2n-3)}{6n(n-1)^2}.$$

This is at least 1 for $n \geq 5$, completing the proof of the lemma. ∎

The rectangular case

Proof of the theorem (the rectangular case). The idea of the proof is to divide the set of lattice paths from $(0, 0)$ to (m, n) into two disjoint sets: paths passing through $(m - 1, n)$, and paths passing through $(m, n - 1)$. If $(1, 0)$ gives a maximum for $f_{m-1,n}(a, b)$ and for $f_{m,n-1}(a, b)$, then we conclude that $(1, 0)$ gives a maximum for $f_{m,n}(a, b)$.

We proceed by a kind of double induction. For a given value of n we induct on m. The case $n = 0$ is trivial. We give the argument for $n = 1$ in detail, and then adapt it to the general situation. We compute that $f_{2,1}(1, 0) = f_{2,1}(1, 1)$. Since $(1, 0)$ gives a maximum for $f_{2,1}(a, b)$ and for $f_{3,0}(a, b)$, we conclude that $(1, 0)$ also gives a maximum for $f_{3,1}(a, b)$. Suppose that $m \geq 3$ and that $(1, 0)$ gives a maximum for $f_{m,1}(a, b)$. Obviously, $(1, 0)$ gives a maximum for $f_{m+1,0}(a, b)$, so we conclude that $(1, 0)$ gives a maximum for $f_{m+1,1}(a, b)$.

We now argue by induction on n. Suppose that for $k \in [1, n - 1]$ and any $m > k$ we know that $(1, 0)$ gives a maximum for $f_{m,k}(a, b)$. We show that $(1, 0)$ also gives a maximum for $f_{m,n}(a, b)$.

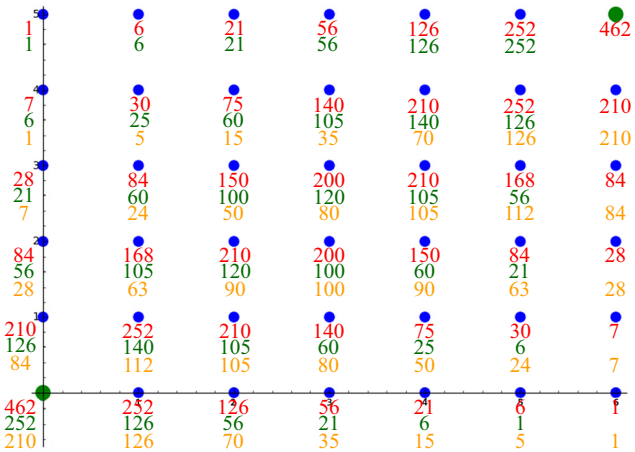


Figure 3 Values of $f_{6,5}(a, b)$ (red), $f_{5,5}(a, b)$ (green), and $f_{6,4}(a, b)$ (orange). Comparing $f_{6,5}(1, 0)$ and $f_{6,5}(1, 1)$, noting that $(1, 0)$ gives the second largest value of $f_{5,5}(a, b)$ and gives a maximum for $f_{6,4}(a, b)$, implies that $(1, 0)$ gives a maximum for $f_{6,5}(a, b)$.

For a fixed n, the base case of our induction is to show that $(1, 0)$ gives a maximum for $f_{n+1,n}(a, b)$. We verify this explicitly for $n = 3$ and $n = 4$. For $n \geq 5$ we use the result of the lemma that $f_{n,n}(1, 0) \geq f_{n,n}(a, b)$ for all pairs (a, b) satisfying $0 < a + b \leq n$ except $(a, b) = (1, 1)$. By the induction hypothesis, $(1, 0)$ gives a

maximum among all $f_{n+1,n-1}(a, b)$. Combining these facts shows that $f_{n+1,n}(1, 0) \geq f_{n+1,n}(a, b)$ for all pairs (a, b) satisfying $0 < a + b \leq n$ except $(a, b) = (1, 1)$. The explicit computation that $f_{n+1,n}(1, 0) = f_{n+1,n}(1, 1)$ shows that $(1, 0)$ gives a maximum for $f_{n+1,n}(a, b)$.

Now we suppose that $(1, 0)$ gives a maximum for $f_{m,n}(a, b)$. By induction $(1, 0)$ gives a maximum for $f_{m+1,n-1}(a, b)$. We conclude that $(1, 0)$ gives a maximum for $f_{m+1,n}(a, b)$, completing the proof. ■

The hypergeometric distribution, the gamma function, and *the Jetsons*

The Hypergeometric Distribution

One of the most appealing aspects of Question 3 is that it takes an elementary subject not obviously related to statistics, counting lattice paths, and leads to a fundamental discrete probability distribution.

Question 4. *Suppose there are $m + n$ students in a class, with $a + b$ girls and the rest boys. If we randomly choose m students, what is the probability that exactly a of them are girls?*

By basic counting, first choosing the girls and then choosing the boys, we see that the answer is exactly the expression $F_{m,n}(a, b)$.

This is the basis for the *hypergeometric distribution*, which is essential to understanding sampling without replacement from a finite population. Usually it is introduced in the following form. In a set of n elements, n_1 are red and the rest are black. If we choose exactly r elements at random without replacement, then the probability that exactly k of our choices are red is given by

$$\frac{\binom{n_1}{k} \cdot \binom{n-n_1}{r-k}}{\binom{n}{r}}.$$

For more of the basics of the hypergeometric distribution and some applications, see [**6**, Section II.6]. Considering more sophisticated types of lattice paths and their generating functions leads to certain hypergeometric series that have number theoretic applications in the theory of partitions [**1**].

The Gamma Function

We have emphasized finding the maximum value of $F_{m,n}(a, b)$ among all points except $(0, 0)$ and (m, n), but it is also interesting to consider a bird's-eye view of this function over its entire domain. In Figure 4 we give an example for $m = 60,\ n = 30$, where the circle at (a, b) is large if the corresponding value of $F_{60,30}(a, b)$ is large.

Points that are not close to the main diagonal $y = \frac{1}{2}x$ do not have many paths passing through them. Given a lattice path from $(0, 0)$ to (m, n) we can compute its maximum distance from this line. How large do we expect this maximum to be? We leave this and more refined statistical questions about these lattice paths to the interested reader.

Another appealing aspect of Question 3 is that it is an elementary example of an optimization problem where the correct first step is not to take a derivative and set it to zero. Attempting to go down this path does lead to interesting mathematics. Binomial coefficients are defined in terms of factorials, which are initially defined only for nonnegative integers. However, there is a natural continuous setting in which to consider

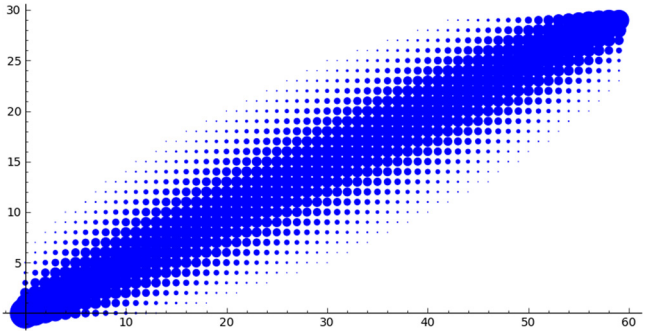


Figure 4 A plot showing relative sizes of $F_{60,30}(a, b)$, where larger points represent larger values.

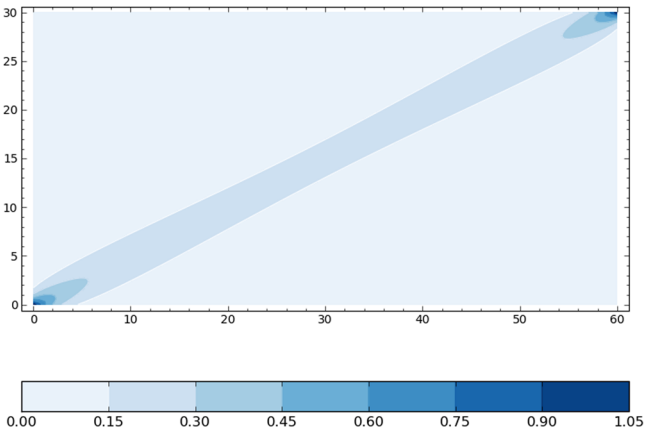


Figure 5 A plot showing the relative sizes of the continuous version of $F_{60,30}(a, b)$ defined by the gamma function.

this problem that involves the gamma function $\Gamma(t)$. This function plays an important role in complex analysis and analytic number theory. Figure 4 suggests that the discrete values of $F_{m,n}(a, b)$ can be continuously interpolated in a nice way.

The *Gamma function* is defined by

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx,$$

for all complex number t except negative integers and zero. It is a standard exercise to show that for n a positive integer, $\Gamma(n) = (n-1)!$. A continuous version of the binomial coefficient is then given by

$$\binom{x}{y} = \frac{\Gamma(x+1)}{\Gamma(y+1) \cdot \Gamma(x-y+1)}.$$

For much more on the gamma function and its role in number theory, see [**7**, Chapter 3].

In Figure 5 we give a contour plot of the continuous version of the function $F_{60,30}(a, b)$. We see the same phenomenon we saw in Figure 4. This function is small away from the line $y = \frac{1}{2}x$ and takes its largest values very close to the origin and to the point (60, 30).

Giving a continuous function that agrees with $F_{m,n}(a, b)$ suggests a way to try to find the maximum value of this function subject to the constraints of Question 3

by taking derivatives. The derivative of the gamma function is given in terms of the polygamma function. This approach leads to some more advanced complex analysis, but it is not clear that this helps us determine where to open our restaurant.

Restaurants in $\mathbb{Z}^n$

We end this paper with a generalization of Question 3 to higher dimensions. We considered all potential locations for our restaurant on a rectangular grid. Suppose we finally reach the situation promised by *The Jetsons* decades ago. You live at $(0, 0, 0)$ and take your aerocar to your office at point (a_1, a_2, a_3). You always follow the rules of space-traffic, only driving along edges of the integer lattice $\mathbb{Z}^3$. Your minimum length drive takes $a_1 + a_2 + a_3$ steps. The number of such drives is given by the trinomial coefficient $\binom{a_1+a_2+a_3}{a_1,a_2,a_3} = \frac{(a_1+a_2+a_3)!}{a_1!a_2!a_3!}$.

Question 5. *Which point in $\mathbb{Z}^3$ other than $(0, 0, 0)$ and (a_1, a_2, a_3) is visited on the maximum number of drives? That is, where is the best place to open your 3-dimensional space cantina? More generally, we can ask the same question in $\mathbb{Z}^n$. If people live at the origin and work at $(a_1, \ldots, a_n)$, where should I open my n-dimensional restaurant in order to maximize visits?*

The main ideas of the discussion at the start of Section carry over to this higher dimensional setting. The ideal location should be both "close" to the origin and not "too far" from the line connecting the origin to the endpoint $(a_1, \ldots, a_n)$. A similar approach of maximizing first over hyperplanes $\sum_{i=1}^{n} x_i = k$ and then over lines orthogonal to them seems promising. We leave this as an exercise for the interested reader.

Acknowledgments All of the graphics in this paper were made in the computer algebra system Sage. The author thanks the staff of Harvard's 2012 General Education probability and statistics course *Fat Chance*: Joe Harris, Dick Gross, and Nathan Pflueger. He thanks Scott Chapman and the two referees for helpful comments. He thanks the Fall 2013 students of Math 244 at Yale for persevering through a particularly brutal homework assignment about lattice paths.

REFERENCES

1. A. K. Agarwal, D. M. Bressoud, Lattice paths and multiple basic hypergeometric series. *Pacific J. Math.* **136** no. 2 (1989) 209–228.
2. R. Beeler, *How to Count. An Introduction to Combinatorics and its Applications*. Springer, Cham, 2015.
3. M. Bóna, *A Walk Through Combinatorics.* Third ed. World Scientific Publishing, Hackensack, NJ, 2011.
4. R. Brualdi, *Introductory Combinatorics.* Fifth ed. Pearson Prentice Hall, Upper Saddle River, NJ, 2010.
5. C. C. Chen, K. M. Koh, *Principles and Techniques in Combinatorics.* World Scientific Publishing, River Edge, NJ, 1992.
6. W. Feller, *An Introduction to Probability Theory and its Applications.* Vol. I. Third ed. John Wiley & Sons, New York 1968.
7. S. J. Miller, R. Takloo-Bighash, *An Invitation to Modern Number Theory.* Princeton Univ. Press, Princeton, NJ, 2006.
8. W. Safire, On Language—Location, Location, Location. *New York Times*, June 28, 2009: MM14.

Summary. We answer a question in real estate that is a variation of a problem that arises in many combinatorics and discrete mathematics courses.

NATHAN KAPLAN (MR Author ID: 799489) received an A.B. from Princeton University in 2007, a Certificate of Advanced Study from Cambridge University in 2008, and a Ph.D. from Harvard University in 2013. He was a Gibbs Assistant Professor at Yale from 2013 to 2015 and is currently an Assistant Professor at University of California, Irvine. He is passionate about undergraduate math research and the New York Mets.

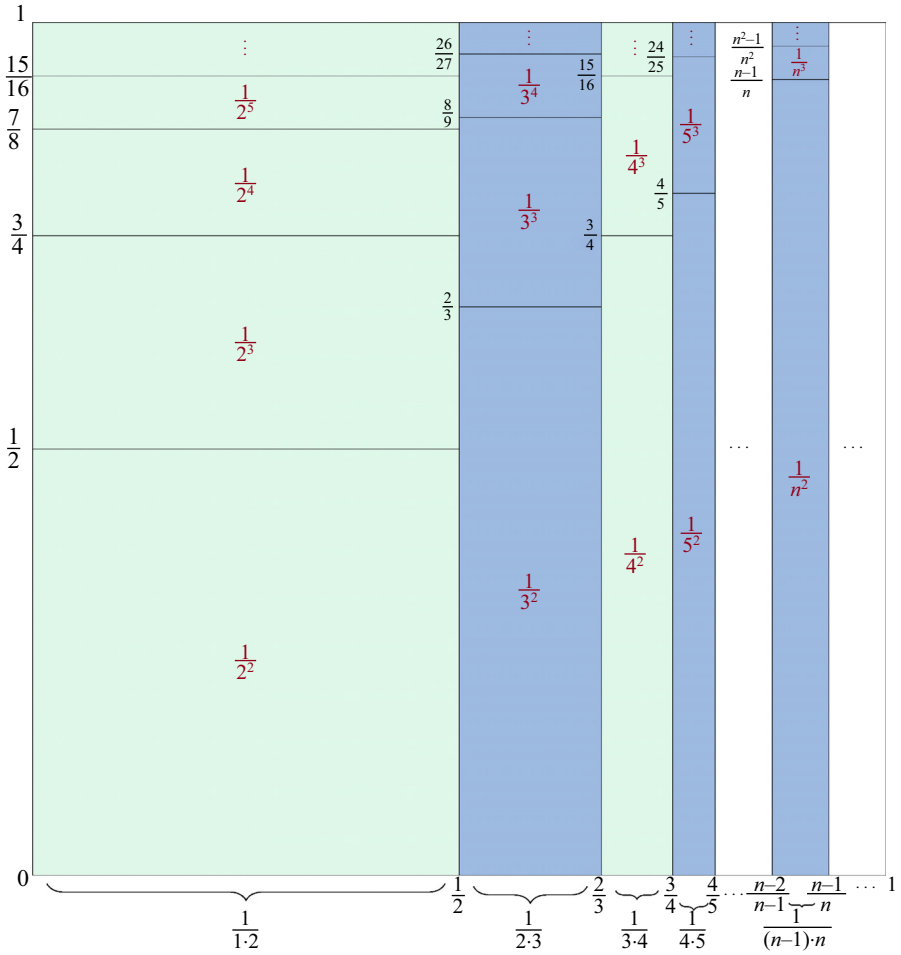
Proof Without Words: Series of Perfect Powers

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

The *multiset* of perfect powers is the collection $\mathbf{P} = \{n^m \mid n > 1;\ m > 1\}$:

$$\mathbf{P} = \{2^2, 2^3, 2^4, \ldots, 3^2, 3^3, 3^4, \ldots, 4^2, 4^3, 4^4 \ldots, 5^2, 5^3, 5^4, \ldots\}.$$

Theorem. *The sum of reciprocals of perfect powers is* 1: $\displaystyle\sum_{b\in\mathbf{P}} \frac{1}{b} = \sum_{n\geq 2}\sum_{m\geq 2} \frac{1}{n^m} = 1.$



Summary. We wordlessly show that the sum of reciprocals of perfect powers (with duplicates included) is 1.

TOM EDGAR (MR Author ID: 821633) is an associate professor of mathematics at Pacific Lutheran University in Tacoma, Washington.

Math. Mag. **90** (2017) 286. doi:10.4169/math.mag.90.4.286.
MSC: Primary 40A05

Imitating the *Shazam* App with Wavelets

EDWARD ABOUFADEL
Grand Valley State University
Allendale, Michigan
aboufade@gvsu.edu

Like searching for a needle in a haystack, suppose that we have a large set of signals (finite sequences of numbers) $\{\mathbf{s_1}, \mathbf{s_2}, \mathbf{s_3}, \dots\}$, and a special signal $\mathbf{q}$ that may or may not be in the collection. How can we find signals in the collection that are similar if not identical to $\mathbf{q}$, and *how can we do this quickly?* A solution to this question is the basis of the *Shazam* smartphone app, where a listener captures a short excerpt of a recorded song with the smartphone's microphone, and in a matter of moments the app reports the name of the song and the artist [**12**]. There the "needle" is the excerpt, and the "haystack" is a vast corpus of popular music. The *Shazam* algorithm is powered by Fourier analysis [**15**], and the purpose of this paper is to present a simpler, wavelet-based method that captures the basic process used by the app.

Solutions to this problem are useful in situations where the description of the "needle" might not be precise or may have noise in it, such as the *Shazam* problem, and where there will be frequent searches of the "haystack." For this presentation, we will use a "haystack" of comparable and accessible signals. The Jaeb Center for Health Research has made a large database of continuous glucose monitor (CGM) data available to the public*. The data comes from a recent study of type-1 diabetes (an autoimmune disorder characterized by the destruction of the islet beta cells in the pancreas by the body's own immune system) that involved 451 patients wearing a CGM for 6 or 12 months [**9**]. In Figure 1 we see an example of a "CGM day" from a type-1 diabetic patient: 288 readings of positive integers—one every five minutes.

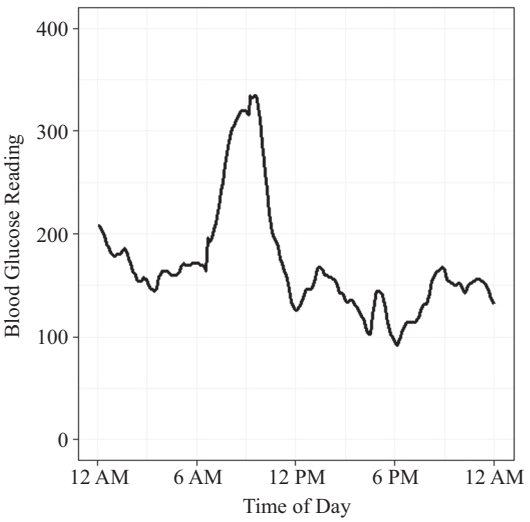


Figure 1 Example of an blood glucose daily chart.

Math. Mag. **90** (2017) 287–295. doi:10.4169/math.mag.90.4.287.
MSC: Primary 65T60

*http://diabetes.jaeb.org/RT_CGMRCTProtocol.aspx

In the following, we will describe how to take a "needle"—a CGM day labeled **q**—and find similar days in our database. The key idea—which is how the *Shazam* app works—is to create reduced, and practically unique, representations of every CGM day in the database, and then to compare these short representations rather than the original signals. Basically, there is an initial investment of time and computing power to generate short signatures for each signal in the database, and then a search can be performed quickly by comparing the signatures. To create the signatures, we will use a visualization called a *wavelet scalogram*. Wavelets are excellent tools for this type of signal reduction for CGM data, because of the rough and irregular quality of the data.

Wavelet filters

For our signal matching algorithm, wavelet filtering is an important tool. Wavelets have been applied in a wide range of areas, such as video cameras [**5**], Internet worm detection [**6**], and the design of low-power pacemakers [**8**]. Wavelets came to prominence in the late 1980s, and Ingrid Daubechies was a key researcher in their development. For good reason, there are families of wavelet filters that are named after her [**11**]. In our simplified version of the *Shazam* process, we will use the Daubechies-6 wavelet filters. Applying a *filter* to a finite time series yields another finite series. Filtering a time series may identify structures in the signal.

There are two Daubechies-6 filters: a *low-pass filter* which creates a "blur" of the input time series, and a *high-pass filter* which reveals details within the time series. The low-pass filter is a weighted average of entries in the time series, leading to a new and shorter time series that appears similar to the original. Entries in the high-pass filter output are close to 0 when the input series is nearly constant, linear, or quadratic, while other behavior in the input series shows up as larger values (in absolute value) in the high-pass output.

Given a signal $\mathbf{s} = \{s_i\}$, i from 1 to n, we take six entries at a time to compute filter outputs. For the low-pass filter [**14**], the calculation is

$$H_k = h_0 s_{2k} + h_1 s_{2k+1} + h_2 s_{2k+2} + h_3 s_{2k+3} + h_4 s_{2k+4} + h_5 s_{2k+5}, \tag{1}$$

where the filter coefficients $\{h_i\}$ are

$$\begin{aligned} &h_0 \approx 0.2352, h_1 \approx 0.5706, h_2 \approx 0.3252, \\ &h_3 \approx -0.0955, h_4 \approx -0.0604, h_5 \approx 0.02491 \end{aligned} \tag{2}$$

and k varies from 1 to $n/2$. These coefficients were derived by Daubechies so that the underlying wavelet functions satisfy certain properties: compact support, orthogonality, and regularity [**1**]. Filtering in this way creates a new signal that is half the length of the input signal. The high-pass filter calculation is

$$G_k = -h_0 s_{2k-4} + h_1 s_{2k-3} - h_2 s_{2k-2} + h_3 s_{2k-1} - h_4 s_{2k} + h_5 s_{2k+1}. \tag{3}$$

Applying these filters yield two new time series of length $n/2$. The low-pass output is a weighted moving average of entries in the original time series, while the high-pass output captures short-term changes in the time series. We call this the first scale of the analysis.

When using wavelet filters, a *pyramid scheme* is often implemented. A pyramid scheme involves applying the filters over-and-over to the low-pass output from the previous scale. So the second scale consists of two new time series of length $n/4$, the third scale has output length $n/8$, etc. There is a point where the resulting low-pass output is quite short and no further filtering is useful.

Before filtering the CGM signals, we apply a common technique known as "padding" the signals [**13**], which is extending the signal s_i in some way for $i < 1$ and/or $i > 288$. This is necessary for some of the low-pass filter calculations for k near 288 and some of the high-pass calculations for k near 1. Some ways to pad signals might be by adding 0's on the end, or extending the series as if it were periodic. (For example, if the original signal has length 28, then assign $s_{29} = s_1$, $s_{30} = s_2$, etc.) Since a CGM day is 24 hours, the periodic approach makes sense here.

When filtering each CGM day $\mathbf{q}$, the first round yields two signals of length 144. Other rounds produce signals of lengths 72, 36, and 18, and the fifth produces length 9. At this point, we no longer have an even number of entries and the low-pass output of length 9 is quite short, so we end the scheme, leaving us with low- and high-pass outputs on five different levels. In Figure 2 we see two examples of CGM days and their corresponding scale-5 "blurs."

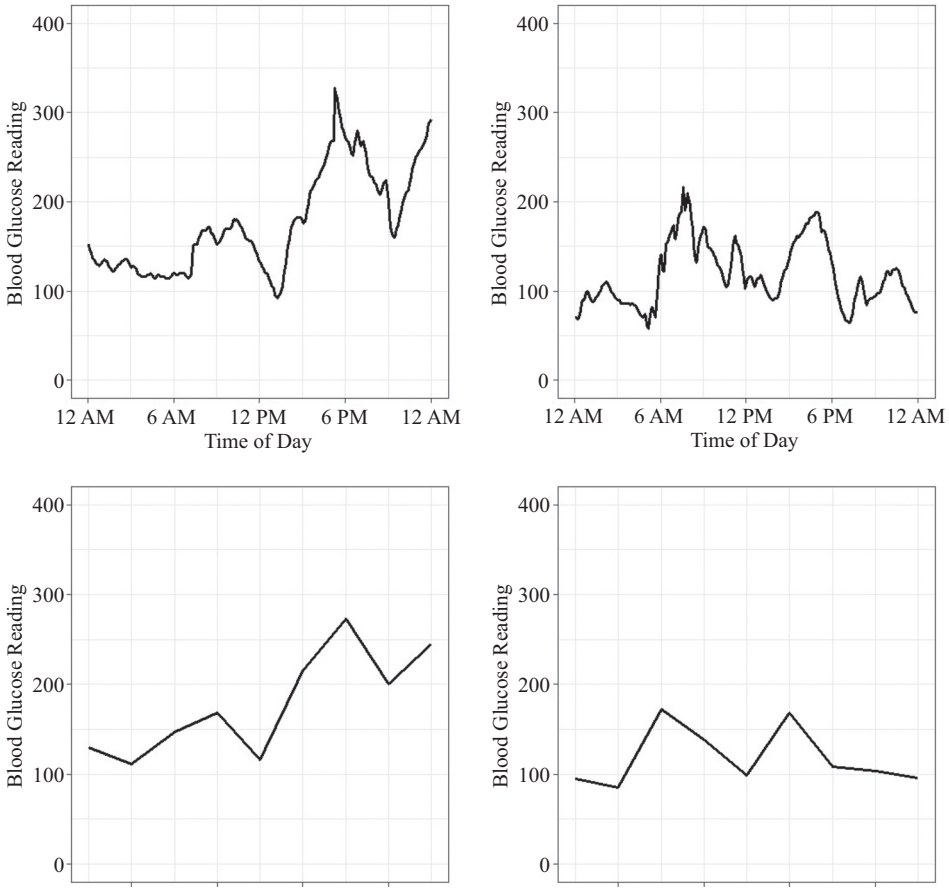


Figure 2 CGM days and scale-5 wavelet blurs.

Creating wavelet scalograms and signatures

If speed was not important, we could compare the original signals using classic distance formulas. For example, for any pair of signals of equal length $\mathbf{q} = \{q_i\}$ and $\mathbf{s} = \{s_i\}$, $i = 1, \ldots, n$, there are several distances, or norms, that we can calculate between these signals, such as the 1-norm, the 2-norm, or the infinity norm (or sup norm):

$$\|\mathbf{q}-\mathbf{s}\|_1 = \sum_{i=1}^{n} |q_i - s_i|,$$

$$\|\mathbf{q}-\mathbf{s}\|_2 = \sqrt{\sum_{i=1}^{n} (q_i - s_i)^2},$$

$$\|\mathbf{q}-\mathbf{s}\|_\infty = \max_{1\le i\le n} |q_i - s_i|.$$

However, to perform a signal match *quickly* is important, and for larger n and a large database of signals, these calculations can take considerable time. In this section we will describe how to reduce the original CGM signals to shorter signatures that can be compared instead.

A *wavelet scalogram* is a visual representation of the high-pass filter output that highlights the most dramatic changes in the signal. For our method, for each CGM day we create a scalogram as follows: We start with the five scales of analysis that are calculated from a CGM signal, giving us 279 high-pass outputs (from scales of length 144, 72, 36, 18, and 9). After identifying the 16 largest and 16 smallest (most negative) entries. we replace those largest entries with the tag $+1$, the smallest entries with the tag -1, and the remaining 263 entries with the tag 0. This is a lossy process (it cannot be reversed) and it is an example of *percentile thresholding*, which is often applied in signal processing to filter output.

After thresholding, the scalogram is created by using our time of day as the x-axis and scale as the y-axis. At each scale we populate a row of rectangles, one for each high-pass output entry, and color the rectangles black for a top-16 entry (a tag of $+1$), gray for the bottom-16 entries (a tag of -1), and white for the rest of the entries. When comparing the scalogram to the original time series, one can see that, in general, black and gray areas of the scalogram correspond to places of interesting behavior in the time series. Black areas usually correspond to where blood glucose is rising, while gray is where blood glucose is falling, and we might get some black or gray rectangles on the left or right end due to the periodic padding of the signal. The scale indicates the length of time of the behavior, with the larger scale capturing behavior over longer periods of time. Two examples of scalograms can be found in Figure 3.

From each scalogram we can create a 71-item vector for each day that we will call a *signature*. The first nine entries will be the nine low-pass output values on the fifth scale, each rounded to the nearest integer. The remaining entries are the tags from the fifth, fourth, and then third scales. In Figure 3, the left example is* 269-36827, and here is the signature for this CGM day, divided up as it was created:

[67, 79, 133, 63, 92, 165, 191, 148, 155;

$-1, 1, -1, 0, 1, 1, -1, 1, 1$;

$0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, -1, -1, 0, -1, 1, 1, 0$;

$0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0,$

$-1, 0, -1, 0, -1, 0]$.

In this way we can represent each CGM signal $\mathbf{q}$ (of length 288) by a shorter, structured signature $\tilde{\mathbf{q}}$ of nine positive integers and 68 tags from the set $\{-1, 0, 1\}$. Once these signatures are calculated, we will use them instead of the original signals to compare CGM days.

*given by patient number in the Jaeb database and "Excel Day," where January 1, 1900 at 12 midnight is "1."

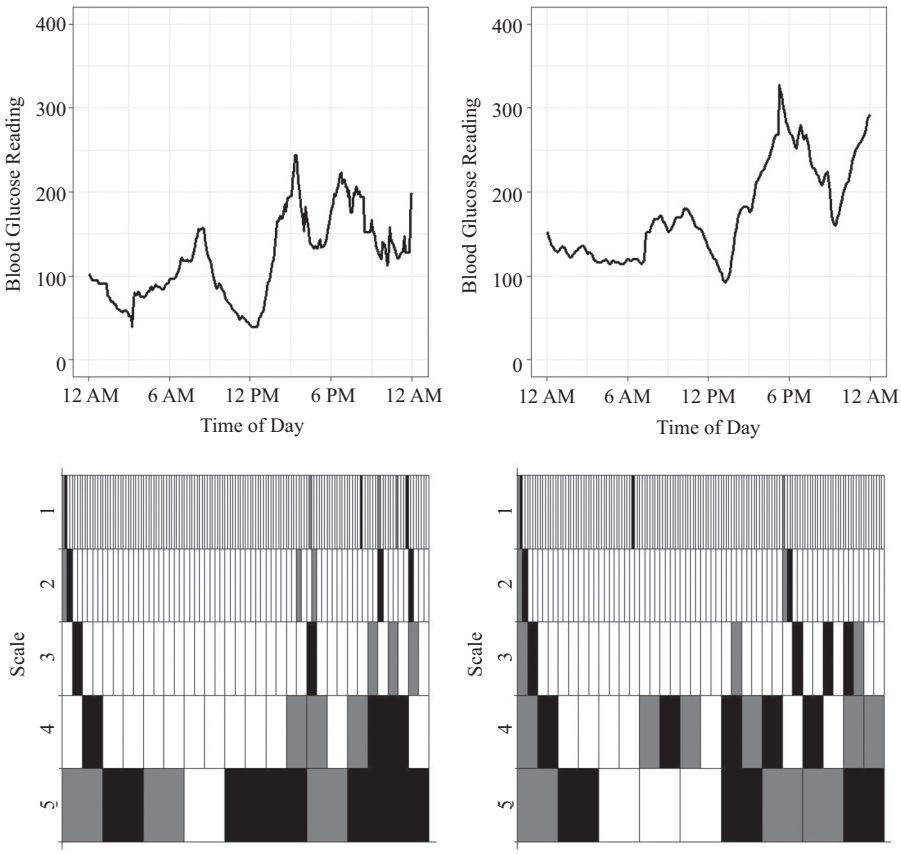


Figure 3 CGM days and scalograms.

Matching CGM days

In this section we will define a measure of similarity S between two signatures $\tilde{\mathbf{q}}$ and $\tilde{\mathbf{s}}$. We will consider these signatures (and the corresponding CGM days) to be "similar" if the measure $S(\tilde{\mathbf{q}}, \tilde{\mathbf{s}})$ is sufficiently small. Starting with a test signal, we will be satisfied with matches that twist and turn in ways like the test signal, and which possess glucose values that are close to the test signal's values during most periods of the day. Figure 4 contains a few examples of what we have in mind, with days that are similar to test day 337-36931.

Our measure S will be defined by calculating a "penalty" that comes from comparing the "high" parts of the signatures, and then a separate penalty by comparing the "low" parts. The "high" parts of the signatures can be used to find signals with coincidental twists and turns. The basic idea is that the more the -1 and $+1$ tags match in the two signatures, the more the signals will have similar shape. We will quantify this by calculating penalties when the tags do not match, giving greater weight to the entries in the fifth scale and fourth scale (the two bottom rows of the spectogram). For signatures $\tilde{\mathbf{q}}$ and $\tilde{\mathbf{s}}$, we define $h(\tilde{\mathbf{q}}, \tilde{\mathbf{s}})$ as follows: for each pair of corresponding entries in the "high" parts of the signatures, we assign the following penalties:

	-1	0	1
-1	0	6	12
0	6	3	6
1	12	6	0

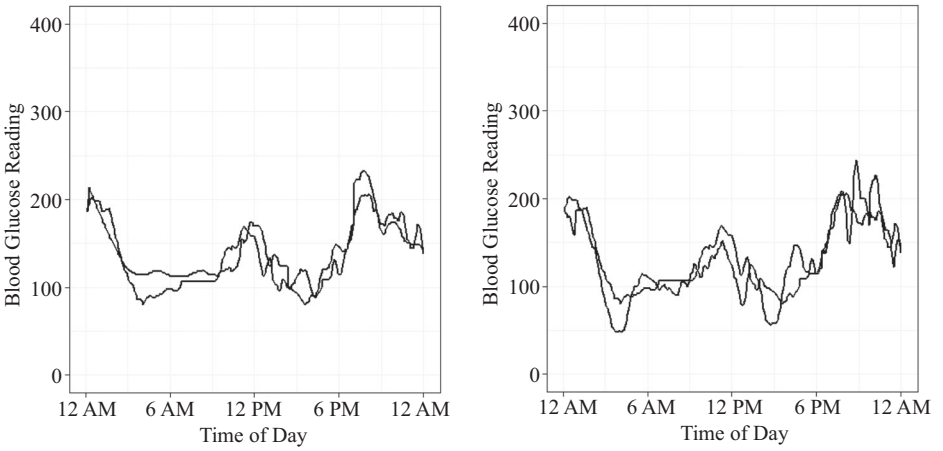


Figure 4 Matched CGM days with 337-36931.

Additionally, for the fourth scale, we double the penalties, and for the the fifth scale, they are quadrupled. Then all penalties are added together to compute h. Because of the uncertainty of the 0 tag, there are penalties whenever it appears, but there is a smaller penalty for a pair of 0's.

If we try using h alone as our similarity measure, matches of a test signal will have comparable ups-and-downs, but can be quite distant from each other vertically. For instance, a CGM signal and a second signal created by adding 100 to the first signal would match perfectly, but this is not what we have in mind for similarity. For this reason, the "low" parts of the signatures must also be taken into account.

A natural approach to compare "low" entries would be to use vector space norms such as the 1-norm, the 2-norm, or the infinity norm that were mentioned above. But each of these three norms are flawed when used to compare signatures. To decide how to use the "low" entries, a discussion of what we will desire to identify as "similar" is necessary.

If two CGM days appeared to follow the same history for most of the 24 hours and diverged for an hour or two, but not in an extreme way, then that would be acceptable as "similar." Consequently, if two signals agreed more or less on seven or eight of their nine low entries, while on the other entries the differences were not extreme, and h for this pair was small, then we would want our similarity measure between these two days to be small. Also, for type-1 diabetes, all differences between CGM signals are not the same. First of all, any reading above 240 is considered a "high," and the target of the type-1 diabetic patient and any caregiver is to maintain blood glucose numbers between 70 and 140 mg/dL. It is then the case that a difference of 10 between 75 and 65 is the difference between "in range" and "low," and this difference is just as, if not more important than, the difference between 65 and 55. In turn, both of these differences are much more important than between 310 and 300.

Consequently, if we think of using these norms to define penalties between CGM days that are "not similar," the use of any one of these norms is problematic. For both the "2-norm" and the "infinity norm," there is a high penalty for isolated corresponding entries that are significantly different. So for two signals that agree for nearly their whole length but differ significantly in one small period of time, the calculation of either of these norms can yield relatively large values. The "1-norm" will take into account all nine differences between the two "low" signatures, but unless we weight the individual differences based on the value of "low" entries, we will either over-penalize or under-penalize certain differences.

Rather than modify the "1-norm," we propose using a combination of these three norms as part of our similarity measure, favoring pairs of CGM days that are assessed

low penalties from all three norms. Of the nine differences between the "low" entries, pairs of signals where the largest of the nine differences are relatively small, and the average of the nine differences are also relatively small, will have a similarity measure which is relatively small. With this in mind, we define the following similarity measure between two different CGM signatures $\tilde{\mathbf{q}}$ and $\tilde{\mathbf{s}}$:

$$S(\tilde{\mathbf{q}}, \tilde{\mathbf{s}}) = 0.01\|\tilde{\mathbf{q}} - \tilde{\mathbf{s}}\|_1 + 0.02\|\tilde{\mathbf{q}} - \tilde{\mathbf{s}}\|_2 + 0.04\|\tilde{\mathbf{q}} - \tilde{\mathbf{s}}\|_\infty + 0.01h(\tilde{\mathbf{q}}, \tilde{\mathbf{s}}), \qquad (4)$$

where the first three norms are calculated on the first nine entries of the signatures, and h applies to the rest of the entries. We also set $S(\tilde{\mathbf{q}}, \tilde{\mathbf{q}}) = 0$, for all $\tilde{\mathbf{q}}$, since $h(\tilde{\mathbf{q}}, \tilde{\mathbf{q}}) \neq 0$. Small coefficients are included in each part of the formula for S to give our measure of similarity a reasonable size when applied to CGM data, and to help with potential future computer calculations (e.g., dealing with a matrix of similarity measures). The coefficients were determined through trial-and-error by visually studying the matches that were identified for selected test signals.

We now have the pieces needed to create a system for matching CGM signals: First, there is a one-time calculation of signatures for all signals in the database. Once this investment in calculation is complete, it is then relatively quick to find matches for our test signal by using the signatures and S. It is this speed of matching that is the rationale for our proposed method. It would be simpler, but significantly longer in time, to just compare all the original signals using one of the three norms above, or some combination of those norms.

Figure 5 shows a second set of examples of matching a test day, in this case day 59-36688 in the Jaeb database. Similar days are identified by values of S less than 10.

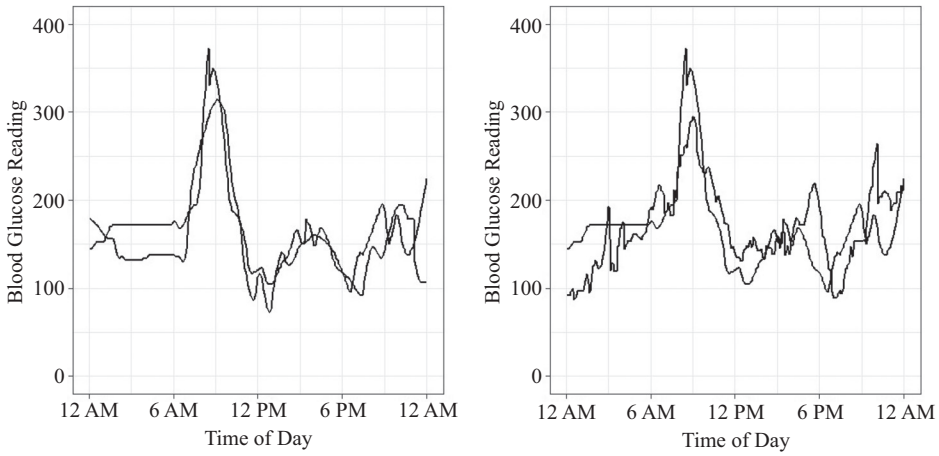


Figure 5 Matched CGM days with 59-36688.

Further discussion

How might this method of matching CGM days be used in practice with patients, caregivers, and doctors? Currently a patient can use CGM data to fine-tune short-term blood glucose management: determining at any time if extra insulin is needed to "correct a high," or if a "fix" of extra carbohydrates is required for low blood sugar. For long-term management, a blood test known as A1C* is done every three months, providing a measure of average blood glucose over the past six months. However, the A1C test doesn't capture issues of range and volatility of CGM readings.

*glycated hemoglobin is known as HbA1C, or just A1C.

One idea is to use the similarity measure S to compare the days of a patient with "exemplary" days—near-perfect glucose days where all CGM readings are between 70 and 140 and there is very little change from one reading to the next. With such a comparison, a patient and treatment team can analyze how close the patient's days are to these "exemplary" days, and this may shed light on issues of range and volatility. The more that is understood about a patient's blood glucose dynamics, the better that variables such as medication, food, and activity can be adjusted [**4, 10**].

Specifically, it might be enlightening to find days with large S values compared to "exemplary" days. An analysis like this can be revealing for some patients, such as patient 484 in the Jaeb database. For this patient, the reported A1C in the Jaeb data files was 6.6%, a very good reading, and it corresponds to an average blood glucose of around 150. However, a study of the patient's CGM data using our similarity measure S shows fewer exemplary days than would be expected. Graphs of individual days, such as 484-37036 and 484-37171 (see Figure 6), suggest that for this patient, there are many highs and lows of short duration that an A1C score would not detect. With this understanding, treatment plans could be adjusted to address these issues of volatility and range.

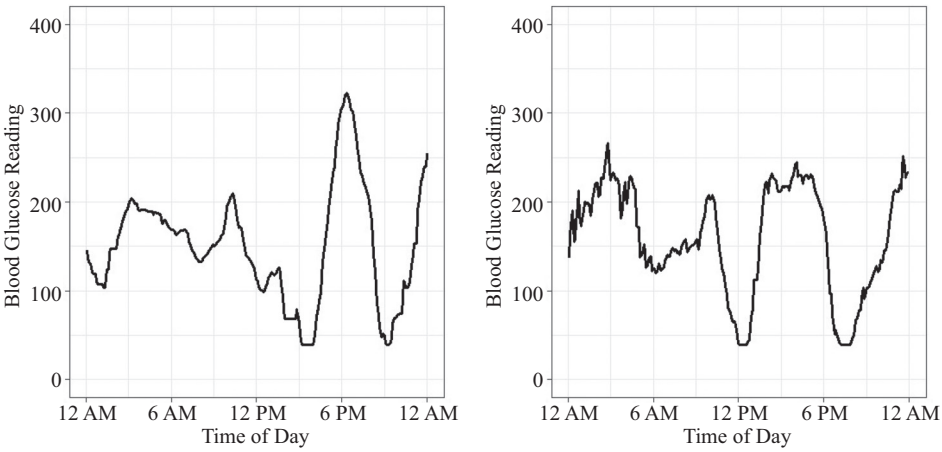


Figure 6 Patient 484: Two days with significant highs and lows, but a good average.

Returning to the *Shazam* app, the methods used to match an audio clip to a song in its database are similar to what has been described in this paper, but more complicated. For instance, Fourier analysis is used instead of wavelet filters, which leads to the creation of *spectograms* rather than scalograms. The signature/similarity measured calculations also involve finding peak values in the spectrogram, but in a significantly different way. Also, modifications are needed to take into account that only a short segment of a song is the test signal, rather than the whole song. Details can be found in [**15**]. The scalograms that we created are based on standard wavelet scalograms [**2**] and the use of wavelet scalograms with audio signals is demonstrated in [**3**]. That paper also incorporates *quantization*, which is a variant on the use of tags that is described in this paper.

Finally, the *Shazam* algorithm and the methods in this paper are part of the field of data science, which has seen massive growth in the past few years. Advances have been fueled by the growing availability of large data sets as sensors and storage devices have gotten smaller and cheaper, and more legacy data is made freely available on the Internet. New innovations are anticipated over the next 10–20 years [**7**], and there are many parts of upper-level collegiate mathematics that can open doors to this discipline.

Appendix: Preprocessing the CGM data

To create our set of CGM signals to search through, we only used the days in the Jaeb database which had at least 209 out of 288 possible CGM readings. This led to 42,799 days coming from 346 patients. For the days that were used, when CGM readings were missing, the previous known reading was used as a reasonable replacement. So the database that we used consisted of 42,799 days, each with 288 CGM positive integer readings.

Acknowledgment The author would like to thank Walter Stromquist, former editor of *Mathematics Magazine*, for his advice to revise early drafts of this paper. The author is also indebted to anonymous reviewers for careful and, at times, challenging critiques of earlier drafts of the manuscript.

REFERENCES

1. E. Aboufadel, S. Schlicker, *Discovering Wavelets*. Wiley, New York, 1999.
2. P. Addison, *The Illustrated Wavelet Transform Handbook*. Institute of Physics Publishing, Philadelphia, PA 2002, doi:10.1201/9781420033397.
3. S. Baluja, M. Covell, Content fingerprinting using wavelets. *3rd European Conference on Visual Media Production CVMP 2006*, 198–207, doi:10.1049/cp:20061964.
4. R. Beaser, *Joslin's Diabetes Deskbook: A Guide for Primary Care Providers.* Second Ed. Joslin Diabetes Center, Boston, MA, 2010.
5. F. Caimi, D. Kocak, F. Dalgleish, J. Watson, Underwater imaging and optics: Recent advances, *Oceans 2008* (2008) 1–9, doi:10.1109/OCEANS.2008.5152118.
6. B. Chen, B.-X. Fang, X.-C. Yun, Wavelet analysis based worm attack early detection, *Int. J. Comp. Sci. Eng. Sys.* **1** no. 3 (2007) 175–186.
7. A. Gandomi, M. Haider, Beyond the hype: Big data concepts, methods, and analytics, *Int. J. Inform. Manag.* **35** no. 2 (2015) 137–144, doi:10.1016/j.ijinfomgt.2014.10.007.
8. S. Haddad, W. Serdijn, *Ultra Low-Power Biomedical Signal Processing: An Analog Wavelet Filter Approach for Pacemakers*. Springer, New York, 2009, doi:10.1007/978-1-4020-9073-8.
9. JDRF CGM Study Group, JDRF randomized clinical trial to assess the efficacy of real-time continuous glucose monitoring in the management of type 1 diabetes: Research design and methods, *Diabetes Technol. Therapeutics* **10** no. 4. (2008) 310–321.
10. JDRF Research Team, JDRF Experts Weigh in on Better Metrics for Blood-Glucose Control, http://jdrf.org/blog/2013/jdrf-experts-weigh-in-on-better/.
11. W. Lawton, Applications of complex valued wavelet transforms to subband decomposition, *Signal Proc.* **41** no. 12 (1993) 3566–3568.
12. J. Medeiros, How Shazam pivoted its business to help US television broadcasters, *Wired UK* **19** no. 10 (2011).
13. J. Trygg, S. Wold, PLS regression on wavelet compressed NIR spectra, *Chemometrics Intell. Lab. Syst.* **42** no. 1–2 (1998) 209–220, doi:10.1016/S0169-7439(98)00013-6.
14. P. VanFleet, *Discrete Wavelet Transforms: An Elementary Approach with Applications*. Wiley-Interscience, New York, 2008.
15. A. Li-Chun Wang, An industrial-strength audio search algorithm, in *Proceedings of the International Conference on Music Information Retrieval ISMIR* (2000) 7–13, doi:10.1109/IITAW.2009.110.

Summary. With the *Shazam* smartphone app, a listener captures a short excerpt of a recorded song with the smartphone's microphone, and in a matter of moments the app reports the name of the song and the artist. Fourier analysis is a key mathematical tool that powers the app. In this paper, we describe a wavelet-based method that captures the basic process used by the Shazam app to search a database of number sequences (signals) to find those that are similar to a test signal. We will describe our implementation with a different source of signals: continuous glucose monitor data from the management of type-1 diabetes.

EDWARD ABOUFADEL (MR Author ID: 613192) is a Professor of Mathematics at Grand Valley State University, and an Assistant Vice President for Academic Affairs. He is an applied mathematician who has worked with students and other collaborators on projects such as detecting potholes with smartphones, using wavelets to read CAPTCHAs, and designing objects for 3D printing.

ACROSS

1. Marina spot
5. Makes a mistake
9. Italian of finite plane fame
13. $y = 2x + 11$, e.g.
14. Crouch
15. Fe
16. End of a Christian prayer
17. Daddy's sister
18. Complexion trouble for teenagers
19. Euclid's fifth was controversial
21. Soot
22. Appropriately
23. Mushroom spore sacs, or nearly a 2^7-character set
24. Slow musical pieces
27. Sets of Hilbert and Birkhoff
29. Pythagorean theater
30. She is a problem to be solved, in song
32. Former EU trade org.
34. The whole is greater ___ the part
35. Loses color
36. Small whirlpool
37. Male heir
38. Main circulatory stem
39. Sailboat tie-down
40. Unsquarable figure
42. Largest inland city in California
43. Check___
44. "Going off on a tangent," for example
46. Not tails
48. Face type of a right pyramid
52. Prefix with symmetry
53. Selling point
54. Vertex
55. In ___ of
56. Says, "Shorry, thass not my drink!"
57. Pythagorean attire
58. $2\mathbb{Z} + 1$
59. Understands
60. Current interdisc. acad. focus

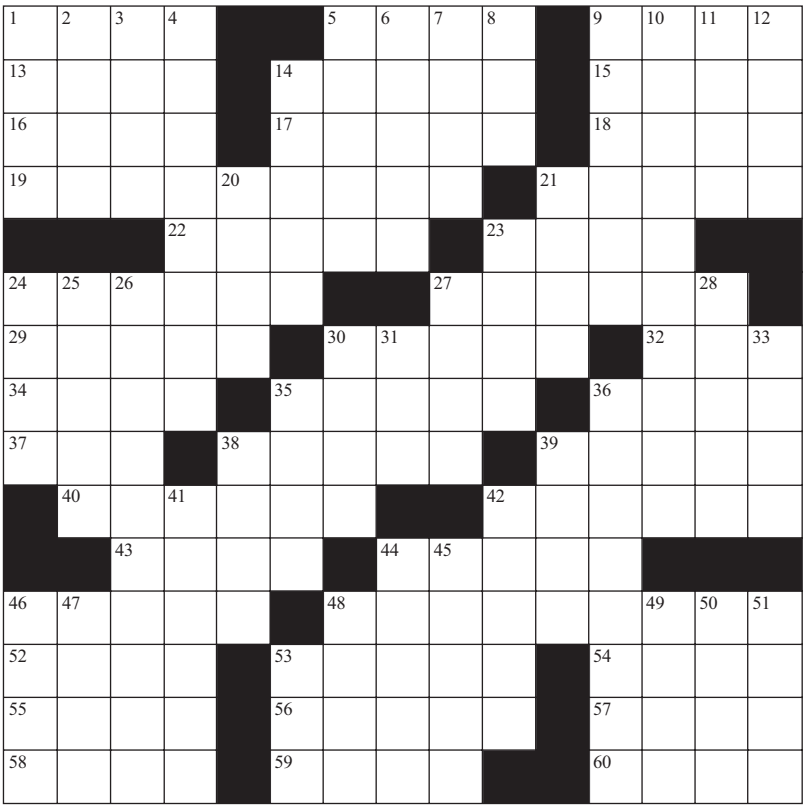
DOWN

1. Gretsky shot
2. VIP transport
3. Chem. suffixes
4. Regular figure constructible with a 72° angle
5. Like vertical angles
6. Like the smallest of the litter
7. $\frac{dy}{dx}$ or $\frac{\Delta y}{\Delta x}$
8. Hog home
9. Debacle
10. Greek mathematician with a spiral
11. $\emptyset$
12. Multiplicative identities of algebraic fields
14. Preserves, as cod
20. Once ___ a time...
21. It spans 11 time zones
23. x- or y-____
24. Oodles
25. Not standing, as a committee
26. Revived, like a zombie
27. Result of Heron's formula
28. Automobile category
30. Filly's mother
31. Ht. of a triangle
33. Prefix with -plasm
35. Singularity of a complex function
36. Euclid's 13-book classic
38. Divisions of a play
39. Critter sometimes confused with a gator
41. Bicycle spoke, for example
42. Dukes, say
44. Unit of MATHEMATICS MAGAZINE
45. These folks are not all talk
46. Xbox first-person shooter game
47. Oklahoma city west of Tulsa
48. Skye or Wight
49. Burglar's take
50. Graph component unrepeated in an Euler path
51. Line at the intersection of two pieces
53. It's often mistaken for a triangle cong. scheme

Math. Mag. **90** (2017) 296–297. doi:10.4169/math.mag.90.4.296.

Geometry

MAUREEN T. CARROLL
University of Scranton
Scranton, PA 18510
maureen.carroll@scranton.edu



Clues start at left, on page 296. The Solution is on page 266.

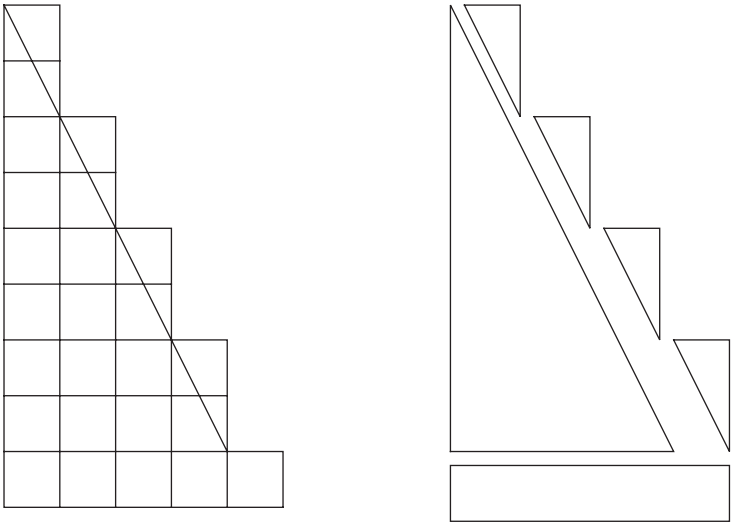
Extra copies of the puzzle can be found at the MAGAZINE's website, www.maa.org/mathmag/supplements.

Crossword Puzzle Creators

If you are interested in submitting a mathematically themed crossword puzzle for possible inclusion in MATHEMATICS MAGAZINE, please contact the editor at mathmag@maa.org.

Proof Without Words: Sums of Odd Integers

SAMUEL G. MORENO
Universidad de Jaén
23071 Jaén, Spain
samuel@ujaen.es



$$1 + 3 + 5 + \cdots + (2n-1) = n + \frac{1}{2}(n-1)(2n-2) + n - 1 = n^2.$$

REFERENCES

1. C. Alsina, R. B. Nelsen, *Charming Proofs. A Journey into Elegant Mathematics*. MAA Press, Washington DC, 2010.
2. R. B. Nelsen, *Proofs without Words: Exercises in Visual Thinking*. MAA Press, Washington DC, 1993.
3. I. Richards, Proofs without Words: Sum of Integers, *Math. Mag.* **5**7 no. 2 (1984) 104.

Summary. The number of unit-squares in the left figure equals the sum of the area of the rectangle (of length n), plus the area of the big triangle (of height $2n-2$), plus the area of the $n-1$ small 2×1 triangles.

SAMUEL G. MORENO (MR Author ID: 721728) graduated from Universidad Autónoma de Madrid (Spain), received an M.S. from Universidad Nacional de San Luis (Argentina) under the direction of Felipe Zó and did his doctoral studies at Universidad de Jaén (Southern Spain), where he is currently associate professor of mathematics. He enjoys spending time with his family, practicing all kind of sports, and spending time in the sun and at the beach when time permits.

Math. Mag. **90** (2017) 298. doi:10.4169/math.mag.90.4.298.
MSC: Primary 05A19

PROBLEMS

EDUARDO DUEÑEZ, *Editor*
University of Texas at San Antonio

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

JOSÉ A. GÓMEZ, Facultad de Ciencias, UNAM, Mexico; CODY PATTERSON, University of Texas at San Antonio; MARÍA LUISA PÉREZ-SEGUÍ, Universidad Michoacana SNH, Mexico; RICARDO A. SÁENZ, Universidad de Colima, Mexico; ROGELIO VALDEZ, Centro de Investigación en Ciencias, UAEM, Mexico; *Assistant Editors*

Proposals

To be considered for publication, solutions should be received by March 1, 2018.

2026. *Proposed by Lokman Gökçe, Turkey.*

The angles of triangle $\triangle ABC$ satisfy $\angle A = 3 \cdot \angle B$. What is the least possible perimeter of $\triangle ABC$ assuming its side lengths are integers?

2027. *Proposed by Marian Tetiva, National College "Gheorghe Roşca Codreanu", Bîrlad, Romania.*

Prove that the equation

$$n(n+1)(n+2)(n+3) = a^2 + b^2$$

admits infinitely many solutions in integers a, b, n.

2028. *Proposed by Oniciuc Gheorghe, Botosani, Romania.*

Is there a nonconstant sequence $\{a_n\}_{n\geq 1}$ in the interval $(0, 1)$ such that

$$\lim_{n\to\infty} (1 + a_1 - a_2)(1 + a_2 - a_3) \cdots (1 + a_{n-1} - a_n)(1 + a_n - a_1) = 1?$$

2029. *Proposed by Kenneth Schilling, University of Michigan-Flint, MI.*

Let $\mathbb{N}$ be the set of natural numbers. For any function $g : \mathbb{N} \to \mathbb{N}$ and a subset Y of $\mathbb{N}$, we denote by $g[Y] = \{g(y) : y \in \mathbb{N}\}$ the forward image of Y by g. Given a natural number n, a collection $\mathcal{C}$ of subsets of $\mathbb{N}$ is called:

Math. Mag. **90** (2017) 299–305. doi:10.4169/math.mag.90.4.299.

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

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

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

- *n-plenary* if there exist functions $f_1, f_2, \ldots, f_n : \mathbb{N} \to \mathbb{N}$ such that, for all $X \in \mathcal{C}$, $f_1[X] \cup f_2[X] \cup \cdots \cup f_n[X] = \mathbb{N}$.
- *strongly n-plenary* if there exist $f_1, f_2, \ldots, f_n : \mathbb{N} \to \mathbb{N}$ such that, for all $X \in \mathcal{C}$, either $f_1[X] = \mathbb{N}$, or $f_2[X] = \mathbb{N}, \ldots$, or $f_n[X] = \mathbb{N}$.

For every natural number n:

(a) Prove that a collection $\mathcal{C}$ is n-plenary if and only if it is strongly n-plenary.

(b) Construct an $(n + 1)$-plenary collection $\mathcal{C}$ that is not n-plenary.

Editor's Note. Problem 2029 was inspired by Problem 1998 in which the concept of plenary collection was introduced. The statement and solution of Problem 1998 appear under Solutions below.

2030. *Proposed by Rajesh Sharma, Himachal Pradesh University, India.*

(a) Prove that a real symmetric 3×3 matrix with equal diagonal entries has a multiple eigenvalue if and only if its nondiagonal entries all have the same absolute value.

(b) Characterize all complex hermitian 3×3 matrices with equal diagonal entries and a multiple eigenvalue.

Quickies

1073. *Proposed by M. Hajja, Amman, Jordan.*

Let $ABCD$ be a convex quadrilateral such that $AB = CD$ and $\angle D > \angle A$. Prove that $\angle B > \angle C$.

1074. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

A *derivative on* $\mathbb{R}$ is a function $f : \mathbb{R} \to \mathbb{R}$ such that f is the derivative of some function $F : \mathbb{R} \to \mathbb{R}$. Prove or disprove: Every derivative f on $\mathbb{R}$ satisfies

$$\liminf_{t \to x} f(t) \leq f(x) \leq \limsup_{t \to x} f(t) \quad \text{for all } x \in \mathbb{R}.$$

Solutions

1996. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

Compute

$$\lim_{x \to 0} \frac{1}{x} \int_0^x \left| \cos \frac{1}{t} \right| dt.$$

Solution by Michel Bataille, Rouen, France.
We show that the value of the limit is $2/\pi$.

For $x \neq 0$, let $L(x) = x^{-1} \int_0^x \left| \cos t^{-1} \right| dt$. The substitution $t \mapsto -t$ in the integral shows that the L is an even function, so it suffices to show that $\lim_{x \to 0^+} L(x) = 2/\pi$. For $x > 0$, the substitution $t \mapsto 1/t$ gives

$$L(x) = \frac{1}{x}\int_{1/x}^{\infty} \frac{|\cos t|}{t^2}\,dt.$$

The function $f(x) = |\cos x|$ is continuous, hence it has a primitive F on $\mathbb{R}$ (i.e., an antiderivative). Let $G(x) = F(x) - 2x/\pi$. For all real x,

$$G(x+\pi) - G(x) = F(x+\pi) - F(x) - \frac{2}{\pi}\cdot\pi = \int_x^{x+\pi} |\cos x|\,dx - 2$$
$$= \int_0^{\pi} |\cos x|\,dx - 2 = 0, \quad \text{since } f(x) = |\cos x| \text{ has period } \pi.$$

Thus, G is a continuous periodic function, hence bounded on $\mathbb{R}$. Let $M = \max\{|G(t)| : t \in \mathbb{R}\}$. For $Y > y > 0$, integrating by parts we obtain

$$\int_y^Y \frac{|\cos t| - \frac{2}{\pi}}{t^2}\,dt = \int_y^Y \frac{1}{t^2}\cdot G'(t)\,dt = \frac{G(Y)}{Y^2} - \frac{G(y)}{y^2} + 2\int_y^Y \frac{G(t)}{t^3}\,dt.$$

Since G is bounded and $\int_y^{\infty} t^{-3}\,dt$ converges, it is clear that $\int_y^{\infty}(|\cos t| - 2/\pi)t^{-2}\,dt$ converges (absolutely, in fact), and

$$\int_y^{\infty} \frac{|\cos t| - \frac{2}{\pi}}{t^2}\,dt = -\frac{G(y)}{y^2} + 2\int_y^{\infty} \frac{G(t)}{t^3}\,dt.$$

Now,

$$\left|L(x) - \frac{2}{\pi}\right| = \left|x^{-1}\int_{x^{-1}}^{\infty} \frac{|\cos t|}{t^2}\,dt - \frac{2}{\pi}\right| = \left|x^{-1}\int_{x^{-1}}^{\infty} \frac{|\cos t| - \frac{2}{\pi}}{t^2}\,dt\right|$$
$$= \left|-\frac{G(x^{-1})}{x^{-1}} + 2x^{-1}\int_{x^{-1}}^{\infty} \frac{G(t)}{t^3}\,dt\right|$$
$$\le Mx + 2Mx^{-1}\int_{x^{-1}}^{\infty} \frac{dt}{t^3} = 2Mx,$$

so

$$x^{-1}\int_0^x \left|\cos t^{-1}\right|\,dt = L(x) \to \frac{2}{\pi}$$

as $x \to 0^+$, hence also as $x \to 0$.

Also solved by Robert Calcaterra, Robin Chapman (UK), Dmitry Fleischman, Nilotpal Ghosh, Jan Greszik, Eugene Herman, Tom Jager, Kee-Wai Lau (Hong Kong), Missouri State University Problem Solving Group, Moubinool Omarjee (France), Northwestern University Problem Solving Group, Paolo Perfetti (Italy), Mehtaab Sawhney, Nicholas C. Singer, and the proposer. There were 2 incomplete or incorrect solutions.

The improper integral of a rational-exponential function — June 2016

1997. *Proposed by Ovidiu Furdui and Alina Sîntămărian, Technical University of Cluj-Napoca, Cluj-Napoca, Romania.*

Calculate

$$\int_0^{\infty} \left(\frac{1 - e^{-x}}{x}\right)^2 dx.$$

Solution by Travis D. Cunningham (student), Carson City, MI.

The integrand is the square of

$$\frac{1-e^{-x}}{x} = \int_0^1 e^{-tx}\,dt,$$

the Laplace transform of the characteristic function of the interval [0, 1]. Thus,

$$\begin{aligned}\int_0^\infty \left(\frac{1-e^{-x}}{x}\right)^2 dx &= \int_0^\infty \left(\int_0^1 e^{-tx}\,dt\right)\left(\int_0^1 e^{-ux}\,du\right)dx \\ &= \int_0^1\int_0^1\int_0^\infty e^{-(t+u)x}\,dx\,du\,dt = \int_0^1\int_0^1 \frac{1}{t+u}\,du\,dt \\ &= \int_0^1 \big(\ln(t+1)-\ln t\big)\,dt = \int_1^2 \ln t\,dt - \int_0^1 \ln t\,dt = (2\ln 2 - 1) - (-1) \\ &= 2\ln 2.\end{aligned}$$

(Recall the standard integral $\int_a^b \ln t\,dt = F(b) - F(a)$ for $a, b \geq 0$, where $F(t) = t \ln t - t$ for $t > 0$, and $F(0) = \lim_{t\to 0^+} F(t) = 0$.) Switching of the order of integration above is justified by Fubini's theorem since the integrand is nonnegative.

Editor's Note. Solutions submitted by readers used a variety of approaches. Bruce E. Davis from St. Louis Community College at Florissant Valley (MO) provided no fewer than four different methods, including the one above.

Also solved by Adnan Ali (student) (India), Michel Bataille (France), Khristo Boyadzhiev, Robert Calcaterra, Robin Chapman (UK), Hongwei Chen, Bruce E. Davis, Prithwijit De (India), Nilotpal Ghosh, Jan Grzesik, Jim Hartman, Eugene Herman, Finbarr Holland (Ireland), Tom Jager, Parviz Khalili, Koopa Koo (Hong Kong), Kee-Wai Lau (Hong Kong), Robert J. Lopez & Herb Bailey, John Mahony (New Zealand), Missouri State University Problem Solving Group, Moubinool Omarjee, Paolo Perfetti, Arthur J. Rosenthal & Andrew A. Barrett, Mehtaab Sawhney, Nicholas C. Singer, Seán Stewart (Australia), Nora Thornber, Utah Weber State University "P, J and R Group", Michael Vowe (Switzerland), Yuanyuan Zhao, and the proposer. There was 1 incomplete or incorrect solution.

1998. *Proposed by Greg Oman, University of Colorado, Colorado Springs, CO.*

Let $\mathbb{N}$ be the set of natural numbers. We call a collection $\mathcal{C}$ of subsets of $\mathbb{N}$ *plenary* if there exists a function $f : \mathbb{N} \to \mathbb{N}$ such that $f[X] = \mathbb{N}$ for all $X \in \mathcal{C}$, where $f[X] = \{f(x) : x \in X\}$ is the set of images of elements of X under f.

(a) Prove that any countable collection of infinite subsets of $\mathbb{N}$ is plenary.

(b) Prove that the collection of all infinite subsets of $\mathbb{N}$ is not plenary.

(c) Are there any uncountable plenary collections?

Solution by Northwestern University Math Problem Solving Group, Evanston, IL.

(a) Assume $\mathcal{C} = \{X_k\}_{k=0}^\infty$ is a countable collection of infinite subsets of $\mathbb{N}$. Let $g : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ be any bijection. (For instance, $g(p,q) = \frac{1}{2}(p+q)(p+q+1) + p$.) Call f a *partial function* $\mathbb{N} \to \mathbb{N}$ if f is a function into $\mathbb{N}$ whose domain $\mathrm{dom}(f)$ is a subset of $\mathbb{N}$. We construct a sequence $\{f_i\}_{i\in\mathbb{N}}$ of partial functions $\mathbb{N} \to \mathbb{N}$ recursively as follows. Let f_0 be the empty function $\emptyset \to \mathbb{N}$. Given $i \in \mathbb{N}$, let $(k_i, m_i) = g^{-1}(i)$. Let n_i be the least element of $X_k \setminus \mathrm{dom}(f_i) = \{j \in X_k \mid j \notin \mathrm{dom}(f_i)\}$, and let f_{i+1} extend f_i by defining $f_{i+1}(n_i) = m_i$. The process can continue indefinitely because $\mathrm{dom}(f_i)$ is clearly a finite subset of $\mathbb{N}$ (with exactly i elements in fact), so $X_k \setminus \mathrm{dom}(f_i)$ must still be infinite, hence nonempty. By the

well-ordering of $\mathbb{N}$ the element n_i above exists, ensuring the recursive construction proceeds indefinitely. Now let the (complete) function $f : \mathbb{N} \to \mathbb{N}$ be defined by $f(n) = f_i(n)$ if $n \in \mathrm{dom}(f_i)$ for some $i \in \mathbb{N}$, and $f(n) = 0$ otherwise. Clearly, f_j is a function extending f_i for $j \geq i$, so f is well defined *a fortiori*. To finish the proof, we show that $f[X_k] = \mathbb{N}$ for any given $k \in \mathbb{N}$. Let $m \in \mathbb{N}$, and let $i = g(k, m)$. Clearly, $(k, m) = (k_i, m_i)$, hence $m = m_i = f_{i+1}(n_i) = f(n_i) \in f[X_k]$ since $n_i \in X_k$. We conclude that $\mathcal{C}$ is plenary.

(b) Let $f : \mathbb{N} \to \mathbb{N}$ be arbitrary. Let $\mathbb{E}$ be the set of even, $\mathbb{O}$ the set of odd natural numbers, so $\mathbb{N}$ is the disjoint union of the infinite sets $\mathbb{E}$ and $\mathbb{O}$. Clearly, $X_0 = f^{-1}[\mathbb{E}]$ and $X_1 = f^{-1}[\mathbb{O}]$ are also disjoint sets with $X_0 \cup X_1 = \mathbb{N}$. Note that $f[X_0] \subset \mathbb{E} \subsetneq \mathbb{N}$ and also $f[X_1] \subsetneq \mathbb{N}$. Since $\mathbb{N}$ is infinite, at least one of X_0, X_1 must be infinite with image under f a proper subset of $\mathbb{N}$, so the collection of all infinite subsets of $\mathbb{N}$ is not plenary.

(c) The answer is yes. We construct one uncountable plenary collection.

Let $\mathcal{C}$ be the collection of subsets of $\mathbb{N}$ that include the set $\mathbb{E}$ of even numbers. Note that $\mathcal{C} = \{\mathbb{E} \cup Y : Y \subset \mathbb{O}\}$ with $\mathbb{O}$ the set of odd numbers; in fact, the mapping $Y \mapsto \mathbb{E} \cup Y$ is a bijection between subsets of $\mathbb{O}$ and members of $\mathcal{C}$. The countable set $\mathbb{O}$ has uncountably many subsets Y by Cantor's theorem, hence $\mathcal{C}$ is uncountable also.

To see that $\mathcal{C}$ is plenary, note that if $f : \mathbb{N} \to \mathbb{N}$ is any function such that $f(2n) = n$ for all $n \in \mathbb{N}$, then $f[X] \supseteq f[\mathbb{E}] = \mathbb{N}$, hence $f[X] = \mathbb{N}$ for all $X \in \mathcal{C}$.

Also solved by Paul Budney, Robert Calcaterra, Robin Chapman, Joseph DiMuro, Jerrold Grossman, and the proposer. There was 1 incomplete or incorrect solution.

1999. *Proposed by Mihàly Bencze, Brasov, Romania.*

For any real number $a > 1$, evaluate

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{m^2 n}{a^m(na^m + ma^n)}.$$

Solution by Adnan Ali (student), A.E.C.S-4, Mumbai, India.

The value of the sum is $S = \frac{1}{2}a^2/(a-1)^4$. The series has nonnegative terms, hence we may swap the order of summation and rename the dummy variables to get

$$S = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{mn^2}{a^n(na^m + ma^n)}.$$

Thus,

$$2S = S + S = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{mn}{na^m + ma^n}\left(\frac{m}{a^m} + \frac{n}{a^n}\right) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{mn}{a^m \cdot a^n} = \left(\sum_{n=1}^{\infty}\frac{n}{a^n}\right)^2.$$

For $|x| < 1$,

$$\sum_{n=1}^{\infty} nx^n = x \cdot \frac{d}{dx}\sum_{n=0}^{\infty} x^n = x \cdot \frac{d}{dx}\left(\frac{1}{1-x}\right) = \frac{x}{(1-x)^2},$$

whence

$$2S = \left(\sum_{n=1}^{\infty} na^{-n}\right)^2 = \left(\frac{a^{-1}}{(1-a^{-1})^2}\right)^2 = \frac{a^2}{(a-1)^4}$$

whenever $a > 1$, because $|a^{-1}| < 1$ in this case. It follows that $S = \frac{1}{2}a^2/(a-1)^4$.

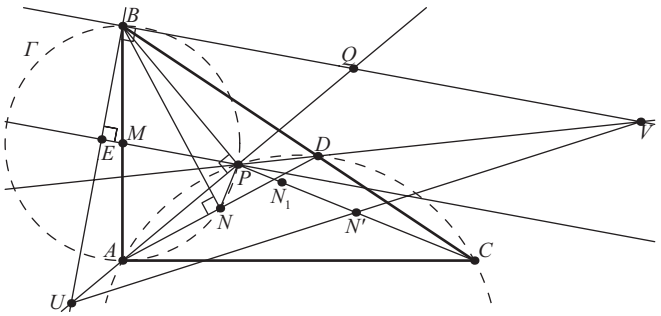
Also solved by Michel Bataille (France), Robert Calcaterra, Robin Chapman (UK), Hongwei Chen, Travis D. Cunningham, Bruce Davis, Tom Jager, Parviz Khalili, Koopa Koo (Hong Kong), James Magliano, John Mahony, Rituraj Nandan, Michael Nathanson, Northwestern University Math Problem Solving Group, Mehtaab Sawhney, Nicholas Singer, Michael Vowe (Switzerland), and the proposer.

2000. *Proposed by Michel Bataille, Rouen, France.*

Let $\triangle ABC$ have a right angle at A. Let M be the midpoint of AB, let D lie on side $\overline{BC}$ so $BD = BA$, and let P lie on the circumcircle of $\triangle ADC$ so that $\angle APB = 90^\circ$. Let U lie on line $\overleftrightarrow{AP}$ so that $\overline{BU}$ is perpendicular to $\overline{MP}$, and let V lie on $\overleftrightarrow{DP}$ so that $\overline{BV}$ is parallel to $\overline{MP}$.

Prove that $PU/PV = BU/BV$, and the line $\overleftrightarrow{CP}$ bisects $\overline{UV}$.

Solution by Peter McPolin, St. Mary's University College, Belfast, UK.



Let Γ be the circle with diameter $\overline{AB}$ and let E be the point of intersection of the line $\overleftrightarrow{MP}$ with its perpendicular line $\overleftrightarrow{BU}$. Since $\overline{AP} \perp \overline{PB}$ and $\overline{PE} \perp \overline{BU}$ by construction, the point P lies on Γ and $\angle BPE = \angle PUB$; moreover, $\angle BPE = \angle BPM = \angle MBP = \angle ABP$ because $\triangle BMP$ is isosceles. We conclude that we have a similarity $\triangle BPU \sim \triangle APB$ between right triangles.

Next, let $\alpha = \angle PAB = \angle UBP$, $\beta = \angle DAP$, and $\gamma = \angle PDA$. Since A, P, D, C are concyclic, $\angle DCP = \angle DAP = \beta$ and $\angle PCA = \angle PDA = \gamma$, hence $\angle BCA = \angle DCA = \angle DCP + \angle PCA = \beta + \gamma$. Let Q be the point of intersection of lines $\overleftrightarrow{PA}$ and $\overleftrightarrow{BV}$. Since $\overline{BQ} \parallel \overline{PM} \perp \overline{BU}$ and $\overline{QU} \perp \overline{BP}$, we have $\angle BQU = \angle UBP = \alpha$.

Angle $\angle BQU = \angle BQP$ is exterior to triangle $\triangle QPV$, so $\alpha = \angle BQP = \angle VPQ + \angle QVP$. Similarly, $\angle VPQ$ is exterior to $\triangle PDA$, so $\angle VPQ = \beta + \gamma$. Thus, $\angle BVP = \angle QVP = \alpha - (\beta + \gamma)$. Since $\triangle BAC$ has a right angle at A, $\angle ABC = 90^\circ - \angle BCA = 90^\circ - (\beta + \gamma)$. We now have, $\angle PBD = \angle ABC - \angle ABP = \angle ABC - (90^\circ - \angle PAB) = (90^\circ - (\beta + \gamma)) - (90^\circ - \alpha) = \alpha - (\beta + \gamma) = \angle BVP$. The similarity $\triangle VPB \sim \triangle BPD$ follows since these triangles have two equal angles.

Now we prove that $PU/PV = BU/BV$. From $\triangle APB \sim \triangle BPU$, we get $PU/PB = BU/BA$; from $\triangle BPD \sim \triangle VPB$, we get $PV/PB = BV/BD = BV/BA$ (since $BD = BA$ by construction). Combining these ratios gives $PU/PV = PU/PB \cdot PB/PV = BU/BA \cdot BA/BV = BU/BV$ as asserted.

We have $\angle DCP = \angle DAP$ in the cyclic quadrilateral $APDC$. Let N be the midpoint of $\overline{AD}$. Since $\triangle ABD$ is isosceles, BN is perpendicular to AD and $\angle BNA$ is a right angle. It follows that the quadrilateral $ANPB$ is also cyclic and $\angle PNB = \angle PAB = \angle UBP$. The similarity $\triangle VPB \sim \triangle BPD$ gives $\angle BDP = \angle PBV = \angle UBV - \angle UBP = 90^\circ - \angle UBP$. Thus, $\angle PDC = 180^\circ - \angle BDP = 180^\circ - (90^\circ - \angle UBP) = 90^\circ + \angle UBP = 90^\circ + \angle PNB = \angle PNA$, hence, $\angle PDC = \angle PNA$. Since triangles $\triangle ANP$, $\triangle CDP$

have two equal angles $\angle NAP = \angle DAP = \angle DCP$ and $\angle PNA = \angle PDC$, their third angle must also be equal: $\angle APN = \angle CPD$. It follows that the reflection N_1 of N in the angle bisector of $\angle APD$ must lie on the line $\overleftrightarrow{CP}$.

Next, we show that $\triangle APD \sim \triangle VPU$. From $\triangle BPU \sim \triangle APB$ we get $PA/PB = PB/PU$, and from $\triangle VPB \sim \triangle BPD$ we get $PD/PB = PB/PV$. Thus, $PD/PA = PD/PB \cdot PB/PA = PB/PV \cdot PU/PB = PU/PV$. The similarity $\triangle APD \sim \triangle VPU$ follows by SAS.

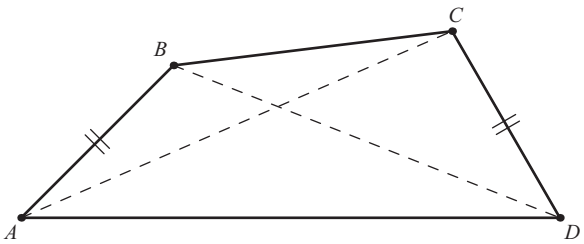
The transformation of $\triangle APD$ into $\triangle VPU$ is achieved by reflecting $\triangle APD$ in the angle bisector of the common angle at P (i.e., of $\angle APD = \angle UPV$), then dilating with center P by factor PU/PD. This transformation maps the midpoint N of $\overline{AD}$ to the midpoint N' of $\overline{UV}$, which is precisely the dilation of N_1 with center P and factor PU/PD. Since N_1 lies on $\overleftrightarrow{CP}$ so does N', concluding the proof.

Also solved by Herb Bailey, Robert Calcaterra, Robin Chapman (UK), Andrea Fanchini (Italy), Ioana Mihaila, and the proposer. There was 1 incomplete or incorrect solution.

Answers

Solutions to the Quickies from page 300.

A1073.



This is one of the fairly rare applications of Propositions 24 and 25 of Book I of Euclid's Elements, sometimes nicknamed the *open mouth theorems*. Taken jointly, these theorems state that if $\triangle ABC$ and $\triangle A'B'C'$ are two triangles in which $AB = A'B'$ and $AC = A'C'$, then $BC > B'C'$ if and only if $\angle A > \angle A'$. (The "if" statement is Proposition 24; the "only if" is Proposition 25.) Thus, in triangles $\triangle DAC$ and $\triangle ADB$ we have $AC > DB$; in triangles $\triangle BCA$ and $\triangle CBD$ we have $\angle B > \angle C$ as desired.

A1074. We prove the statement. Let f be the derivative of $F : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \mathbb{R}$. By the mean value theorem, for any positive integer n there exists $x_n \in (x_0, x_0 + 1/n)$ such that

$$\frac{F(x_0 + 1/n) - F(x_0)}{1/n} = F'(x_n) = f(x_n).$$

Thus,

$$f(x_0) = F'(x_0) = \lim_{n\to\infty} \frac{F(x_0 + 1/n) - F(x_0)}{1/n} = \lim_{n\to\infty} f(x_n).$$

Since $x_n \to x_0$ as $n \to \infty$ (and $x_n \neq x_0$ for all n), $f(x_0) = \lim_{n\to\infty} f(x_n)$ is a cluster value of f at x_0, hence $\liminf_{t\to x_0} f(t) \leq f(x_0) \leq \limsup_{t\to x_0} f(t)$. As x_0 was arbitrary, this concludes the proof.

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.

Knudson, Kevin, There are only 15 pentagonal tilings (probably), *Forbes* (26 June 2017), https://www.forbes.com/sites/kevinknudson/2017/06/26/there-are-only-15-pentagonal-tilings-probably/#338258b54cb6.

Wolchover, Natalie, Pentagon tiling proof solves century-old math problem, https://www.quantamagazine.org/pentagon-tiling-proof-solves-century-old-math-problem-20170711/.

Wolchover, Natalie, Marjorie Rice's secret pentagons, https://www.quantamagazine.org/marjorie-rices-secret-pentagons-20170711/.

Lipton, R. J., and K. W. Regan, Kitchen tile catalog complete, https://rjlipton.wordpress.com/2017/07/16/kitchen-tile-catalog-complete/.

Rao, Michaël, Exhaustive search of convex pentagons which tile the plane, https://perso.ens-lyon.fr/michael.rao/publi/penta.pdf. Presentation slides at http://www.crm.umontreal.ca/2017/Pavages17/pdf/Rao1_slides.pdf.

Amateur mathematician Marjorie Rice, who discovered in the 1970s four previously unknown families of tessellations by pentagons, died in July. Dementia prevented her from learning of a proof a few months earlier that there are only 15 such families, by Michaël Rao (CNRS and École Normale Supérieure de Lyon). Rao concludes his lecture slides with further work to do: recheck the proof, recheck the code (≈5,000 lines in C++), reproduce the exhaustive search, find a formal proof.

Willemain, Thomas Reed, *Working on the Dark Side of the Moon: Life Inside the National Security Agency*, Mill City Press, Maitland, FL; xii+119 pp, $15.99(P). ISBN 978-1-62952-872-4.

Author Willemain, "an electrical engineer turned applied statistician and software entrepreneur," was a sabbatical visitor on the Mathematics Research staff of the National Security Agency (NSA) and later at the Institute for Defense Analysis. This book describes day-to-day life in "a dream environment for a professor on sabbatical. . . the best of the academic world without some of its drawbacks." The writing is breezy, the sentiments are sincere, and his enthusiasm for the people and the mission of the NSA is palpable. The book was cleared by the NSA (redacted passages are replaced by black bars); the author offers a Webpage that gives an account of his prepublication experience.

Rougetet, Lisa, The prehistory of Nim game [sic] (36 pp). In *G4G11 Gift Exchange Book*. www.gathering4gardner.org/g4g11gift/Rougetet_Lisa-Prehistory_of_Nim.pdf.

This paper traces description of various Nim-type games back, to Luca Pacioli around 1500, and forward, through various European works on recreational mathematics. The paper concludes with description of a West-African grid-pattern blocking game, Tiouk-Tiouk, that is analogous to Nim.

Math. Mag. **90** (2017) 306–307, doi:10.4169/math.mag.90.4.306.

Pitici, Mircea, *The Best Writing on Mathematics 2016*, Princeton University Press, 2017; xxii+377 pp, $32.95(P). ISBN 978-0-691-17529-4.

This year's compendium of rich and remarkable mathematical writing offers the added pleasure of color illustrations for many of its pieces. Notable articles are a spirited defense of "strong" non-Platonism (Derek Abbott), a curious fact about stacking wine bottles (Burkard Polster), news of proof of the Umbral Moonshine Conjecture (Erica Klarreich), mathematical objects at the Metropolitan Museum of Art in New York (Joseph Dauben and Marjorie Senechal), and what depth may mean in mathematics (John Stillwell).

Nadis, Steve, The universe according to Emmy Noether: How a deceptively simple theorem from a century ago still shapes modern physics, *Discover* 38 (5) (June 2017) 49–53.

Neuenschwander, Dwight E., *Emmy Noether's Wonderful Theorem*, revised and updated edition, Johns Hopkins University Press, 2017; xvi+317 pp, $30(P). ISBN 978-1-4214-42267-1.

"[E]very 'continuous' symmetry in nature has a corresponding conservation law, and vice versa. . . . The theorem provides an explicit mathematical formula for finding the symmetry that underlies a given conservation law and, conversely, finding the conservation law that corresponds to a given symmetry." So author Nadis summarizes Noether's theorem, explaining that it provides a common principle for conservations of momentum, angular momentum, energy, and electric charge, and symmetries of space translation, rotational invariance, time translation, and color symmetry of quarks. Nadis gives no technical details, but Neuenschwander does, in a book for physics majors with a strong background in mathematics; the book does not shy away from Lie groups and the study of invariants. This new edition delves into distinctions between two Noether theorems and adds more exercises, references, and details.

Narayanan, Arvind, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder, *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction*, Princeton University Press, 2016; xxvii+304 pp, $49.50. ISBN 978-0-691-17169-2.

This book explains how Bitcoin works, its reliance on decentralization, how to mine Bitcoins (and how and why it gets more difficult over time), and how this "cryptocurrency" solves typical problems of a currency (e.g., theft, double spending). The authors suggest that the reader have a basic understanding of computer science and some programming experience. Online supplementary materials include homework questions and programming assignments.

Stewart, Ian, *Significant Figures: The Lives and Work of Great Mathematicians*, Basic Books, 2017; vi+300 pp, $27. ISBN 978-0-465-09612-1.

A previous generation of mathematicians was inspired by E.T. Bell's *Men of Mathematics* (1937), marred as it is by fanciful conjecture and inexact history. Stewart has prepared a readable collection of snapshots for the next generation. He relates the mathematics—the "most important—or interesting—discoveries and concepts" of 25 "significant figures" of mathematics. Stewart does not include any living mathematicians and ends with the relatively recent Benôit Mandelbrot and William Thurston, after treating Gödel and Turing.

Posamentier, Alfred S., and Stephen Krulik, *Strategy Games to Enhance Problem-Solving Ability in Mathematics*, World Scientific, 2017; xiii+121 pp, $24(P). ISBN 978-9-81-314634-1.

The authors present several families of games of strategy: tic-tac-toe games, blocking games, games with ongoing changes in strategy, and miscellaneous games, some of which were new to me. The book sets out the rules, the goal, and a "sample simulation" of the game. Suggestions for strategies are included for selected games. The authors assert a strong connection between good at problem solving and being good at games of strategy, and between the kinds of thinking involved in games and that practiced by mathematicians. They also point out similarities between heuristics for game strategies and those for general problem solving. But they do not cite sources that delve into those relationships, nor ones that support the claim of the title that playing strategy games enhances mathematical problem-solving ability. I wish that there were some mention of mathematical theory about games (e.g., every finite tree game has a natural outcome), plus a bibliography with further details about the origins and strategies for the games.

Carl B. Allendoerfer Awards

The Carl B. Allendoerfer Awards, established in 1976, are made to authors of articles of expository excellence published in *Mathematics Magazine*. The Awards are named for Carl B. Allendoerfer, a distinguished mathematician at the University of Washington and president of the Mathematical Association of America, 1959–60.

Brian Conrey, James Gabbard, Katie Grant, Andrew Liu, and Kent Morrison

"Intransitive Dice," *Mathematics Magazine*, Volume 89, Number 2, April 2016, pages 133–143.

Intuition suggests that transitivity should hold in matters of strength. This intuition fails spectacularly in an example described by Martin Gardner in 1970 and originally due to Bradley Efron a few years earlier. The example consists of four six-sided dice labeled A–D, with all faces having numbers belonging to {1, 2, 3, 4, 5, 6}, such that A beats B with probability 2/3, B beats C with probability 2/3, C beats D with probability 2/3, and D beats A with probability 2/3.

The fundamental question asked in "Intransitive Dice" is: How rare is this? In other words, given a random set of dice, how likely is it that one could put them into a cycle that is intransitive? The question is a tantalizing one, and the authors deftly move from the concrete to the abstract in their search for the answer.

The authors first look for intransitive triples of dice. The frequency of ties in the setting of six-sided dice may leave one unsatisfied, so the authors go on to consider triples of n-sided dice, allowing numbers in $\{1, \ldots, n\}$ and imposing the condition that the sum of the numbers on each die be $n(n+1)/2$. The authors conjecture that as n grows, the probability of a tie goes to 0, while the probability of an intransitive triple goes to 1/4. After giving some computational evidence for these conjectures, they prove that as the number of sides grows, the probability of an intransitive cycle when there are no ties is 1/4.

The authors conclude by returning to the four-dice setting of the original example and taking n-sided dice where the sum of all numbers on each die is $n(n+1)/2$. They present both heuristic and computational evidence that as n grows the probability of an intransitive cycle approaches 3/8. Even more provocatively, they conjecture that for k such dice, the limiting value approaches an expression in k that in turn goes to 1 as k grows. As they write, "... our intuition that intransitive dice are rare and that larger sets are even rarer is completely unfounded. They are common for three dice and almost unavoidable as the number of dice grows."

It's not exactly common for high school students to participate in a Math Circle that leads to a published article in a mathematics journal, but in the case of "Intransitive Dice" we have just that. The authors of this article hit on all of the important modes of mathematical research. They collect data, generalize patterns, look for conjectures, and even prove theorems. All of this is tied together in a fun article that touches on many areas of mathematics and keeps the reader engaged to the end.

Math. Mag. **90** (2017) 308–310. doi:10.4169/math.mag.90.4.308.

Response from Conrey, Gabbard, Grant, Liu, and Morrison

It is both an honor and a pleasure to be recognized with the Allendoerfer Award this year. For all five of us, this work has been an unusual experience. James, Katie, and Andrew were high school students in Morgan Hill, California, when they met Brian through the outreach program of the American Institute of Mathematics. They were looking for advice and mentoring for a science fair project in mathematics, and Brian suggested a topic that he liked to use in his "Untuition" talk aimed at high school students and teachers. For their science fair project, which eventually received an Honorable Mention for first place in the 2013 California State Science Fair, they focused on systematic construction of sets of intransitive dice. From this work, a precise question was formulated. Suppose that you have three random dice A, B, and C, with A stronger than B and B stronger than C. How likely is it that A is stronger than C? They decided to concentrate on "one-step" dice, which are as close to the standard die as possible. After they figured out completely how two one-step dice do not tie, then it was possible to write a computer program to find the patterns for triples of one-step dice that do not have any ties, and this enabled us to count the intransitive triples with enough accuracy to prove that as n goes to infinity, the answer to the question (for one-step dice) is that the two possibilities are equally likely. Further computer experiments convinced us that the intransitivity phenomenon is generic for larger sets of proper dice (i.e., those whose n faces use numbers between 1 and n with total $n(n+1)/2$). We have not found a proof or even a promising way to attack this grand conjecture, but we hope that our article stimulates interest in the conjecture and ultimately leads to its resolution.

Biographical Notes

Brian Conrey is the executive director of the American Institute of Mathematics in San Jose, a position he has held since 1997. He has taught at the University of Illinois and Oklahoma State University. He received his B.S. from Santa Clara University and his Ph.D. from the University of Michigan. His research interests are in analytic number theory and random matrix theory.

Kent Morrison received his B.A. and Ph.D. degrees from the University of California, Santa Cruz. After 30 years on the faculty of Cal Poly, San Luis Obispo, including nine years as department chair, he is now Professor Emeritus. Since 2009 he has been affiliated with the American Institute of Mathematics (AIM), where he helps in various capacities and directs the AIM Open Textbook Initiative to encourage the development and use of open source and open access textbooks for undergraduate mathematics courses. He has found it difficult to stick to one area of mathematics but in recent years his interests have been in combinatorics, probability, and game theory, and especially in mixtures of these areas. "Intransitive Dice" is his fifth article in *Mathematics Magazine* and his third to have student coauthors.

James Gabbard is an undergraduate at the University of Southern California where he is majoring in mechanical engineering and applied mathematics. In his spare time he enjoys backpacking and performing with the USC marching band.

Katie Grant is pursuing a B.S. in management science and a minor in computer science at the University of California, San Diego. Her favorite classes have been Decisions under Uncertainty, Econometrics, and Introduction to Probability. In addition to her studies she has worked in an internship with TD Ameritrade and is a registered investment advisor.

Shang-Chi Andrew Liu attends the University of California, Los Angeles, where he is majoring in political science with a double minor in philosophy and cognitive science. In his spare time he enjoys playing tennis and exploring the city of LA with his friends.

Vladimir Pozdnyakov and Michael Steele

"Buses, Bullies, and Bijections," *Mathematics Magazine*, Volume 89, Number 3, June 2016, pages 167–176.

"In times past in a country far away, passengers on intercity buses were all assigned seat numbers, and the buses were always full. There were many casual people, and there were a few people who were real sticklers for the rules. This is where our problem begins."

Suppose n people are assigned to n seats on a bus such that person i is assigned to seat i, for $1 \leqslant i \leqslant n$. Persons 2 through n enter the bus and take seats randomly. When person A enters, A sits in his or her assigned seat if it is available, otherwise A forces the person in his or her seat to move, and A requires the displaced person to take his or her assigned seat, possibly forcing someone else to move. This process continues until there are no more displaced persons. The authors ask: What is the probability that person 2 will have to move? This simple question starts the reader on a journey through the world of bijections on the symmetric group S_n.

The answer to the question comes from observing there is a bijection between permutations that have 1 and 2 in the same cycle and those that do not. The authors elegantly tackle more involved questions about bumping passengers and the associated bijections, permutations, and cycles. One interesting question concerns the probability that exactly m people out of the first k to enter the bus get displaced. The surprising result is that the answer is $\frac{1}{k}$ which does not depend on m or n.

The authors conclude with a glimpse at how tools like those developed in their paper can address other questions. Topics touched on include derangements, Spitzer's identity from combinatorial probability, and finally the famous Robinson–Knuth–Schensted correspondence, which is a bijection from permutations to pairs of Young tableaux. The authors demonstrate that there is much to be learned by examining certain bijections on the set of permutations.

The entire paper is intriguing and accessible. The writing is engaging, the explanations lucid, and the results remarkable in their simplicity. The authors clearly took great delight in sharing this story with the reader. There is almost a literary wink, where the authors say if you thought that was neat, wait until you see what's next.

Response from Pozdnyakov and Steele

What a surprise and a delight it is to hear that we have been awarded the Allendoerfer prize! When we began this project, we knew we had a story that we enjoyed telling to friends and students, but we had no idea that it might be acknowledged in this special way. We are both long-time readers of *Mathematics Magazine*, and, even from high school days, we have had a love for expositions of interesting, accessible mathematics. It is marvelous to hear that we have made a real contribution to this generations-long conversation where we have been eager listeners for so many years. We are genuinely moved. Thank you!

Biographical Notes

Vladimir Pozdnyakov received his Ph.D. in statistics in 2001 from the University of Pennsylvania. Since that time, he has taught at the University of Connecticut, where he is currently professor of statistics. His research is mostly in applied probability, and he has a particular interest in the discovery and exploitation of martingale tricks.

J. Michael Steele received his Ph.D. in mathematics from Stanford University in 1975. He has taught at U.B.C., Stanford, CMU, Princeton, and the Wharton School of the University of Pennsylvania. He has worked in many parts of probability theory, and he is the author of several books, including *The Cauchy Schwarz Master Class* published by the MAA.

46^{th} United States of America Mathematical Olympiad and 8^{th} United States of America Junior Mathematical Olympiad

GABRIEL CARROLL
Stanford University
Stanford, CA 94305
gdc@stanford.edu

DOUG ENSLEY
Mathematical Association of America
Washington DC 20036
densley@maa.org

The United States of America Mathematical Olympiad (USAMO) and Junior Olympiad (USAJMO) are high-level contests in the style of the International Mathematical Olympiad (IMO) offered by the Committee on the American Mathematics Competitions of the Mathematical Association of America. The two competitions are administered simultaneously, this year on April 19 and 20. Each competition uses the IMO format, consisting of three problems for each of two days, with an allowed time of 4.5 hours each day.

The USAMO is used to select a team of six students to represent the nation in the IMO, and the level of the problems reflects the level expected on the IMO competition. The USAJMO, offered to students in grade 10 and below, is used to identify students to train for participation in future IMO competitions. In setting problems for the USAJMO, the Committee strives to provide a nicely balanced link between the computational character of the AIME problems and the more advanced proof-oriented problems of the USAMO.

The 2017 contests included two common problems. On the first day, problem USAJMO1 was the same as USAMO1, and on the second day, USAJMO6 and USAMO4 were identical. This year, 285 students sat for the USAMO contest and 222 for the USAJMO. More information is available on the AMC section of the MAA website.

USAMO Problems

1. Prove that there are infinitely many distinct pairs (a, b) of relatively prime integers $a > 1$ and $b > 1$ such that $a^b + b^a$ is divisible by $a + b$.
2. Let $m_1, \ldots, m_n$ be a collection of n positive integers, not necessarily distinct. For any sequence of integers $A = (a_1, \ldots, a_n)$ and any permutation $w = w_1, \ldots, w_n$ of $m_1, \ldots, m_n$, define an *A-inversion* of w to be a pair of entries w_i, w_j with $i < j$ for which one of the following conditions holds:
   - $a_i \geq w_i > w_j$,
   - $w_j > a_i \geq w_i$, or
   - $w_i > w_j > a_i$.

Math. Mag. **90** (2017) 311–319. doi:10.4169/math.mag.90.4.311.

Show that, for any two sequences of integers $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$, and for any positive integer k, the number of permutations of $m_1, \ldots, m_n$ having exactly k A-inversions is equal to the number of permutations of $m_1, \ldots, m_n$ having exactly k B-inversions.

3. Let ABC be a scalene triangle with circumcircle Ω and incenter I. Ray AI meets $\overline{BC}$ at D and meets Ω again at M; the circle with diameter $\overline{DM}$ cuts Ω again at K. Lines MK and BC meet at S, and N is the midpoint of $\overline{IS}$. The circumcircles of $\triangle KID$ and $\triangle MAN$ intersect at points L_1 and L_2. Prove that Ω passes through the midpoint of either $\overline{IL_1}$ or $\overline{IL_2}$.

4. Let $P_1, \ldots, P_{2n}$ be $2n$ distinct points on the unit circle $x^2 + y^2 = 1$ other than $(1, 0)$. Each point is colored either red or blue, with exactly n of them red and n of them blue. Let $R_1, \ldots, R_n$ be any ordering of the red points. Let B_1 be the nearest blue point to R_1 traveling counterclockwise around the circle starting from R_1. Then let B_2 be the nearest of the remaining blue points to R_2 traveling counterclockwise around the circle from R_2, and so on, until we have labeled all of the blue points $B_1, \ldots, B_n$. Show that the number of counterclockwise arcs of the form $R_i \to B_i$ that contain the point $(1, 0)$ is independent of the way we chose the ordering $R_1, \ldots, R_n$ of the red points.

5. Let $\mathbf{Z}$ denote the set of all integers. Find all real numbers $c > 0$ such that there exists a labeling of the lattice points $(x, y) \in \mathbf{Z}^2$ with positive integers for which

   - only finitely many distinct labels occur, and
   - for each label i, the distance between any two points labeled i is at least c^i.

6. Find the minimum possible value of

$$\frac{a}{b^3+4} + \frac{b}{c^3+4} + \frac{c}{d^3+4} + \frac{d}{a^3+4},$$

given that a, b, c, d are nonnegative real numbers such that $a + b + c + d = 4$.

Solutions

1. Let n be an odd positive integer, and take $a = 2n - 1, b = 2n + 1$. Then $a^b + b^a \equiv 1 + 3 \equiv 0 \pmod 4$, and $a^b + b^a \equiv -1 + 1 \equiv 0 \pmod n$. Therefore, $a + b = 4n$ divides $a^b + b^a$.

   This problem and solution were proposed by Gregory Galperin.

2. It suffices to show the result for $B = (0, 0, \ldots, 0)$ since then any sequence is equivalent to any other sequence via B. We first show that the result holds for all sequences of the form $A = (a, a, \ldots, a)$ for some a.

   For each positive integer i, define the ith **lifting map** B_i on the permutations of $m_1, \ldots, m_n$ by $B_i(w_1, \ldots, w_n) = v_1, \ldots, v_n$ where $v_j = i$ if and only if $w_{n+1-j} = i$ and where the subsequence of v consisting of all entries not equal to i (taken in order) is equal to the subsequence of w consisting of all entries not equal to i.

   Lemma. Let $A_{i-1} = (i - 1, i - 1, \ldots, i - 1)$ and $A_i = (i, i, \ldots, i)$. Then the number of A_{i-1}-inversions of w equals the number of A_i-inversions of $B_i(w)$. Moreover, B_i is a bijection on the permutations of m, showing the result in this case.

   Proof of Lemma. It is easy to see that B_i is a bijection for any i since we can reverse the map.

   Now, note that any A_{i-1}-inversions between entries not equal to i in w are still A_i-inversions in $B_i(w)$ and vice versa. Notice also that there are no A_{i-1}-inversions

in w having i as the left entry. Similarly, there are no A_i-inversions having i as the *right* entry in $B_i(w)$.

On the other hand, in w, any non-i entry to the *left* of an i forms an A_{i-1}-inversion with that i. And in $B_i(w)$, any non-i entry to the *right* of an i forms an A_i-inversion with that i. Since the positions of the i's are reversed from w to $B_i(w)$, the number of inversions involving an i are equal in each case, and the result follows. □

For $j > i$, we denote $B_{i\to j} := B_j \circ B_{j-1} \circ \cdots \circ B_{i+2} \circ B_{i+1}$ and denote $B_{j\to i} := B_{i\to j}^{-1}$. Also, let $B_{i\to i}$ be the identity permutation.

Now, for $A = (a_1, \ldots, a_n)$ and for a permutation w of $m_1, \ldots, m_n$, we define $\phi_A(w)$ as follows. Let $w^{(1)} = B_{0\to a_1}(w)$ and, inductively, for $i > 1$ let $w^{(i)}$ be the result of applying $B_{a_{i-1}\to a_i}$ to the last $n - i + 1$ terms of $w^{(i-1)}$ and leaving the first $i - 1$ terms unchanged. Finally, let $\phi_A(w) = w^{(n)}$.

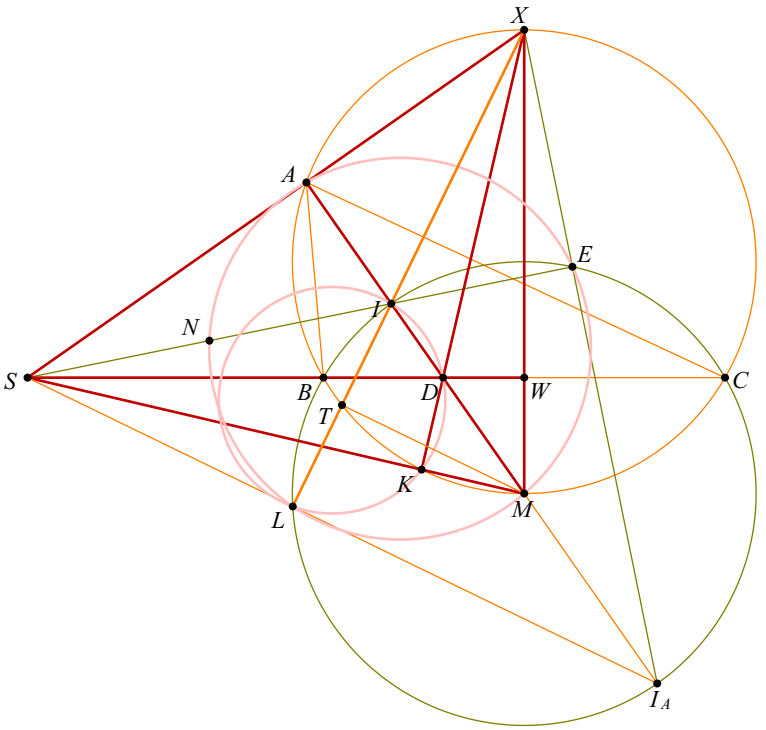
Lemma. The number of A-inversions of $\phi_A(w)$ is equal to the number of B-inversions of w where $B = (0, 0, \ldots, 0)$.

Proof of Lemma. This is a consequence of the definition of ϕ_A: At any step $w^{(i)}$ in the process of computing $\phi_A(w)$, we consider the sequence $A^{(i)}$ formed by changing the last $n - i + 1$ terms of the previous sequence $A^{(i-1)}$ (starting at $A^{(0)} = (0, 0, \ldots, 0)$) from a_{i-1} to a_i. Then we have $A^{(n)} = A$, and at each step, the number of $A^{(i)}$-inversions of $w^{(i)}$ is equal to the number of $A^{(i-1)}$-inversions of $w^{(i-1)}$, via an application of the previous lemma. The result follows. □

Now, ϕ_A is a bijection, being a composition of bijections, and we are done.

This problem and solution were proposed by Maria Monks Gillespie.

3. Let W be the midpoint of $\overline{BC}$, and let X be the point on Ω opposite M. Observe that line KD passes through X, and thus lines BC, MK, XA concur at the orthocenter of $\triangle DMX$, which is S. Denote by I_A the A-excenter of ABC.



Next, let E be the foot of the altitude from I to $\overline{XI_A}$; observe that E lies on the circle ω centered at M through I, B, C, I_A. Then, S is the radical center of ω, Ω,

and the circle with diameter $\overline{IX}$; hence, line SI passes through E; accordingly, I is the orthocenter of $\triangle XSI_A$; denote by L the foot of the altitude from X to $\overline{I_A S}$.

We claim that this L lies on both the circumcircle of $\triangle KID$ and $\triangle MAN$. It lies on the circumcircle of $\triangle MAN$ since this circle is the nine-point circle of $\triangle XSI_A$. For the other, note that $\triangle MWI \sim \triangle MIX$ since they share the same angle at M and $MW \cdot MX = MB^2 = MI^2$. Consequently, $\angle IWM = \angle MIX = 180^\circ - \angle LIM = 180^\circ - \angle MLI$, enough to imply that quadrilateral $MWIL$ is cyclic. But lines IL, DK, and WM meet at X, so power of a point in cyclic quadrilaterals $DKMW$ and $MWIL$ gives $XD \cdot XK = XM \cdot XW = XI \cdot XL$, hence $KDIL$ is cyclic as needed.

All that remains to show is that the midpoint T of $\overline{IL}$ lies on Ω. But this follows from the fact that $\overline{TM} \parallel \overline{I_A L} \implies \angle MTX = 90^\circ$, thus the problem is solved.

This problem and solution were proposed by Evan Chen.

4. Assume the points have been labeled $P_1, P_2, \ldots, P_{2n}$ in order, going counterclockwise from $(1, 0)$. Now, write out the color of each point in order, and replace each R with a $+1$ and each B with a -1 to get a list $p_1, \ldots, p_{2n}$ of $+1$'s and -1's. Consider the partial sums $s_k = p_1 + \cdots + p_k$ of this sequence, and choose the index k with s_k minimal (breaking ties by smallest k). Rotate the circle clockwise so that points $P_1, \ldots, P_k$ are moved past $(1, 0)$; the resulting sequence of $+1$'s and -1's from the new orientation now has all nonnegative partial sums, and the total sum is 0.

   In the rotated diagram, arc $R_1 \to B_1$ does not cross $(1, 0)$, for otherwise the sequence ends with a string of $+1$'s, and the partial sums before those $+1$'s would be negative. Furthermore, the sequence of entries from R_1 to B_1 looks like $+1, +1, +1, \ldots, +1, -1$, and so removing R_1 and B_1 is equivalent to removing a consecutive pair of a $+1$ and -1, so the partial sums remain all nonnegative. It follows that the next pairing also doesn't cross $(1, 0)$, and so on. Thus, no counterclockwise arcs $R_i \to B_i$ cross $(1, 0)$.

   Finally, note that all the red points among $P_1, \ldots, P_k$ are paired with blue points in this same subsequence since there are no crossings in the rotated picture. Let m be the difference between the number of blue and red points among $P_1, \ldots, P_k$. Then, exactly m blue points in $P_1, \ldots, P_k$ were matched with red points from $P_{k+1}, \ldots, P_{2n}$. Therefore, when we rotate the circle back to its original position, there are exactly m crossings, no matter how the red points were ordered. Since k and m are independent of the ordering, the proof is complete.

   This problem and solution were proposed by Maria Monks Gillespie.

5. The answer is that such a labeling is possible for any $c < \sqrt{2}$.

   Let any $c < \sqrt{2}$ be given. We can partition $\mathbf{Z}^2$ into two subsets

$$L_1 = \{(x, y) \mid x + y \text{ is odd}\} \quad \text{and} \quad L_1' = \{(x, y) \mid x + y \text{ is even}\}.$$

   Both L_1 and L_1' are square lattices with unit length $\sqrt{2}$. Hence, we can similarly partition L_1' into two square lattices L_2 and L_2' with unit length $\sqrt{2}^2$, then partition L_2' into two square lattices L_3 and L_3' with unit length $\sqrt{2}^3$, and so forth. Hence, for any $N \geq 1$, $\mathbf{Z}^2$ can be partitioned into $N + 1$ square lattices $L_1, L_2, \ldots, L_N, L_N'$ with unit lengths $\sqrt{2}, \sqrt{2}^2, \ldots, \sqrt{2}^N, \sqrt{2}^N$, respectively. Take N large enough so that $c^{N+1} \leq \sqrt{2}^N$. For $i = 1, \ldots, N$, label all points in L_i by i, and then label all points in L_N' by $N + 1$. Any two points in L_i lie at least $\sqrt{2}^i > c^i$ apart, while any two points in L_N' lie at least $\sqrt{2}^N \geq c^{N+1}$ apart, so this is a valid labeling.

Now suppose that $c \geq \sqrt{2}$. For a nonnegative integer m, define

$$R_m = \{(x, y) \mid 1 \leq x \leq 2^a,\ 1 \leq y \leq 2^b\} \subseteq \mathbf{Z}^2,$$

where

$$(a, b) = \begin{cases} (\frac{m}{2}, \frac{m}{2}) & \text{if } m \text{ is even,} \\ (\frac{m-1}{2}, \frac{m+1}{2}) & \text{if } m \text{ is odd.} \end{cases}$$

We will show by induction that R_m does not have a valid labeling using only labels at most m, which will prove that $\mathbf{Z}^2$ has no valid labeling. The case $m = 0$ is trivial.

Suppose $m > 0$ is odd and that R_{m-1} does not have a valid labeling using only $1, \ldots, m-1$ (the inductive hypothesis) but that R_m does have a valid labeling using only $1, \ldots, m$. Consider this labeling of R_m. Since $R_m \supseteq R_{m-1}$, some point (x_0, y_0) with $y_0 \leq 2^{(m-1)/2}$ must be labeled m. But then (x_0, y_0) lies directly below a translate R' of R_{m-1} inside R_m. The distance between (x_0, y_0) and any point in R' is at most

$$\sqrt{(2^{\frac{m-1}{2}} - 1)^2 + (2^{\frac{m-1}{2}})^2} < \sqrt{2}^m \leq c^m,$$

so no points in R' can be labeled m. But by the inductive hypothesis, R' has no valid labeling using only $1, \ldots, m-1$, which is a contradiction.

Now suppose $m > 0$ is even and that R_{m-1} does not have a valid labeling using only $1, \ldots, m-1$, but R_m does have a valid labeling using only $1, \ldots, m$. By the inductive hypothesis, some point (x_0, y_0) with $\frac{1}{4} \cdot 2^{m/2} < y_0 \leq \frac{3}{4} \cdot 2^{m/2}$ must be labeled m (since the corresponding rows of R_m form a rotated copy of R_{m-1}). But then (x_0, y_0) lies either directly to the left or to the right of a translate R' of R_{m-1} inside R_m. The distance between (x_0, y_0) and any point of R' is less than

$$\sqrt{(\tfrac{3}{4} \cdot 2^{\frac{m}{2}})^2 + (2^{\frac{m-2}{2}})^2} = \tfrac{\sqrt{13}}{4} \cdot \sqrt{2}^m < \sqrt{2}^m \leq c^m,$$

so no points in R' can be labeled m. But by the inductive hypothesis, R' has no valid labeling using only $1, \ldots, m-1$, which is a contradiction. This completes the proof.

This problem and solution were proposed by Ricky Liu.

6. We will show that the minimum is $\frac{2}{3}$. We have

$$\frac{4a}{b^3+4} = a - \frac{ab^3}{b^3+4} \geq a - \frac{ab}{3}$$

since AM-GM implies $b^3 + 4 = (b^3/2) + (b^3/2) + 4 \geq 3b^2$. Then

$$\begin{aligned} \frac{a}{b^3+4} + \frac{b}{c^3+4} + \frac{c}{d^3+4} + \frac{d}{a^3+4} &\geq \frac{a+b+c+d}{4} - \frac{ab+bc+cd+da}{12} \\ &= \frac{a+b+c+d}{4} - \frac{(a+c)(b+d)}{12} \\ &\geq \frac{a+b+c+d}{4} - \frac{(a+b+c+d)^2/4}{12} \\ &= \frac{4}{4} - \frac{4}{12} = \frac{2}{3}. \end{aligned}$$

The minimum is realized when, for example, $a = b = 2$ and $c = d = 0$.

This problem and solution were proposed by Titu Andreescu.

USAJMO Problems

1. Same as USAMO 1.
2. Consider the equation

$$(3x^3 + xy^2)(x^2y + 3y^3) = (x - y)^7.$$

   (a) Prove that there are infinitely many pairs (x, y) of positive integers satisfying the equation.
   (b) Describe all pairs (x, y) of positive integers satisfying the equation.

3. Let ABC be an equilateral triangle and let P be a point on its circumcircle. Let lines PA and BC intersect at D; let lines PB and CA intersect at E; and let lines PC and AB intersect at F. Prove that the area of triangle DEF is twice the area of triangle ABC.
4. Are there any triples (a, b, c) of positive integers such that $(a - 2)(b - 2)(c - 2) + 12$ is a prime that properly divides the positive number $a^2 + b^2 + c^2 + abc - 2017$?
5. Let O and H be the circumcenter and the orthocenter of an acute triangle ABC. Points M and D lie on side BC such that $BM = CM$ and $\angle BAD = \angle CAD$. Ray MO intersects the circumcircle of triangle BHC in point N. Prove that $\angle ADO = \angle HAN$.
6. Same as USAMO 4.

Solutions

2. The following analysis solves both parts at once. Write the equation as

$$x(3x^2 + y^2)y(x^2 + 3y^2) = (x - y)^7,$$

which is equivalent to

$$(x^3 + 3xy^2)(3x^2y + y^3) = (x - y)^7.$$

Let $x^3 + 3xy^2 = a$ and $3x^2y + y^3 = b$. Then we have $(ab)^3 = (x - y)^{21} = (a - b)^7$. Let $d = \gcd(a, b)$. Then $a = du$ and $b = dv$ for some relatively prime positive integers u and v. Hence,

$$(uv)^3 = d(u - v)^7.$$

But $\gcd(u - v, uv) = 1$. It follows that $u - v = 1$ and $d = (uv)^3$. Hence, $u = k + 1$ and $v = k$, where k is a positive integer, and so $a = (k + 1)^4k^3$ and $b = k^4(k + 1)^3$. Then

$$(x - y)^3 = a - b = [k(k + 1)]^3$$

and

$$(x + y)^3 = a + b = [k(k + 1)]^3(2k + 1).$$

It follows that $2k + 1 = n^3$ for some odd integer $n > 1$ and that $x + y = nk(k + 1)$ and $x - y = k(k + 1)$. Hence,

$$(x, y) = \left(\frac{(n + 1)k(k + 1)}{2}, \frac{(n - 1)k(k + 1)}{2}\right)$$

where $k = \frac{n^3-1}{2}$. Thus,

$$(x, y) = \left(\frac{(n+1)(n^6-1)}{8}, \frac{(n-1)(n^6-1)}{8} \right)$$

where n is an odd integer greater than 1, and it is easy to check that these are solutions to the given equation. Hence, these pairs describe all the solutions.

This problem and solution were proposed by Titu Andreescu.

3. We assume the points are configured so that P is on minor arc $\widehat{BC}$ of the circle. Scale to set $AB = 1$. Then $[ABC] = \sqrt{3}/4$. Set $b = PB$, $c = PC$, $e = PE$, and $f = PF$. Note that $\angle FBD = \angle ECD = \angle BPC = 120^\circ$. Hence,

$$[DEF] = [BCEF] - [FBD] - [ECD] = \frac{1}{2} \sin 120^\circ (BE \cdot CF - BF \cdot BD - CE \cdot CD).$$

It suffices to show that $[DEF] = \sqrt{3}/2$ or

$$2 = (BE \cdot CF - BF \cdot BD - CE \cdot CD) = (b+e)(c+f) - BF \cdot BD - CE \cdot CD.$$

Because $\angle FBC = \angle BPC$ and $\angle FCB = \angle PCB$, triangles FCB and BCP are similar, implying that

$$\frac{FC}{BC} = \frac{CB}{CP} = \frac{BF}{PB} \quad \text{or} \quad \frac{c+f}{1} = \frac{1}{c} = \frac{BF}{b}.$$

Thus, $c + f = 1/c$ and $BF = b/c$. Analogously, $b + e = 1/b$ and $CE = c/b$. It remains to show that

$$2 = (b+e)(c+f) - BF \cdot BD - CE \cdot CD = \frac{1}{bc} - \frac{b}{c} \cdot BD - \frac{c}{b} \cdot CD.$$

Note that $\angle BPD = \angle CPD = 60^\circ$, so we have $BD/CD = BP/CP$ by the angle-bisector theorem. Consequently, we have $BD = b/(b+c)$ and $CD = c/(b+c)$. Thus, we want to show that

$$\begin{aligned} 2 &= \frac{1}{bc} - \frac{b}{c} \cdot BD - \frac{c}{b} \cdot CD = \frac{1}{bc} - \frac{b^2}{c(b+c)} - \frac{c^2}{b(b+c)} \\ &= \frac{1}{bc} - \frac{b^3+c^3}{bc(b+c)} = \frac{1-b^2-c^2+bc}{bc}, \end{aligned}$$

or $b^2 + c^2 + bc = 1$, which is true by applying the law of cosines in $\triangle BPC$.

This problem was proposed by Titu Andreescu, Luis Gonzalez, and Cosmin Pohoata. Solution by USA(J)MO packet reviewers.

4. Suppose (a, b, c) is such a triple. WLOG assume $a \le b \le c$. Put $p = (a-2)(b-2)(c-2) + 12$ and $s = a^2 + b^2 + c^2 + abc - 2017 > 0$. The prime p also divides the difference

$$\begin{aligned} &a^2 + b^2 + c^2 + abc - 2017 - (a-2)(b-2)(c-2) - 12 \\ &= (a+b+c)^2 - 4(a+b+c) + 4 - 2025 \\ &= (a+b+c-2)^2 - 45^2 \\ &= (a+b+c-47)(a+b+c+43). \end{aligned}$$

We first rule out $a = 1$: If $b = 1$ also, then the prime $p = c + 10$ divides $(c-45)(c+45)$ so divides either 35 or 55, hence $p = 11$ and $c = 1$. If $b \ge 3$,

then $p > 0$ only if $c \leq 14$. But in all such cases, $s < 0$. And $b = 2$ implies $p = 12$, which is not prime. Thus, $a = 1$ is impossible. Also, $a = 2$ is impossible, again because 12 is not prime.

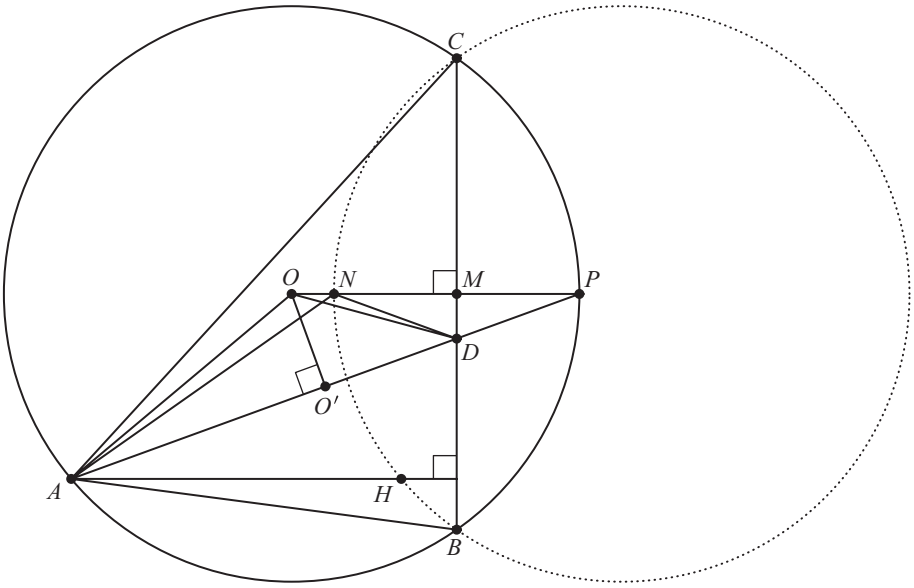
Now, let $x = a - 2$, $y = b - 2$, $z = c - 2$. We now know that $1 \leq x \leq y \leq z$ and $(x + 2) + (y + 2) + (z + 2) > 47$. So $x + y + z \geq 41$, and therefore $z \geq 14$. The prime $xyz + 12$ cannot divide $(x + 2) + (y + 2) + (z + 2) - 47$ since $xyz > x + y + z - 41$. Indeed, this latter inequality follows from $x(yz - 1) \geq yz - 1 \geq y + z - 2$ (the latter holds by $(y - 1)(z - 1) \geq 0$).

Hence, $xyz + 12$ divides $(x + 2) + (y + 2) + (z + 2) + 43$. They cannot be equal: x, y, z must all be odd, otherwise $xyz + 12$ is not prime, but then $(x + 2) + (y + 2) + (z + 2) + 43$ is even and so not equal to $xyz + 12$. Thus, $2(xyz + 12) \leq x + y + z + 49$, implying $2yz - 1 \leq x(2yz - 1) \leq y + z + 25$. It follows that $(2y - 1)(2z - 1) \leq 53$. If $y = 1$, then $z \leq 27$; otherwise, $y \geq 2$ and $2z - 1 \leq 53/3 < 18$, implying $x, y, z \leq 9$. In both cases, the requirement $x + y + z \geq 41$ is violated.

Thus, the desired triples (a, b, c) do not exist.

This problem and solution were proposed by Titu Andreescu.

5. The key idea is to prove that $ADNO$ is cyclic. Once this is proven, the problem follows by noticing that $\angle ADO = \angle ANO = \angle HAN$, where the latter holds due to the fact that $ON \| AH$.



To prove the concyclicity one can use power of a point. First, extend segment AD to meet the circumcircle (denoted by Ω) of $\triangle ABC$ at P. Then P is the midpoint of minor arc $\overset{\frown}{BC}$ of Ω, and P, M, N, O all lie on the perpendicular bisector of $\overline{BC}$. Next, notice that the reflection of H across line BC lies on Ω. This implies that the circumcircle of $\triangle BHC$ is the reflection of Ω across line BC, so N is simply the reflection of P. Hence, M is the midpoint of $\overline{PN}$. Next, let O' denote the orthogonal projection of O on AD. Clearly, $OO'DM$ is cyclic, so power of a point yields $PM \cdot PO = PD \cdot PO'$. Then

$$PN \cdot PO = 2 \cdot PM \cdot PO = 2 \cdot PO' \cdot PD = PA \cdot PD,$$

which by power of a point gives the concyclicity of $ADNO$. This completes the proof.

This problem was proposed by Ivan Borsenco. Solution by Titu Andreescu and Cosmin Pohoata.

The top 12 students (in alphabetical order) on the 2017 USAMO were:

Zachary Chroman	The Nueva School	CA
Andrew Gu	Pittsburgh Allderdice High School	PA
James Lin	Phillips Exeter Academy	NH
Daniel Liu	Redmond High School	WA
Michael Ren	Phillips Academy	MA
Victor Rong	Marc Garneau Collegiate Institute	ON
Ashwin Sah	Jesuit High School	OR
Mihir Singhal	Palo Alto High School	CA
Alec Sun	Phillips Exeter Academy	NH
Kada Williams	Radnoti Miklos High School	Hungary
Yuan Yao	Phillips Exeter Academy	NH
William Zhao	Richmond Hill High School	ON

The top 11 students (in alphabetical order) on the 2017 USAJMO were:

Vincent Bian	Poolesville High School	MD
Xinyang Chen	Carson Middle School	PA
Kevin Cong	Ridge High School	NJ
Daniel Hu	Los Altos High School	CA
Benjamin Kang	Thomas Jefferson High School	VA
Le Nguyen	Wasatch Academy	UT
Brian Liu	Holmdel High School	NJ
Benjamin Qi	Princeton High School	NJ
Brandon Wang	Saratoga High School	CA
Andrew Yao	Weston High School	MA
Jonathan Zhou	Monta Vista High School	CA

Statement of Ownership, Management, and Circulation (All Periodicals Publications Except Requester Publications)

UNITED STATES POSTAL SERVICE®

1. Publication Title	2. Publication Number	3. Filing Date
Mathematics Magazine	0 0 2 5 – 5 7 0 X	August 23, 2017
4. Issue Frequency	5. Number of Issues Published Annually	6. Annual Subscription Price
Bimonthly except July and August	5	$296

7. Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4®)*
Mathematical Association of America
1529 Eighteenth St., NW
Washington, DC 20036

Contact Person: Beverly Joy Ruedi
Telephone *(Include area code)*: 540-582-2208 (home office)

8. Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*
Mathematical Association of America
1529 Eighteenth St., NW
Washington, DC 20036

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor *(Do not leave blank)*

Publisher *(Name and complete mailing address)*
Mathematical Association of America
1529 Eighteenth St., NW
Washington, DC 20036

Editor *(Name and complete mailing address)*
Michael Jones
1529 Eighteenth Street, NW
Washington, DC 20036

Managing Editor *(Name and complete mailing address)*
Beverly Ruedi
1529 Eighteenth St., NW
Washington, DC 20036

10. Owner *(Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)*

Full Name	Complete Mailing Address
Mathematical Association of America	1529 18th St., NW, Washington, DC 20036

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → ☐ None

Full Name	Complete Mailing Address

12. Tax Status *(For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)*
The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes:
☒ Has Not Changed During Preceding 12 Months
☐ Has Changed During Preceding 12 Months *(Publisher must submit explanation of change with this statement)*

13. Publication Title	14. Issue Date for Circulation Data Below
Mathematics Magazine	June 2017

15. Extent and Nature of Circulation			Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies *(Net press run)*			2440	2500
b. Paid Circulation *(By Mail and Outside the Mail)*	(1)	Mailed Outside-County Paid Subscriptions Stated on PS Form 3541 (Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)	1819	1925
	(2)	Mailed In-County Paid Subscriptions Stated on PS Form 3541 *(Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)*	0	0
	(3)	Paid Distribution Outside the Mails Including Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Paid Distribution Outside USPS®	0	0
	(4)	Paid Distribution by Other Classes of Mail Through the USPS (e.g., First-Class Mail®)	0	0
c. Total Paid Distribution *[Sum of 15b (1), (2), (3), and (4)]* →			1819	1925
d. Free or Nominal Rate Distribution *(By Mail and Outside the Mail)*	(1)	Free or Nominal Rate Outside-County Copies included on PS Form 3541	0	0
	(2)	Free or Nominal Rate In-County Copies Included on PS Form 3541	0	0
	(3)	Free or Nominal Rate Copies Mailed at Other Classes Through the USPS (e.g., First-Class Mail)	0	0
	(4)	Free or Nominal Rate Distribution Outside the Mail *(Carriers or other means)*	100	100
e. Total Free or Nominal Rate Distribution *(Sum of 15d (1), (2), (3) and (4))*			100	100
f. Total Distribution *(Sum of 15c and 15e)* →			1919	2025
g. Copies not Distributed *(See Instructions to Publishers #4 (page #3))* →			521	475
h. Total *(Sum of 15f and g)*			2440	2500
i. Percent Paid *(15c divided by 15f times 100)* →			94.7%	95%

* If you are claiming electronic copies, go to line 16 on page 3. If you are not claiming electronic copies, skip to line 17 on page 3.

Statement of Ownership, Management, and Circulation (All Periodicals Publications Except Requester Publications)

UNITED STATES POSTAL SERVICE®

16. Electronic Copy Circulation	Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Paid Electronic Copies →	15790	15840
b. Total Paid Print Copies (Line 15c) + Paid Electronic Copies (Line 16a) →	17609	17765
c. Total Print Distribution (Line 15f) + Paid Electronic Copies (Line 16a) →	17709	17865
d. Percent Paid (Both Print & Electronic Copies) (16b divided by 16c × 100) →	99%	99%

☒ **I certify that 50% of all my distributed copies (electronic and print) are paid above a nominal price.**

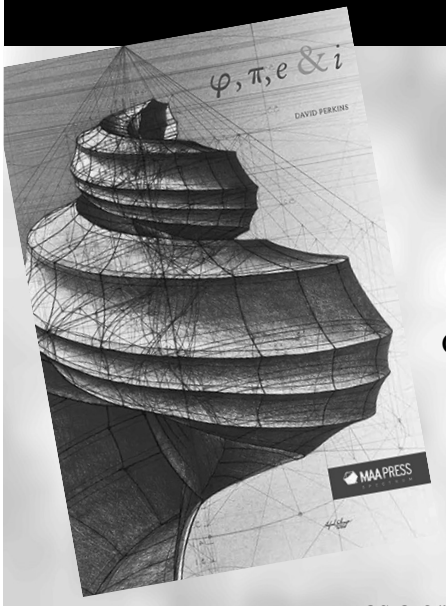
17. Publication of Statement of Ownership

☒ If the publication is a general publication, publication of this statement is required. Will be printed in the October 2017 issue of this publication. ☐ Publication not required.

18. Signature and Title of Editor, Publisher, Business Manager, or Owner
Beverly Joy Ruedi
Managing Editor
Date: 8/23/2017

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

New in the eBooks Store



φ, π, e, and i

David Perkins

Certain constants occupy precise balancing points in the cosmos of number, like habitable planets sprinkled throughout our galaxy at just the right distances from their suns. This book introduces and connects four of these constants (φ, π, e and i). But here we discuss their properties, as a group, at a level appropriate for an audience armed only with the tools of elementary calculus. Since the book is written with the goal that an undergraduate student can read the book solo, the author provides endnotes throughout, in case the reader is unable to work out some of the missing steps. Those endnotes appear in the last chapter, Extra Help. Each chapter concludes with a series of exercises, all of which introduce new historical figures or content.

Spectrum
e-ISBN: 978-1-61444-525-8

ebook: $25.00
191 pages, 2017

To order an ebook go to www.maa.org/ebooks/PPE.

Order a paperbound book at store.maa.org.

Catalog Code: PPE
ISBN: 978-0-88385-589-8

List: $50.00
MAA Member: $37.50

MAA PRESS

MAA
MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth St., NW
Washington, DC 20036

CONTENTS

ARTICLES

PROBLEMS AND SOLUTIONS

REVIEWS

NEWS AND LETTERS