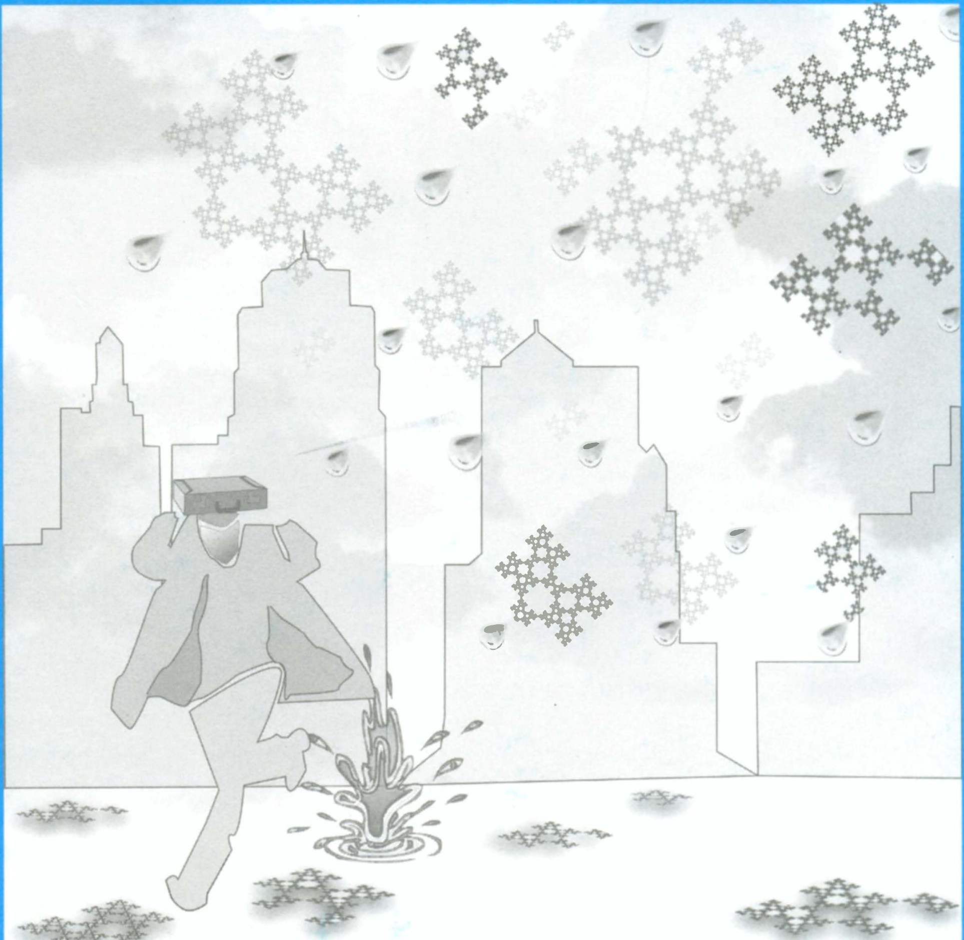
Vol. 82, No. 4, October 2009

MATHEMATICS MAGAZINE



Running in the Fractal Rain

- A Fractal Made of Golden Sets
- A Survey of Euler's Constant
- Keeping Dry: The Mathematics of Running in the Rain

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

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.

A more detailed statement of author guidelines appears in this *Magazine*, Vol. 74, pp. 75–76, and is available from the Editor or at www.maa.org/pubs/mathmag.html. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Please submit new manuscripts by email to Editor-Elect Walter Stromquist at mathmag@maa.org. A brief message with an attached PDF file is preferred. Word-processor and DVI files can also be considered. Alternatively, manuscripts may be mailed to Mathematics Magazine, 132 Bodine Rd., Berwyn, PA 19312-1027. If possible, please include an email address for further correspondence.

Cover image: Running in the Fractal Rain, by Hunter Cowdery, supervised by Jason Challas. Hunter studies illustration at San Jose State University, having completed his coursework at West Valley College, where Jason teaches art.

AUTHORS

Joseph B. Dence studied mathematics, physics, and chemistry at Bowling Green State University (B.A. 1963), and physical organic chemistry at the California Institute of Technology (Ph.D. 1968). Besides teaching chemistry (Florida State University) and mathematics and doing research at a variety of institutions, he also has earned an M.A. in mathematics from the University of Missouri, St. Louis, where he recently retired from the chemistry faculty. Ever oscillating between his two favorite disciplines, he pursues professional interests in number theory, physical chemistry, and chemical endocrinology. Leisure activities include golf, billiards, classical music (Mahler), and his wife Carmen's Colombian dance troupe.

Thomas P. Dence studied mathematics at Bowling Green State University (B.S.Ed. 1967), the University of Colorado (M.A. 1969), and Colorado State University (Ph.D. 1974). After a post-doctoral fellowship at New Mexico State University, he has been at Bowling Green, Southern Utah University, California State, Los Angeles, the University of Georgia, and now at Ashland University, where he had been department chairman for nine years. In 2003 he was recognized by the Ohio Section MAA as its Distinguished Professor of the year, and then served as President of the Section in 2006–07. Research interests include classical analysis, number theory, and using programming to explore mathematics. Nonmathematical interest include racquetball, gardening, photograph, hiking out west, and trying to maintain a respectable distance behind his three sons with his golf tee shots.

Marc Frantz majored in painting at the Herron School of Art, where he received his B.F.A. in 1975. He had a thirteen-year career as a painter and picture framer, mixed with self-taught courses in mathematics and physics. On the basis of exams, he entered graduate school in mathematics at Indiana University Purdue University Indianapolis (IUPUI), where he received his M.S. in 1990. Since then he has taught mathematics at IUPUI and Indiana University, where he is currently a research associate. He loves the visual approach to mathematics, especially links between mathematics and art.

Dan Kalman received his Ph.D. from the University of Wisconsin in 1980, and has been at American University since 1993. Prior to that he had academic appointments (University of Wisconsin, Green Bay; Augustana College; Sioux Falls) and worked for eight years in the aerospace industry in Southern California. Kalman is a past Associate Executive Director of the MAA, author of a book published by the MAA, and frequent contributor to MAA journals. He delights in puns and word play of all kinds, and is an avid fan of Douglas Adams, J.R.R. Tolkien, and Gilbert and Sullivan.

Bruce Torrence received his Ph.D. from the University of Virginia in 1991. He has been at Randolph-Macon College since 1993, where he is Garnett Professor of Mathematics and chair of the mathematics department. He is the coauthor (with his wife and colleague Eve Torrence) of an introductory text on the *Mathematica* software program. Torrence is currently co-editor of *Math Horizons*. He is an avid cyclist, and enjoys photography and the visual arts.

Vol. 82, No. 4, October 2009

MATHEMATICS MAGAZINE

MATHEMATICS MAGAZINE (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D.C. 20036 and Montpelier, VT, bimonthly except July/August.
The annual subscription price for *MATHEMATICS MAGAZINE* to an individual member of the Association is $131. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount; and new members receive a 20% dues discount for the first two years of membership.)

Subscription correspondence and notice of change of address should be sent to the Membership/ Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036. Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

Advertising correspondence should be addressed to

MAA Advertising
1529 Eighteenth St. NW
Washington DC 20036

Phone: (866) 821-1221
Fax: (202) 387-1208
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:

MAA Service Center, *maahq@maa.org*

All at the address:

The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036

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, N.W., Washington, D.C. 20036-1385.

Printed in the United States of America

ARTICLES

A Fractal Made of Golden Sets

MARC FRANTZ
Indiana University
Bloomington, IN 47405
mfrantz@indiana.edu

doi:10.4169/193009809X468670

The famous golden rectangle is a rectangle whose sides have the short-to-long ratio

$$\gamma = (\sqrt{5} - 1)/2 \approx 0.618,$$

which is called the golden ratio [**5**] or its reciprocal [**4**], depending on the author. Almost any discussion of the golden rectangle will include the image in FIGURE 1, which is constructed by starting with a golden rectangle, partitioning it into a square and a smaller golden rectangle, and then repeating the process indefinitely. We will refer to any such set as a *golden set*.

Authors occasionally include the golden set in discussions of self-similarity [**5, 6**], because it is the union of a square and a smaller copy of itself, scaled by a factor of γ. Some popularizations of the golden ratio [**4, 7**] also include examples of truly self-similar, planar fractals to illustrate properties of γ. Now the golden set is not truly self-similar, because it is not a union of smaller copies of itself, but it suggests a problem: Can we construct a self-similar fractal made of golden sets?

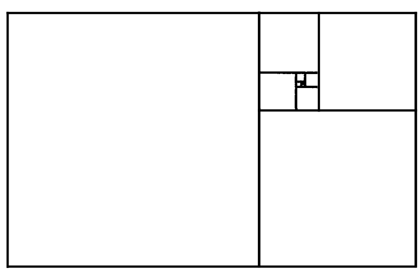


Figure 1 A golden set

In answer to this question, we construct a self-similar fractal that is the closure of a countable union of golden sets. For aesthetic reasons, we want our fractal to consist of *visible* golden sets; it shouldn't be such a jumble of them that their presence is obscured. We present some examples that have a nice symmetry and even give us a little more than we bargained for.

The ideas presented here make a good starting point for classroom exercises on fractals. In particular, verifying that a set is a subset of a given fractal often requires some creativity and lots of practice with similarity transformations (similitudes). We suggest some specific exercises in the last section.

As a preview of our results, the left side of FIGURE 2 shows a self-similar fractal set G with Hausdorff dimension $D \approx 1.44$. (The Hausdorff dimension of a plane set is usually thought of as a measure of its area-filling tendency, with 2 being the maximum. For more on Hausdorff dimension see the discussions of equations (2) and (5).) As we will show, G is symmetric with respect to the line $y = \gamma x$, as indicated. The right side of FIGURE 2 shows a golden set S (black) superimposed on G. It turns out, as the reader can guess by inspection of the figure, that S is a subset of G. In fact, a closer look will suggest the presence of many golden sets in G, at many different scales. We will show as promised that G can be constructed as the closure of a countable union of golden sets, all having their sides parallel or perpendicular to those of S. Obviously, then, G is also the closure of a countable union of golden rectangles.

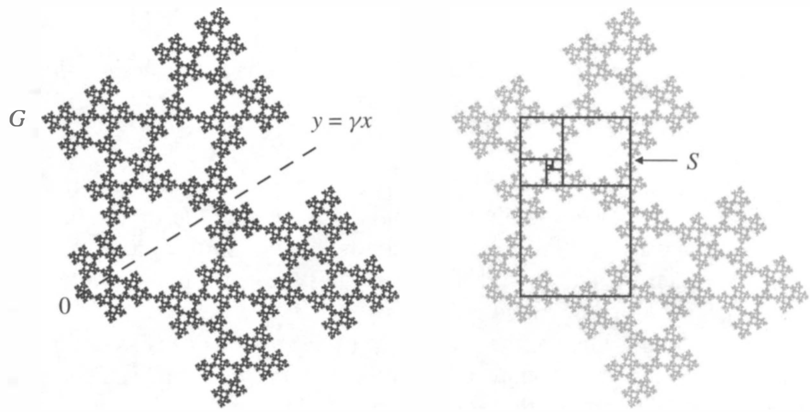


Figure 2 The set G and a golden set S superimposed on G

Because of the symmetry of G, this particular construction of it has an interesting emergent feature, suggested in FIGURE 3. In parts (a) through (e) we see successive unions of 1, 3, 9, 27, and 81 copies of S, all aligned with the coordinate axes. But gradually there appears the suggestion of an entirely different collection of golden subsets of G—the reflections of the original ones in the line $y = \gamma x$. One such reflected set is shown in part (f). Thus, by constructing golden sets aligned with the coordinate axes, we get for free a whole new collection of golden sets running in a different direction. The pleasing thing is, we can see both groups of sets even though they overlap each other. And what about the convex hulls of each of those H-shaped clusters? They are golden rectangles, of course. In the last section we introduce a double version of G called H, whose convex hull is a golden rectangle; we hope that FIGURE 8 will appeal to lovers of all things golden.

Lest we get carried away with the "golden" properties of the golden set, we should mention an enjoyable article by Clement Falbo **[1]** which points out that any rectangle except the square can be repeatedly subdivided in this manner to form a spiral of smaller and smaller similar rectangles. In a later section we present a family $\{H_\alpha\}$ of fractals that are made of more general spirals of this type.

Some necessary tools

In this section we present some needed facts about fractals. The most important tool is Proposition 1, which gives a sufficient condition for a set A to be a subset of a self-similar set K.

Let $X = (X, d)$ be a complete metric space, and let $\mathbb{X}$ be the set of all compact subsets of X. Given a point x in X and a subset A of X, the distance $\text{dist}(x, A)$ from x

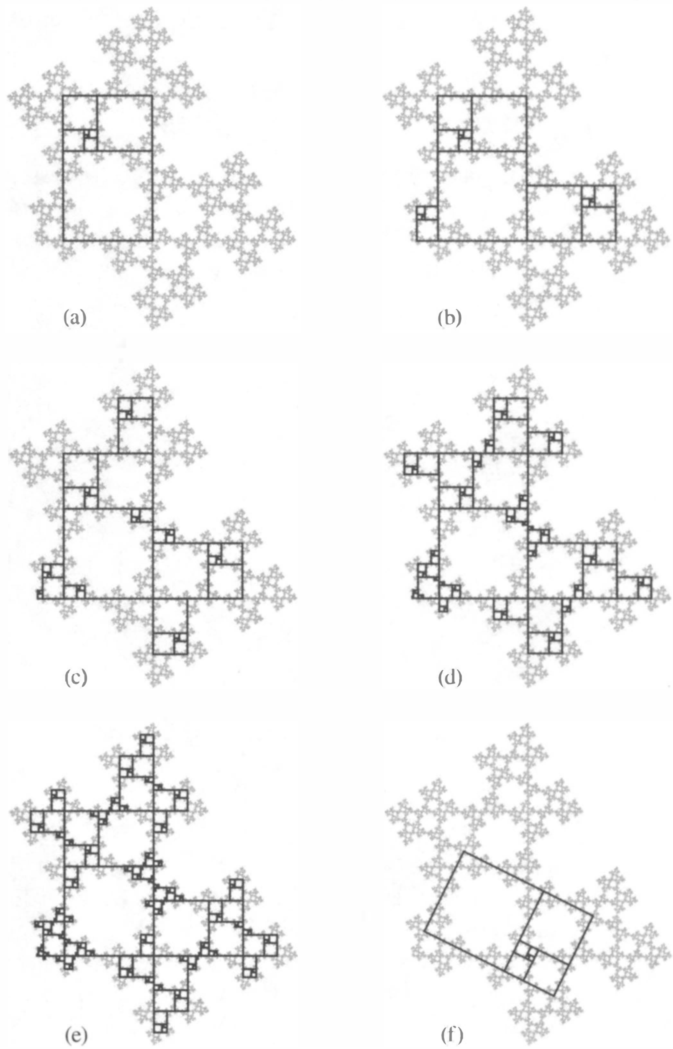


Figure 3 Construction of G by golden sets in parts (a)–(e), and an "emergent" golden set in part (f)

to A is defined by $\text{dist}(x, A) = \inf_{a\in A} d(x, a)$. The Hausdorff metric $\delta : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ is defined for all $A, B \in \mathbb{X}$ by

$$\delta(A, B) = \max\left\{\max_{a\in A} \text{dist}(a, B), \max_{b\in B} \text{dist}(b, A)\right\}.$$

Thus $\delta(A, B) < \varepsilon$ if and only if every point of A is within ε units of some point of B, and vice versa.

THEOREM 1. (HUTCHINSON [**3**]) *Let $X = (X, d)$ be a complete metric space and let $\mathcal{F} = \{f_1, \ldots, f_N\}$ be a finite set of contraction mappings on X. Then there exists a unique, closed, bounded set K such that*

$$K = \bigcup_{n=1}^{N} f_n(K). \tag{1}$$

The set K is compact and is the closure of the set of fixed points of finite compositions $f_{n_1} \circ \cdots \circ f_{n_p}$ of members of $\mathcal{F}$.

For arbitrary $A \subset X$ let $\mathcal{F}(A) = \bigcup_{n=1}^{N} f_n(A)$. For any closed, bounded, nonempty $A \subset X$, $\mathcal{F}^m(A) \to K$ in the Hausdorff metric.

Observe that (1) is equivalent to $K = \mathcal{F}(K)$. The theorem follows from the fact that $\mathcal{F}$ is a contraction mapping on the complete metric space $\mathbb{X}$. The following proposition gives us some practical tests for knowing when we can use a set A to construct K from the "inside out" with unions, or from the "outside in" with intersections. Although these results are probably not new, we give a proof for completeness.

PROPOSITION 1. *Let $\mathcal{F}$ and K be as defined in Theorem 1, and let A be a bounded, nonempty subset of X.*

The following are equivalent:

(a1) $A \subset E \subset \mathcal{F}(E)$ *for some bounded* $E \subset X$.

(a2) $A \subset K$.

(a3) $K = \mathrm{cl}\left(\bigcup_{m=0}^{\infty} \mathcal{F}^m(A)\right)$.

And the following are equivalent:

(b1) $\mathcal{F}(E) \subset E \subset A$ *for some closed, bounded, nonempty* $E \subset X$.

(b2) $K \subset A$.

(b3) $K = \bigcap_{m=0}^{\infty} \mathcal{F}^m(A)$.

Proof. It is clear that (a3) $\Rightarrow$ (a2) $\Rightarrow$ (a1) (with $E = K$), so let us prove (a1) $\Rightarrow$ (a2) $\Rightarrow$ (a3). From (a1) we have by induction that $A \subset \mathcal{F}^m(E) \subset \mathcal{F}^m(\mathrm{cl}(E))$ for all $m \in \mathbb{N}$. Thus if $a \in A$, $\mathrm{dist}(a, K) \le \delta(\mathcal{F}^m(\mathrm{cl}(E)), K) \to 0$ by Theorem 1. Therefore $a \in K$, which proves (a2). Theorem 1 then implies that $\mathcal{F}(A) \subset K$, and by induction, $K \supset \bigcup_{m=0}^{\infty} \mathcal{F}^m(A)$. For any closed subset of A, say a singleton $\{a\}$, we also have $\mathcal{F}^m(\{a\}) \to K$. It follows that any point $k \in K$ is a limit point of $\bigcup_{m=0}^{\infty} \mathcal{F}^m(A)$, which proves (a3).

Again it is obvious that (b3) $\Rightarrow$ (b2) $\Rightarrow$ (b1) with $E = K$, so let us prove (b1) $\Rightarrow$ (b2) $\Rightarrow$ (b3). From (b1) we have by induction that $\mathcal{F}^m(E) \subset E$ for all $m \in \mathbb{N}$. Thus if $k \in K$, $\mathrm{dist}(k, E) \le \mathrm{dist}(k, \mathcal{F}^m(E)) \le \delta(K, \mathcal{F}^m(E)) \to 0$. Therefore $k \in E \subset A$, which proves (b2). (This was also proved by Hutchinson [**3**, p. 727].) By Theorem 1 we then have $K \subset \mathcal{F}^m(A)$ for all $m \in \mathbb{N}$, and hence $K \subset \bigcap_{m=0}^{\infty} \mathcal{F}^m(A)$. To prove (b3), let $a \in \bigcap_{m=0}^{\infty} \mathcal{F}^m(A)$ and observe that $\mathrm{dist}(a, K) \le \delta(\mathcal{F}^m(A), K) \le \delta(\mathcal{F}^m(\mathrm{cl}(A)), K) \to 0$. ■

Remark It's important to note that the condition $A \subset \mathcal{F}(A)$ is sufficient but not necessary for A to be a subset of K. For example, let K be the Cantor middle thirds set, with $f_1(x) = x/3$ and $f_2(x) = x/3 + 2/3$. It's well known that $A = \{1/3\}$ is a subset of K, but $A \not\subset \mathcal{F}(A) = \{1/9, 7/9\}$. However, we can still apply the test by considering a carefully chosen superset E of A. For example, let $E = \{1/3, 1\}$. Then $E \subset \mathcal{F}(E)(= \{1/9, 1/3, 7/9, 1\})$, hence $A \subset E \subset K$. That's the key trick (and the fun) involved in the approach that follows. We need to develop enough familiarity with a given fractal K to conjecture that a given set A is a subset of it, and if $A \not\subset \mathcal{F}(A)$ we must then use a little creativity to construct a superset E of A such that $E \subset \mathcal{F}(E)$.

To keep the sets $f_n(K)$ from overlapping too much, it is customary to require $\mathcal{F}$ to satisfy the *open set condition*: there exists a nonempty open set $U \subset X$ such that $\mathcal{F}(U) \subset U$ and $f_m(U) \cap f_n(U) = \emptyset$ whenever $m \neq n$. Observe that by Proposition 1 we then have $K \subset \mathrm{cl}(U)$, hence $f_n(K) \subset f_n(\mathrm{cl}(U))$ for each n. Intuitively, the condition $f_m(U) \cap f_n(U) = \emptyset$ allows the sets $f_n(K)$ to touch without overlapping too much [**2**, §8.3]. We will call K *self-similar* if $\mathcal{F}$ satisfies the open set condition and $f_1, \ldots, f_N$ are contracting similitudes.

It will be convenient for us to use complex numbers, so we put $X = \mathbb{C}$. In this setting if $f_1, \ldots, f_N : \mathbb{C} \to \mathbb{C}$ are contracting similitudes of the form $f_n(z) = \alpha_n z + \beta_n$, where $\alpha_n, \beta_n \in \mathbb{C}$ and $0 < |\alpha_n| < 1$ for $n = 1, \ldots, N$, and if $\mathcal{F}$ satisfies the open set condition, then the Hausdorff dimension of K is the unique positive number D satisfying

$$\sum_{n=1}^{N} |\alpha_n|^D = 1. \tag{2}$$

In the simplest cases, the value of D corresponds with our intuitive notion of dimension. For instance, any square region K is tiled by four congruent images of itself under similitudes with scale factor $1/2$, thus satisfying (1) and the open set condition with $U = \mathrm{int}(K)$. Since $D = 2$ is the unique positive number satisfying $\sum_{n=1}^{4} |1/2|^D = 1$, the square K has Hausdorff dimension 2. On the other hand, the Cantor set just discussed has a non-integer Hausdorff dimension $D = \log_3 2 \approx 0.63$, since $\sum_{n=1}^{2} |1/3|^{\log_3 2} = 1$. We generally think of the Hausdorff dimension of a plane set as measuring the area-filling tendency of the set (even though counterexamples exist to bludgeon that intuition). Thus we would expect a set of dimension 1.9 to look more filled-in than one of dimension 1.1. Nevertheless, a plane set cannot have positive area (Lebesgue 2-dimensional measure) if its Hausdorff dimension is less than 2. A thorough discussion of Hausdorff dimension can be found in Falconer [**2**].

The golden fractal G

We will be doing arithmetic with γ, so the following rules will be convenient. Since γ satisfies the quadratic $\gamma^2 + \gamma - 1 = 0$, we find by recursion that

$$\gamma^1 = \gamma, \qquad \gamma^2 = 1 - \gamma, \qquad \gamma^3 = 2\gamma - 1, \qquad \gamma^4 = 2 - 3\gamma. \tag{3}$$

In general, if F_n denotes the nth Fibonacci number, with $F_0 = 0$, it is easy to show by induction that

$$\gamma^n = (-1)^n F_{n-1} + (-1)^{n-1} F_n \gamma$$

for $n \in \mathbb{N}$.

Before defining the set G, it will be convenient first to define a set K similar to G, using complex notation. Let $w_1, w_2, w_3 : \mathbb{C} \to \mathbb{C}$ be the contracting similitudes

$$w_1(z) = \gamma^3 z, \qquad w_2(z) = i\gamma z + \gamma, \qquad w_3(z) = -i\gamma z + \gamma. \tag{4}$$

By Theorem 1 there exists a unique compact set $K \subset \mathbb{C}$ satisfying

$$K = \bigcup_{n=1}^{3} w_n(K).$$

FIGURE 4 shows the set K. As suggested by the figure, it is easy to show that $\mathcal{W} = \{w_1, w_2, w_3\}$ satisfies the open set condition with $U = (0, 1) \times (-\gamma, \gamma)$. Specifically, we find using (3) and (4) that $w_1(U) = (0, \gamma^3) \times (-\gamma^4, \gamma^4)$, $w_2(U) = (\gamma^3, 1) \times (0, \gamma)$, and $w_3(U) = (\gamma^3, 1) \times (-\gamma, 0)$. Now let ψ be the real number satisfying $\psi^3 + 2\psi = 1$. We claim that the Hausdorff dimension D of K is given by

$$D = \log_\gamma \psi \approx 1.44. \tag{5}$$

To verify this, we plug D and the scale factors γ^3, γ, γ of w_1, w_2, w_3 into (2) to get

$$(\gamma^3)^D + \gamma^D + \gamma^D = (\gamma^{\log_\gamma \psi})^3 + 2\gamma^{\log_\gamma \psi} = \psi^3 + 2\psi = 1.$$

It is apparent that K is symmetric about the real axis. To verify this, we can consider a two-point set $A = \{a, \bar{a}\}$ which is symmetric about the real axis, where $\bar{a}$ denotes the conjugate of a. We then observe that

$$\mathcal{W}(A) = \{\gamma^3 a, \gamma^3 \bar{a}, i\gamma a + \gamma, i\gamma \bar{a} + \gamma, -i\gamma a + \gamma, -i\gamma \bar{a} + \gamma\}.$$

Thus $\mathcal{W}$ maps any conjugate pair to a set of conjugate pairs. Consequently, if $\mathcal{W}^k(A)$ is symmetric about the real axis for some $k \in \mathbb{N}$, then $\mathcal{W}^{k+1}(A)$ is also, and by induction so is $\mathcal{W}^m(B)$ for all $m \in \mathbb{N}$. Since we also have $\mathcal{W}^m(A) \to K$ in the Hausdorff metric, an easy argument shows that K is symmetric about the real axis, as claimed.

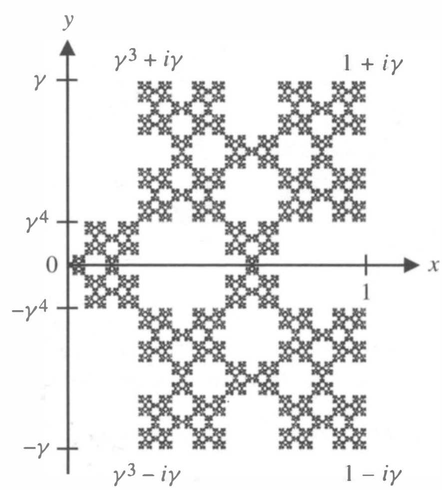


Figure 4 The set K

Having established the Hausdorff dimension and symmetry of K, we are now ready to define G. Let $h : \mathbb{C} \to \mathbb{C}$ be the similitude

$$h(z) = (1 + i\gamma)z = \sqrt{1 + \gamma^2}\, e^{i \arctan \gamma} z,$$

and let G be the set

$$G = h(K).$$

Then G is also compact with Hausdorff dimension D. Observe that h dilates K and rotates it through an angle of $\arctan \gamma$, resulting in the appearance of G as shown in FIGURE 2. In particular, G is symmetric about the line $y = \gamma x$, as claimed. To define

G in terms of contracting similitudes, observe that since h is one-to-one,

$$G = h(K) = h\left(\bigcup_{n=1}^{3} w_n(K)\right) = h\left(\bigcup_{n=1}^{3} w_n \circ h^{-1}(G)\right)$$
$$= \bigcup_{n=1}^{3} h \circ w_n \circ h^{-1}(G).$$

Therefore, we define the set $\mathcal{F} = \{f_1, f_2, f_3\}$ of contracting similitudes by $f_n = h \circ w_n \circ h^{-1}$ for $n = 1, 2, 3$, and using (4) we obtain

$$f_1(z) = \gamma^3 z, \qquad f_2(z) = i\gamma z + \gamma + i\gamma^2, \qquad f_2(z) = -i\gamma z + \gamma + i\gamma^2. \tag{6}$$

The set G is then the unique compact subset of $\mathbb{C}$ satisfying

$$G = \bigcup_{n=1}^{3} f_n(G).$$

Building G from within Refer to FIGURE 5. We begin by showing that G contains the golden rectangle whose extended sides form the black set E on the left side of Figure 5. To do so, we first specify E and then show that $\mathcal{F}(E)$, the set consisting of solid, dashed, and dotted black lines on the right side of FIGURE 5, satisfies $E \subset \mathcal{F}(E)$, allowing us to apply Proposition 1. Let E be the set consisting of the four closed line segments $\overline{z_j, z_k}$:

$$E = \overline{z_1, z_4} \cup \overline{z_2, z_7} \cup \overline{z_3, z_6} \cup \overline{z_5, z_8},$$

where

$$z_1 = 0, \quad z_2 = (2 - 3\gamma) + i(16\gamma - 10),$$
$$z_3 = \gamma + i(2 - 4\gamma), \quad z_4 = 2 - \gamma,$$
$$z_5 = (4 - 5\gamma) + i(2 - 2\gamma), \quad z_6 = \gamma + i2\gamma,$$
$$z_7 = (2 - 3\gamma) + i(8\gamma - 4), \quad z_8 = (3\gamma - 2) + i(2 - 2\gamma).$$

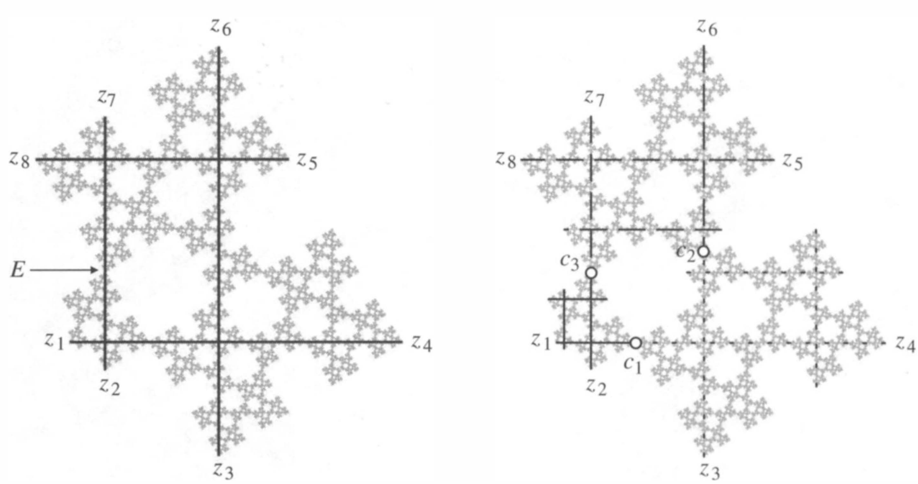


Figure 5 On the left, the set E (black), and on the right, the set $\mathcal{F}(E)$, which is the union of $f_1(E)$ (solid black), $f_2(E)$ (dashed), and $f_3(E)$ (dotted). As the figure suggests, $E \subset \mathcal{F}(E)$.

(To determine the values of the points $z_1, \ldots, z_8$ we determine their counterparts in the set K of FIGURE 4, and then apply the transformation h.)

Observe that the four line segments of E bound a rectangular region $A = [2-3\gamma, \gamma] \times [0, 2-2\gamma]$, whose sides, using (3), have lengths $4\gamma - 2 = 2(2\gamma - 1) = 2\gamma^3$ and $2 - 2\gamma = 2(1-\gamma) = 2\gamma^2$, with ratio γ. Thus the boundary $R = \partial A$ of A is a golden rectangle; this is also illustrated in FIGURE 6. Since $R \subset E$, we can use Proposition 1 to show that $R \subset G$ by showing that $E \subset \mathcal{F}(E)$. To this end, the right side of FIGURE 5 shows $\mathcal{F}(E)$, which is the union of $f_1(E)$ (solid black), $f_2(E)$ (dashed), and $f_3(E)$ (dotted). That drawing also shows the points $z_1, \ldots, z_8$, along with three more points:

$$c_1 = 7\gamma - 4, \quad c_2 = \gamma + i(1-\gamma), \quad c_3 = (2-3\gamma) + i(4-6\gamma).$$

Using the similitudes f_n in (6) and the arithmetic rules in (3), we find that

$$z_1 = f_1(z_1), \qquad z_2 = f_1(z_3), \qquad z_3 = f_3(z_4), \qquad z_4 = f_3(z_6),$$

$$z_5 = f_2(z_3), \qquad z_6 = f_2(z_4), \qquad z_7 = f_2(z_5), \qquad z_8 = f_2(z_6),$$

and

$$c_1 = f_1(z_4) = f_3(z_3), \qquad c_2 = f_2(z_1) = f_3(z_1), \qquad c_3 = f_1(z_6) = f_2(z_8).$$

It is now easy to argue that $E \subset \mathcal{F}(E)$. For example, we can conclude from the above that

$$\overline{z_1, z_4} = \overline{z_1, c_1} \cup \overline{c_1, z_4} = f_1(\overline{z_1, z_4}) \cup f_3(\overline{z_3, z_6}) \subset \mathcal{F}(E).$$

We leave the rest of the proof as an exercise.

It now follows that the golden rectangle R is a subset of G, where again

$$R = \partial([2-3\gamma, \gamma] \times [0, 2-2\gamma]). \tag{7}$$

Thus R has vertices $b_1 = 2 - 3\gamma$, $b_2 = \gamma$, $b_3 = \gamma + i(2-2\gamma)$, and $b_4 = (2-3\gamma) + i(2-2\gamma)$ (FIGURE 6). We are now ready to define the golden set S shown in FIGURE 2: Let

$$S = \bigcup_{m=0}^{\infty} f_2^m(R). \tag{8}$$

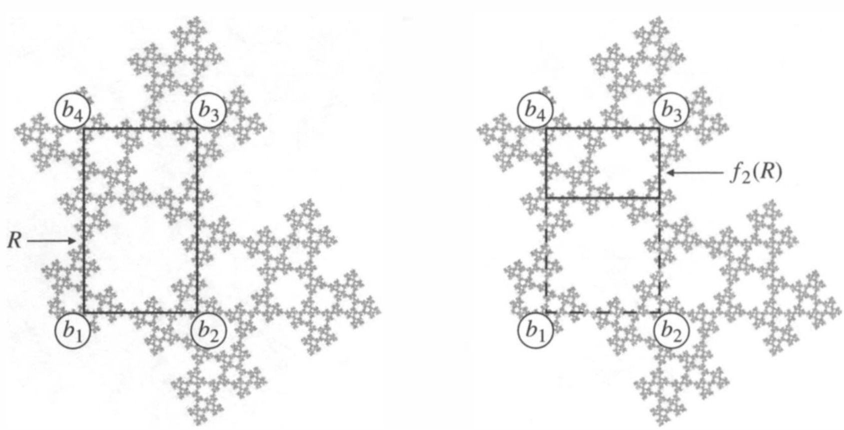


Figure 6 The golden rectangles R and $f_2(R)$

By Theorem 1, S is a subset of G. To show that S is a golden set, we determine $f_2(R)$ (FIGURE 6) by computing $f_2(b_2) = b_3$ and $f_2(b_3) = b_4$. Since f_2 involves a rotation through $\pi/2$, this means that the golden rectangle $f_2(R)$ is the first iteration of the construction of a golden set whose outer golden rectangle is R. It then follows easily that (7) and (8) define a golden subset S of G.

Finally, Proposition 1 shows that

$$G = \mathrm{cl}\left(\bigcup_{m=0}^{\infty} \mathcal{F}^m(S)\right).$$

Since each set $\mathcal{F}^m(S)$ is a finite union of golden sets, G is the closure of a countable union of golden sets, as claimed. Moreover, since the mappings f_1, f_2, f_3 involve scalings and/or rotations through $\pm\pi/2$, all the golden sets in this union have their sides parallel or perpendicular to those of S. Parts (a) through (e) of FIGURE 3 illustrate the respective sets $\mathcal{F}^0(S), \ldots, \mathcal{F}^4(S)$. As mentioned earlier, the symmetry of G about the line $y = \gamma x$ implies that $S' \subset G$, where S' is the reflection of S in the line $y = \gamma x$. Thus we also have $G = \mathrm{cl}\left(\bigcup_{m=0}^{\infty} \mathcal{F}^m(S')\right)$. That is, G is also the closure of a countable union of golden sets, none of whose sides are parallel to those of S.

Remark These results suggest that we can think of the set G as a kind of "improved" golden set, in the following sense. FIGURE 7(a) illustrates the classic method of recursively drawing the golden rectangle R to obtain the golden set S. We start with R, inscribe in it a smaller copy of R (namely $f_2(R)$) scaled by a factor of γ, and then repeat the procedure at each stage on the newest rectangle. As we have observed, this procedure fails to obtain a truly self-similar set. FIGURE 7(b) shows a simple modification of this drawing procedure, which results in the self-similar fractal G. We start with R and then add to it three scaled copies of R (namely $f_1(R)$, $f_2(R)$, and $f_3(R)$)

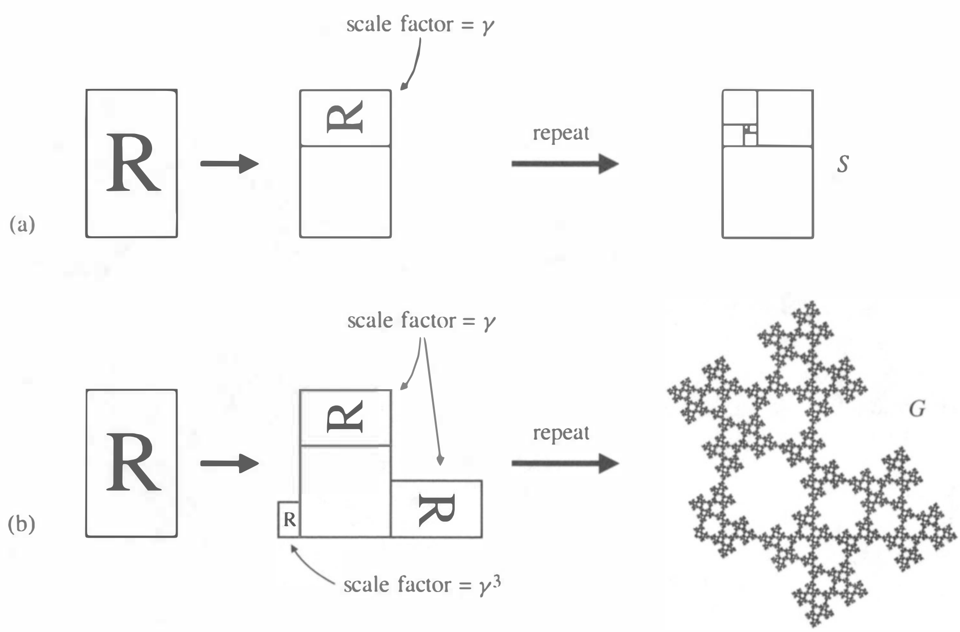


Figure 7 The old-fashioned method (a) of recursively drawing a golden rectangle R results in the not-quite self-similar golden set S. The modified method (b) results in the self-similar fractal G.

as indicated. We then repeat the procedure at each stage on the new crop of rectangles. At the kth stage there are 3^k new rectangles, however, so the drawing procedure demands more and more work. We leave it as an exercise to show that FIGURE 7(b) accurately depicts the sets $f_n(R)$. In terms of our notation, of course, FIGURES 7(a) and (b) illustrate the respective relations $S = \bigcup_{k=0}^{\infty} f_2^k(R)$ and $G = \mathrm{cl}\left(\bigcup_{k=0}^{\infty} \mathcal{F}^k(R)\right)$.

Rectangular recursion

We chose the scale and orientation of the set G for their convenience in proving that G contains a golden set. For summarizing all of the above results, however, we introduce a new set H which has an additional nice property. The set H is just the union of K in FIGURE 4 with its reflection in the imaginary axis (compare with FIGURE 8). We distill our results in the following proposition and FIGURE 8, referring to H rather than G. Intuitively, H can be built from the "outside in" by intersecting finite unions of golden rectangular regions, or from the "inside out" by uniting golden sets (which are not aligned with the golden rectangular regions). We leave the proof of Proposition 2 as an exercise, using Proposition 1, the above results, and FIGURE 8 as a guide.

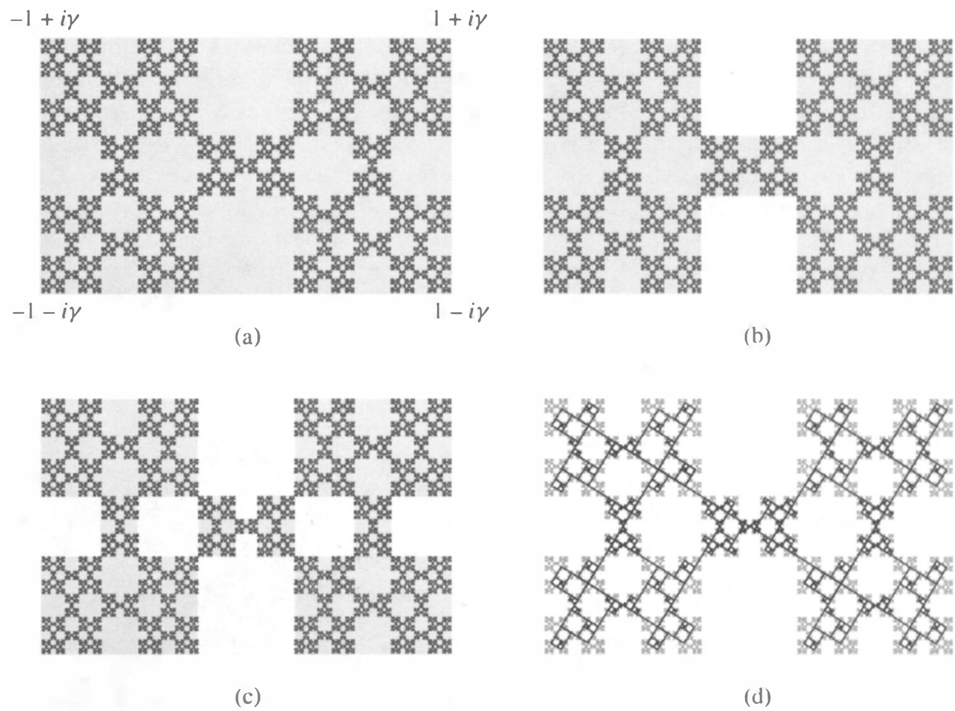


Figure 8 Parts (a)–(c) show the set H being built from the "outside in" as a countable intersection of finite unions of golden rectangular regions, starting with the shaded region $[-1, 1] \times [-\gamma, \gamma]$. Part (d) shows H being built from the "inside out" as the (closure of) a countable union of golden sets.

PROPOSITION 2. *Let $g_1, g_2, g_3 : \mathbb{C} \to \mathbb{C}$ be the contracting similitudes*

$$g_1(z) = \gamma^3 z, \qquad g_2(z) = i\gamma z + \gamma, \qquad g_3(z) = i\gamma z - \gamma, \tag{9}$$

and let H be the unique compact subset of $\mathbb{C}$ satisfying $H = \bigcup_{n=1}^{3} g_n(H)$. Then H has the following properties:

(i) *H has Hausdorff dimension* $D = \log_\gamma \psi \approx 1.44$, *where* ψ *is the real number satisfying* $\psi^3 + 2\psi = 1$.

(ii) *H is the countable intersection of finite unions of closed, golden rectangular regions aligned with the coordinate axes. In each finite union, the interiors of the golden rectangles are pairwise disjoint.*

(iii) *H is the union of the set K and its reflection in the imaginary axis. Thus H is the closure of a countable union of golden sets having sides parallel to the line* $y = -\gamma x$, *or alternatively, the line* $y = \gamma x$.

Further explorations Applications of Proposition 1 are appropriate for many students, in particular, conjecturing that a polygon or other interesting figure is a subset of a fractal, and then proving the conjecture. One place to start is with the structure of the set H: proving, for instance, that H is made of triangles of a certain shape. Classic fractals are also good subjects. For example, does the Sierpiński carpet of FIGURE 9 contain the black line segment? Can you show that the Sierpiński carpet is made of diamonds?

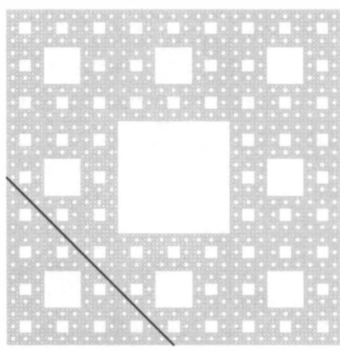


Figure 9 Does the line segment belong to the Sierpiński carpet?

One structure in H that gamma-philes will quickly notice has to do with the missing limit point of the golden set. This point is famously known to be the intersection of appropriately chosen diagonals of any mother-daughter pair of diminishing golden rectangles (FIGURE 10). When—as is usually done, and as we have done here—a golden set is defined as a union of golden rectangles (and not the closure of the union), the limit point of the diminishing rectangles is missing from the set. According to Mario Livio [4], mathematician Clifford Pickover suggests that we call this point the "Eye of

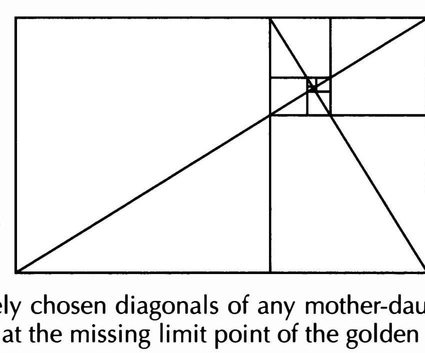


Figure 10 Appropriately chosen diagonals of any mother-daughter pair of diminishing golden rectangles meet at the missing limit point of the golden set. Many such diagonals are visible in FIGURE 8.

God" because of divine properties attributed to γ. We have established enough already (FIGURE 5) to start a proof that H contains every such pair of diagonals with respect to the shaded golden rectangles in FIGURE 8. Moreover, these pairs of diagonals also appear to meet at the "Eyes of God" of the tilted golden sets pictured in FIGURE 8(d). More things to prove!

Another good problem, of course, is to find more fractals made of golden sets. We know of another such example, and there may be more.

Yet another avenue of exploration is the following generalization of the maps in (9). Using (3) we find that $\gamma^3 = 1 - 2\gamma^2$ and $\gamma = 1 - \gamma^2$. Consequently, the following contracting similitudes, for $0 < \alpha < \sqrt{2}/2$, reduce to (9) when $\alpha = \gamma$:

$$z \mapsto (1 - 2\alpha^2)z, \qquad z \mapsto i\alpha z + (1 - \alpha^2), \qquad z \mapsto i\alpha z - (1 - \alpha^2). \tag{10}$$

The action of these maps can be depicted in terms of a rectangular initiator with corners at $\pm 1 \pm i\alpha$, analogous to the progression in FIGURE 8(a) and (b). A warm-up exercise is to draw a picture of this and verify its consistency with (10). The maps in (10) determine a family of roughly H-shaped fractals H_α that are (up to closure) unions of spirals of similar rectangles analogous to the golden set. Imitating the procedures in this article, one can prove results about the whole family $\{H_\alpha\}$ in general, or concentrate on special cases. For example, the case when $\alpha = 1/2$ is rather nice; the Hausdorff dimension of $H_{1/2}$ is the same as that of the Sierpiński triangle, but in this case $H_{1/2}$ can be characterized as the closure of a union of triangles that have the shape of ... well, enough said.

REFERENCES

1. Clement Falbo, The golden ratio—A contrary viewpoint, *College Math. J.* **36**(2) (2005) 123–134.
2. Kenneth J. Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, New York, 1985.
3. John E. Hutchinson, Fractals and self-similarity, *Indiana Univ. Math. J.* **30** (1981) 713–747. doi:10.1512/iumj.1981.30.30055
4. Mario Livio, *The Golden Ratio: The Story of Phi, The World's Most Astonishing Number*, Broadway Books, New York, 2002.
5. Manfred R. Schroeder, *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*, W. H. Freeman, New York, 1991.
6. K. K. Tung, *Topics in Mathematical Modeling*, Princeton University Press, Princeton, NJ, 2007.
7. Hans Walser, *The Golden Section*, translated from the second (1996) German edition by Peter Hilton with the assistance of Jean Pedersen, *Spectrum*, MAA, Washington, DC, 2001.

TWITTER MATH—a tiny new journal of the MAA!

- Tweaching Calculus, 140 Characters at a Time
- Tweal Analysis, a Brief, Brief Course
- Twelegraphic Tweeviews
- TweeTEX Tips: Using Twenvironments

(If you don't know how to receive this journal and submit articles, you are insufficiently cool.)

EDITOR'S (MORE SERIOUS) NOTE: Readers may have noticed the new codes at the beginning of each article. These are Digital Object Identifiers (doi), used to identify the electronic version of the article in a manner that is presumed to be more stable than a URL. Read more at www.doi.org.

A Survey of Euler's Constant

THOMAS P. DENCE
Ashland University
Ashland, OH 44805
tdence@ashland.edu

JOSEPH B. DENCE
St. Louis, MO

doi:10.4169/193009809X468689

The mathematical constant $\gamma = \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \ln(n)\right) = 0.5772156\ldots$, known as Euler's constant, is not as well known as its cousins π, e, i, but is still important enough to warrant serious consideration in the circles of applied mathematics, calculus, and number theory.

Some authors will occasionally refer to γ as the Euler-Mascheroni constant, so named after the Italian geometer Lorenzo Mascheroni (1750–1800), who actually introduced the symbol γ for the constant (although there is controversy about this claim) and also computed, though with error, the first 32 digits [**16, 34**]. Sometimes one will find in older texts the symbols C (this was Euler's constant of integration) and A (also from Mascheroni) to represent the constant, but these notations seem to have disappeared in the modern era [**27**].

Our aim in this article is to present a survey of γ that is both manageable by, and enlightening to, those who favor mathematics at the undergraduate level. To try and follow in the footsteps of the big boys π and e is quite a chore, but this brief historical description of γ and colorful portfolio of applications and surprising appearances in a multitude of settings is both impressive and mathematically educational.

Defining and evaluating the constant

Calculus students can approximate the integral $\int_1^n (1/x)\,dx = \ln(n)$ by inscribed and circumscribed rectangles, and hence obtain the inequalities (for any integer $n > 1$)

$$\frac{1}{n} < \sum_{k=1}^{n}\frac{1}{k} - \ln(n) < 1,$$

so if the limit exists,

$$0 \le \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \ln(n)\right) \le 1.$$

Furthermore, a little more algebraic manipulation gives the approximation

$$\sum_{k=1}^{n}\frac{1}{k} - \ln(n) \approx \frac{1}{2} + \frac{1}{2n},$$

so now one could presume that the limit is close to 0.5 [**23**]. Thoughts like these must have led Euler to christen the birth of γ by writing

$$\gamma = \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \ln(n)\right).$$

Editor's Note: We trust that readers will not find it too jarring to see γ, used in the previous article to mean something else entirely, appear here as Euler's constant. The usage in this article is probably the more familiar.

Several of today's calculus textbooks (though certainly not all of them) make mention of this limit primarily for two reasons: (1) it's an instance of the indeterminate expression $\infty - \infty$ for two divergent sequences and (2) it has an automatic connection with the fascinating harmonic series.

From the limiting definition of γ, we know that for any positive integer n, the sum $\sum_{k=1}^{n} 1/k$ is approximately equal to $\ln(n) + \gamma$. So, for instance, with $n = 1000$, the sum $\sum_{k=1}^{1000} 1/k$ is approximately $\ln(1000) + \gamma \approx 7.4849709$, while the TI-83 calculator gives $\sum_{k=1}^{1000} 1/k = 7.4854709$. We obtain a more accurate sum if we make use of the well-known Euler-Maclaurin formula (used extensively by *Mathematica* for series summations), which is a real gem for comparing definite integrals with infinite sums. One way to apply this flexible tool [**26**] gives

$$\sum_{k=1}^{n} \frac{1}{k} \approx \ln(n) + \gamma + \frac{1}{2n} - \frac{1}{12n^2},$$

in which case the value of the right-hand side (when $n = 1000$) is 7.485470861. Actually, the above improved approximation formula could be extended to either

$$\sum_{k=1}^{n} \frac{1}{k} \approx \ln(n) + \gamma + \frac{1}{2n} - \sum_{k=1}^{\infty} \frac{B_{2k}}{2k} \frac{1}{n^{2k}}$$

or

$$\sum_{k=1}^{n} \frac{1}{k} \approx \ln\left(n + \frac{1}{2}\right) + \gamma + \sum_{k=1}^{\infty} \frac{\left(1 - \frac{1}{2^{2k-1}}\right) B_{2k}}{2k\left(n + \frac{1}{2}\right)^{2k}},$$

where the B_{2k} are the Bernoulli numbers [**14, 23**]. One must be careful in using these approximations since both summations involving B_{2k} diverge because B_{2k} grows so rapidly. Hence the series must be truncated to get a good approximation for γ.

If, in fact, we use the basic limit definition to approximate the value of γ, we find that the rate of convergence is linear and, hence, rather slow because the difference $\sum_{k=1}^{n} \frac{1}{k} - \ln(n) - \gamma$ is approximately $\frac{1}{2n}$ [**27, 35**]. TABLE 1 presents some data.

TABLE 1

n	$\sum_{k=1}^{n} \frac{1}{k} - \ln(n)$	$\sum_{k=1}^{n} \frac{1}{k} - \ln(n) - \gamma$
10	.626383161	$.049167496 \approx 5 \times 10^{-2}$
100	.582207332	$.004991667 \approx 5 \times 10^{-3}$
1000	.577715582	$.000499917 \approx 5 \times 10^{-4}$
10000	.577265664	$.000049999 \approx 5 \times 10^{-5}$

In some relatively new results by De Temple [**15**] and Baxley [**4**], this convergence rate has been improved significantly to a quadratic rate (the difference is now approximately $\frac{1}{25n^2}$), as illustrated in TABLE 2, by simply replacing $\ln(n)$ by $\ln\left(n + \frac{1}{2}\right)$.

Here is a nice application of the idea: Given a positive number A, find the smallest positive integer N so that the Nth partial sum of the harmonic series exceeds A [**4, 5**]. For this we find that $N = 1 + \text{int}(e^{A-\gamma} - .5)$. Thus, in particular, if $A = 4$, then $N = 1 + \text{int}(e^{3.4227844} - .5) = 31$, and the 31st partial sum $S_{31} = 4.02724$ while $S_{30} = 3.99499$. Similarly, if $A = 18$ then $N = 36{,}865{,}412$ and if $A = 100$, then $N \approx 1.51 \times 10^{43}$.

TABLE 2

n	$\sum_{k=1}^{n} \frac{1}{k} - \ln\left(n+\frac{1}{2}\right)$	$\sum_{k=1}^{n} \frac{1}{k} - \ln\left(n+\frac{1}{2}\right) - \gamma$
10	.577592997	$.0003773319 \approx 4 \times 10^{-4}$
100	.577219790	$.0000041252 \approx 4 \times 10^{-6}$
1000	.577215707	$.0000000416 \approx 4 \times 10^{-8}$
10000	.577215665	$.0000000004 \approx 4 \times 10^{-10}$

Identities and expressions that involve $\ln(n)$, $\ln\left(n+\frac{1}{2}\right)$, and $\ln(n+1)$ are not uncommon, and frequently provide unique insight. De Temple **[13]**, for example, gives us a pair of sequences $\{p_n\}$, $\{q_n\}$ where

$$p_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln(n+1) + \frac{1}{2(n+1)}$$

$$q_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln\left(n + \frac{1}{2}\right).$$

Each sequence converges to γ, with $\{p_n\}$ increasing and $\{q_n\}$ decreasing, and they satisfy the inequalities $p_n < p_{n+1} < q_{n+1} < q_n < p_n + \frac{1}{8n^2}$.

Identities

There are a variety of ways to represent γ; all of them involve some form of an infinite process, whether an explicit limit, an infinite series, or an integral. For example, γ is equal to each of the following:

(1) $-\int_0^1 \ln\left(\ln\left(\frac{1}{x}\right)\right) dx$ **[22, 23, 34]**

(2) $-\int_0^\infty \frac{\ln(x)}{e^x}\, dx$ **[16, 23]**

(3) $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n)(2n)!} - \int_1^\infty \frac{\cos(x)}{x}\, dx$ **[16, 23]**

(4) $\lim_{x\to 1^+} \sum_{n=1}^{\infty} \left(\frac{1}{n^x} - \frac{1}{x^n}\right)$ **[16, 29]**

(5) $1 - \sum_{n=1}^{\infty} \left[\sum_{m=2^{n-1}+1}^{2^n} \frac{n}{(2m-1)(2m)} \right]$ **[7, 20]**

(6) $\int_0^1 \left(\frac{1}{\ln(x)} + \frac{1}{1-x}\right) dx$ **[22, 32]**

(7) $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n(n!)} - \int_1^\infty \frac{1}{xe^x}\, dx$ **[25]**

(8) $-\int_0^1 \int_0^1 \frac{1-x}{(1-xy)\ln(xy)}\, dx\, dy$ **[31]**

(9) $\displaystyle\int_0^{\infty}\left[\frac{1}{1+x}-e^{-x}\right]\frac{dx}{x}$ **[22, 34]**

(10) $\displaystyle\int_0^{1}\left[1-e^{-x}-e^{-1/x}\right]\frac{dx}{x}$ **[22, 27]**.

Most of these expressions evaluate nicely on a graphing calculator, with the exception of (4) which is quite distasteful; it's extremely difficult to get a decent approximation, primarily because of the extremely slow convergence of the series for the zeta function for values of the exponent close to 1. Expressions (1) and (6) give 5-place accuracy, as does (2) if you integrate over the interval [0, 20], while (3) gives 3-place accuracy if you sum the first 6 terms and then subtract the integral on [0, 1500]. In (9) if you express the integral as a sum of two integrals, the first integral over the domain [0, 20] and the second as the integral over $[20, \infty)$, but drop the term e^{-x} from this latter integral whose value is now $\ln(21/20)$, you get a value for γ that's accurate to 7 decimal places. Expression (7) yields 7-place accuracy if you sum the first 10 terms and then subtract the integral from [1, 20]. Finally, (10) is especially attractive because the integrand is so well behaved (it has a removable discontinuity at $x = 0$) over the interval [0, 1], and the calculator result is accurate to 7 places.

Mathematics is inundated with instances where seemingly unrelated concepts or functions are woven together to form some beautiful tapestry. Examples involving π are practically endless; a must read here is the fact-filled article by Castellanos [**10**] in this MAGAZINE. So it should not come as any great surprise to find similar relationships involving γ. One such illustration involves identity (8) from above, because making a single change of sign in the denominator produces the equality [**31**]

$$\ln\frac{4}{\pi} = -\int_0^1\int_0^1 \frac{1-x}{(1+xy)\ln(xy)}\,dx\,dy.$$

A second illustration, also involving π, centers around an infinite product expansion of $\pi/2$ similar to Viète's formula. This formula, as given by Sondow [**32**], is

$$\frac{\pi}{2} = \left(\frac{2}{1}\right)^{1/2}\left(\frac{2^2}{1\cdot 3}\right)^{1/4}\left(\frac{2^3\cdot 4}{1\cdot 3^3}\right)^{1/8}\left(\frac{2^4\cdot 4^4}{1\cdot 3^6\cdot 5}\right)^{1/16}\cdots,$$

where the quantity in parentheses in the nth factor is the product

$$\prod_{k=0}^{n}(k+1)^{(-1)^{k+1}\binom{n}{k}}.$$

Now, by simply altering the above fractional exponent in the nth factor to be $1/(n+1)$ instead of $1/2^n$, we get an infinite product expansion for e^{γ} [**28, 31**],

$$e^{\gamma} = \left(\frac{2}{1}\right)^{1/2}\left(\frac{2^2}{1\cdot 3}\right)^{1/3}\left(\frac{2^3\cdot 4}{1\cdot 3^3}\right)^{1/4}\left(\frac{2^4\cdot 4^4}{1\cdot 3^6\cdot 5}\right)^{1/5}\cdots.$$

The convergence here is quite slow; the product of the first 20 terms only yields the approximate value of .566 for γ.

The number e^{γ} has several other interesting representations. One is due to Franz Mertens (1840–1927) and involves counting the number of primes by use of the sieve of Eratosthenes [**33**],

$$e^{\gamma} = \lim_{n\to\infty}\frac{1}{\ln(n)}\prod_{p\le n}\left(\frac{p}{p-1}\right),$$

where the product is taken over all primes $p \le n$.

Relationships with classical functions

Euler spent considerable time studying the gamma function, although it was Adrien-Marie Legendre (1752–1833) who named it and supplied the symbol Γ. The popular definition for the function is [**36**]

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

with initial domain of $x > 0$. Furthermore, the domain can then be extended to all negative, nonintegral reals by the recursive relationship $\Gamma(x+1) = x\,\Gamma(x)$. An equivalent integral representation for the gamma function, with $x > 0$, is

$$\Gamma(x) = \int_0^1 \left(\ln\left(\frac{1}{t}\right)\right)^{x-1} dt,$$

while another definition, due to Karl Weierstrass (1815–1897), provides a connection between γ and Γ, namely [**3**],

$$\Gamma(x) = \frac{e^{-\gamma x}}{x} \prod_{n=1}^{\infty} \frac{e^{x/n}}{(1 + x/n)}.$$

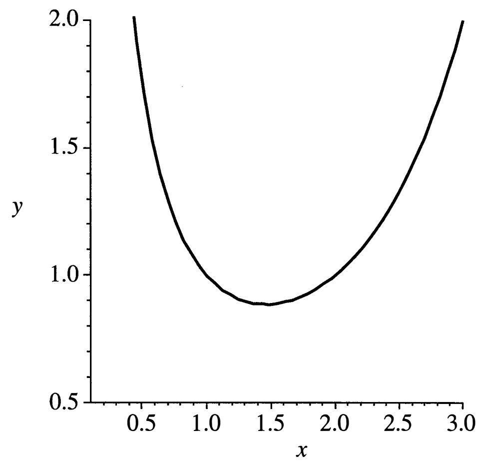


Figure 1 Graph of $y = \Gamma(x)$, $x > 0$

The graph of $y = \Gamma(x)$ in the first quadrant is a smooth curve (FIGURE 1), which is concave upward and passes through the points (1, 1) and (2, 1). Clearly the critical point of the curve is located between $x = 1$ and $x = 2$, but our current interest is with the slopes of the tangent lines at $x = 1, 2, 3, \ldots$. By computing the logarithmic derivative of Γ from Weierstrass' definition, we get

$$\frac{\Gamma'(x)}{\Gamma(x)} = -\gamma - \frac{1}{x} + \sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{x+n}\right),$$

which then gives the surprising value $\Gamma'(1) = -\gamma$ [**27**]. In general, for any integer $n \geq 2$, the derivatives have the values

$$\Gamma'(n) = (n-1)! \left[\sum_{k=1}^{n-1} \frac{1}{k} - \gamma\right],$$

so, in particular, $\Gamma'(2) = 1 - \gamma$, $\Gamma'(3) = 3 - 2\gamma$, and $\Gamma'(4) = 11 - 6\gamma$.

The gamma function, and hence Euler's constant, has mathematical connections with many other functions that play vital roles in understanding some of the deep issues in higher mathematics. A few of the classic functions from analysis include **[23]**

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x} \qquad \text{(zeta function)}$$

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2}\, du \qquad \text{(Laplace's error function)}$$

$$\operatorname{Ci}(x) = -\int_x^{\infty} \frac{\cos(u)}{u}\, du \qquad \text{(cosine integral)}$$

$$\operatorname{Li}(x) = \int_0^x \frac{1}{\ln(u)}\, du \qquad \text{(logarithmic integral).}$$

We note that Γ is related to the circular functions by the beautiful and simple formula

$$\Gamma(x)\,\Gamma(1-x) = \frac{\pi}{\sin(\pi x)}$$

for all real nonintegral x, and that Γ is related to the zeta function by **[3, 23]**

$$\Gamma(x)\,\zeta(x) = \int_0^{\infty} \frac{u^{x-1}}{e^u - 1}\, du$$

for every real x except the integers $x \leq 1$.

Euler's constant is related to the zeta function by several simple expressions

$$1 - \gamma = \sum_{n=2}^{\infty} \frac{\zeta(n) - 1}{n}$$

$$1 - \ln\left(\frac{3}{2}\right) - \gamma = \sum_{n=1}^{\infty} \frac{\zeta(2n+1) - 1}{(2n+1)4^n}$$

which, in fact, Euler used to calculate γ to, respectively, 5 and 12 decimal places **[21]**. The cosine integral has an equivalent formulation as **[1, 22]**

$$\operatorname{Ci}(x) = \gamma + \ln(x) + \int_0^x \frac{\cos(t) - 1}{t}\, dt,$$

and the logarithmic integral can be written

$$\operatorname{li}(x) = \gamma + \ln(\ln(x)) + \sum_{n=1}^{\infty} \frac{(\ln(x))^n}{n \cdot n!}.$$

Gauss used the logarithmic integral to approximate $\pi(x)$, the number of primes less than or equal to x, since he conjectured $\operatorname{li}(x)$ was asymptotic to $\pi(x)$. The British mathematician (and colleague of G. H. Hardy) J. E. Littlewood (1885–1977) studied $\operatorname{li}(x)$ and made important contributions concerning its distribution. Even though it appeared that $\pi(x) < \operatorname{li}(x)$ for all x, Littlewood proved that the inequality would be reversed (in fact, infinitely often) with the first instance occurring somewhere before x reaches $10^{10^{10^{34}}}$, known in the literature, henceforth, as the Skewes number **[9]**. Questions concerning the distribution of primes, which led to the Prime Number Theorem, and the zeros of the zeta function, which led to the Riemann Hypothesis, have been at the pinnacle of mathematics research for the past two hundred years.

Not every discussion of a topic related to γ has to involve high-level mathematics. Freshmen calculus students today have many technology resources to draw on. Their graphing calculators and computer algebra systems all come equipped with a large number of standard functions, including the function that returns the fractional part of a number. This is `fpart(x)` on the TI calculators and `frac` in *Maple*; it is commonly denoted in the literature by $\{x\}$.

The graph of $y = \{x\}$ for $x \geq 1$ is a sawtooth-like curve, with a jump discontinuity at each integer $n \geq 2$, and each segment has length $\sqrt{2}$ with slope $= 1$. Our software can easily graph $\{x\}/x^2$, also with discontinuities at each integer $n \geq 2$. Shown in FIGURE 2, this curve reminds us of the Sidney Opera House. The area between this curve and the x-axis must then be a number bounded above by $\frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots + \frac{1}{n^2} + \cdots$, which is $\frac{\pi^2}{6} - 1 \approx 0.645$. In fact, the exact value is

$$\int_1^\infty \frac{\{x\}}{x^2}\, dx = 1 - \gamma \approx 0.423.$$

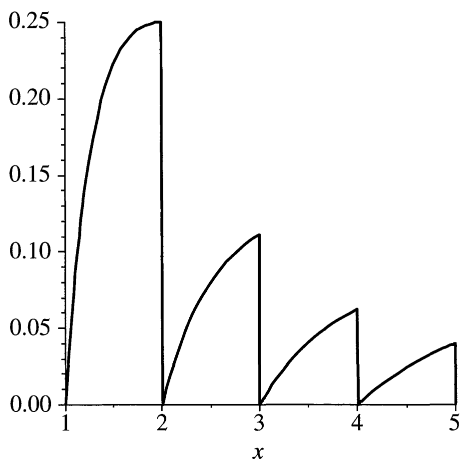


Figure 2 Graph of $y = \{x\}/x^2$, $x \geq 1$

Similarly, we can show that the integral $\int_0^1 \left\{\frac{1}{x}\right\} \ln(x)\, dx$ is equal to $\gamma + \gamma_1 - 1$, where $\gamma_1 \approx -0.072816$ is known as a "generalized Euler constant" (one of many) **[18]**. This constant is defined by

$$\gamma_1 = \lim_{n\to\infty} \left[\sum_{k=1}^{n} \frac{\ln(k)}{k} - \frac{\ln^2(n)}{2} \right].$$

The generalized Euler constants γ_n are important because they occur as coefficients in the Laurent series expansion of the zeta function, and in general are defined by **[2, 17]**

$$\gamma_n = \lim_{k\to\infty} \left[\sum_{i=1}^{k} \frac{\ln^n(i)}{i} - \frac{\ln^{n+1}(k)}{n+1} \right].$$

We note in particular that $\gamma_0 = \gamma$.

One last interesting connection between γ and analysis involves solutions to certain differential equations. Certain solutions to the ordinary differential equation

$$x^2\frac{d^2y}{dx^2}+x\frac{dy}{dx}+(x^2-p^2)y=0, \qquad p\geq 0$$

are known as Bessel functions of order p. The functions are frequently encountered in applied mathematics and physics, as well as in analytic number theory **[24]**. If, in particular, $p=0$, one series solution to $xy''+y'+xy=0$ is termed $Y_0(x)$, where $Y_0(x)$ has the complicated form, which surprisingly includes γ,

$$Y_0(x)=\frac{2}{\pi}\left[(\ln(x/2)+\gamma)\,J_0(x)+\sum_{k=0}^{\infty}(-1)^{k+1}H_k\frac{(x/2)^{2k}}{(k!)^2}\right],$$

where the function

$$J_0(x)=\sum_{k=0}^{\infty}\frac{(-1)^k x^{2k}}{2^{2k}(k!)^2} \quad \text{and} \quad H_k=\sum_{m=1}^{k}\frac{1}{m}, \qquad H_0=0.$$

The function $Y_0(x)$ is known as the Bessel function of the second kind and of order zero.

Number theory

Some discussion has already been presented that addresses connections between γ and number theory, most notably with the distribution of the primes. There are some other interesting applications. The German mathematician Peter Dirichlet (1805–1859) proved in 1838 that the totality of divisors of the integers from 1 to n numbers approximately $n\left[\ln(n)+2\gamma-1\right]$, or to rephrase it **[11, 34]**,

$$\frac{1}{n}\sum_{k=1}^{n}\tau(k)\approx\ln(n)+2\gamma-1,$$

where τ is the standard "number of divisors" function. In particular, if $n=2500$ then

$$\frac{1}{n}\sum_{k=1}^{n}\tau(k)\approx 7.9840 \quad \text{and} \quad \ln(n)+2\gamma-1\approx 7.9783,$$

and what this says is that, on the average, each integer from 1 to 2500 has about 8 divisors.

Suppose we select an integer n and a set $\mathcal{S}$ of positive integers that are all less than or equal to n and are in some arithmetic progression. We then compute the quotient n/s, for each s in $\mathcal{S}$, and let d be the decimal by which n/s falls short of being an integer; thus $d=\left\lceil\frac{n}{s}\right\rceil-\frac{n}{s}$. In 1898 the Belgian mathematician Charles de la Vallée Poussin showed that the average of all d values should be close to γ, and the approximation improves as $n\to\infty$. For instance, if $n=10000$ and $\mathbf{S}=\{3,7,11,15,\ldots,9999\}$ then

$$\frac{1}{2500}\sum_{s\in\mathcal{S}}\left(\left\lceil\frac{10000}{s}\right\rceil-\frac{10000}{s}\right)=0.57759.$$

Clearly, we could let $\mathbf{S}=\{1,2,3,4,\ldots,n\}$ since the integers are in arithmetic progression, but in fact the result also holds if $\mathcal{S}$ consists of all the primes $p\leq n$.

Another application of γ to a situation that students might encounter concerns the summing of rearrangements of the alternating harmonic series with variable signs. Using "little-oh" notation, where $a_n = o(1)$ means $\lim_{n\to\infty} a_n = 0$, we can write

$$\sum_{k=1}^{n} \frac{1}{k} = \ln(n) + \gamma + o(1).$$

Students in calculus learn that the alternating harmonic series converges to $\ln(2)$,

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = \ln(2),$$

but what about, say, the rearranged series

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots,$$

where the signs follow the pattern $+, +, -$? If we look at the $3n$th partial sum

$$\begin{aligned}
S_{3n} &= 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \cdots + \frac{1}{4n-3} + \frac{1}{4n-1} - \frac{1}{2n} \\
&= \sum_{k=1}^{4n} \frac{1}{k} - \sum_{k=1}^{2n} \frac{1}{2k} - \sum_{k=1}^{n} \frac{1}{2k} \\
&= \left[\ln(4n) + \gamma + o(1)\right] - \frac{1}{2}\left[\ln(2n) + \gamma + o(1)\right] - \frac{1}{2}\left[\ln(n) + \gamma + o(1)\right] \\
&= \ln(4) + \ln(n) - \frac{1}{2}\ln(2) - \ln(n) + o(1) \\
&= \frac{3}{2}\ln(2) + o(1) \qquad \text{since } o(1) + o(1) = o(1)
\end{aligned}$$

and, hence, $S_{3n} \to \frac{3}{2}\ln(2)$. Formulas for other related series follow similarly. Students may wish to determine the sum of the rearrangement where p positive terms are followed by n negative terms, and then the pattern repeats [**12**].

There are a variety of puzzle-type problems whose solutions involve the harmonic series [**23**]. Included here are the jeep-crossing-the-desert problem, the domino-stacking-with-maximum-overhang problem (see the front cover of Dan Bonar's book *Real Infinite Series*, published by the MAA [**6**]), and the problem of the worm crawling on a rubber rope (FIGURE 3). In this latter example, which appeared in one of Martin Gardner's "Mathematical Games" columns in *Scientific American* from a number of years ago, there is a worm, call him Willy, who crawls at the constant rate of one inch/second [**19**]. After every second, the rope is instantly stretched an additional yard in length. If Willy starts at the left end of the rope, which is initially a yard long, the question is whether he ever makes it to the right end (which seems quite impossible), and if so, when?

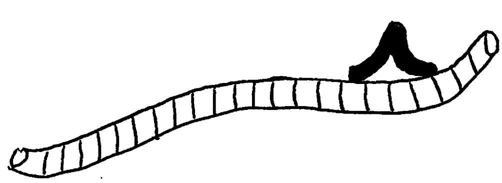


Figure 3 Worm on rubber rope

It follows that after t seconds, and before the rope is instantly stretched while Willy is taking a quick rest, his distance from the left end of the rope (i.e., back to his starting point) can be given by $t\sum_{k=1}^{t}\frac{1}{k}$. So if Willy is ever to reach the other end of the rope, then at that time the distance from the left end must agree with the total length of the rope, or

$$t\sum_{k=1}^{t}\frac{1}{k}=36t.$$

But $\sum_{k=1}^{t}\frac{1}{k}=36$ can be approximated by $\ln(t)+\gamma=36$, so $t\approx e^{36-\gamma}$. Hence, Willy does traverse the entire rope, but he's awfully old when he finishes his trek!

In closing this account of the highly functional constant γ, we would be remiss in not mentioning the all-important, and still open question of whether γ is a rational or an irrational number. Nobody in their right mind would have any support or rationale for proclaiming γ to be rational, since nothing points in that direction. Some sophisticated work with continued fractions has been used to analyze both the character of γ and e^{γ} (which Euler thought was equally as important a constant as γ) [**8**]. Furthermore, it is known [**8, 16, 23, 27**] that if either γ or $e^{\gamma}=p/q$, with integers p, q and $(p,q)=1$, then the denominator would have to be quite large. In fact, it would have to have thousands of digits; a recent result states that q would have to be larger than 10^{242080} [**7**]. Sondow gives some interesting, but deep, criteria for γ being either rational or irrational [**30**]. Furthermore, it has been shown that if γ were an algebraic number satisfying an eighth degree polynomial equation with integer coefficients, then the Euclidean norm of these coefficients would have to be a huge number, which goes against the grain for γ's being algebraic [**27**]. The safe bet is that γ is irrational, but in mathematics this claim has to be proved. This is the mathematician's way of saying, "it isn't over till the fat lady sings."

REFERENCES

1. Jean Anglesio, The integrals are Euler's constant, Prob.10561, *Amer. Math. Monthly* **105** (1998) 278–279. doi:10.2307/2589101
2. T. M. Apostol, Formulas for higher derivatives of the Riemann zeta function, *Math. Comp.* **44** (1985) 223–232. doi:10.2307/2007806
3. Emil Artin, *The Gamma Function*, Holt, Rinehart & Winston, New York, 1964, pp. 11–17.
4. J. V. Baxley, Euler's constant, Taylor's formula, and slowly converging series, this MAGAZINE **65** (1992) 302–313.
5. R. P. Boas, Jr. and J. W. Wrench, Jr., Partial sums of the harmonic series, *Amer. Math. Monthly* **78** (1971) 864–870. doi:10.2307/2316476
6. Daniel D. Bonar and Michael J. Khoury, *Real Infinite Series*, MAA, Washington, DC, 2006.
7. George Boros and V. Moll. *Irresistible Integrals: Symbolics, Analysis, and Experiments in the Evaluation of Integrals*, Cambridge Univ. Press, Cambridge, UK, 2004, 173–185.
8. Richard P. Brent, Computation of the regular continued fraction for Euler's constant, *Math. Comp.* **31** (1977) 771–777. doi:10.2307/2006010
9. David M. Burton, *Elementary Number Theory*, 6th ed., McGraw Hill, New York, 2007, p. 337.
10. Dario Castellanos, The ubiquitous π, this MAGAZINE **61** (1988) 67–98.
11. John Conway and Richard Guy, *The Book of Numbers*, Springer-Verlag, NY, 1996, p. 260.
12. Thomas P. Dence, Sums of simple rearrangements of the alternating harmonic series, *Math. Gaz.* **92** (2008) 511–514.
13. D. W. De Temple, A geometric look at sequences that converge to Euler's constant, *Coll. Math. J.* **37** (2006) 128–131.
14. D. W. De Temple and Shun-Hwa Wang, Half integer approximations for the partial sums of the harmonic series, *J. Math. Anal. App.* **160** (1991) 149–156. doi:10.1016/0022-247X(91)90296-C
15. D. W. De Temple, A quicker convergence to Euler's constant, *Amer. Math. Monthly* **100** (1993) 468–470. doi:10.2307/2324300
16. William Dunham, *Euler: The Master of Us All*, MAA, Washington, DC, 1999, pp. 35–36.

→ Rolfe Ferguson, An application of Stieltjes integration to the power series coefficients of the Riemann zeta function, *Amer. Math. Monthly* **70** (1963) 60–61. doi:10.2307/2312789
18. Ovidiu Furdui, Solutions, #164, *MO. J. Math. Sci.* **19**(2) (2007) 156–158.
19. Martin Gardner, "The Rubber Rope: Time Travel and Other Mathematical Bewilderments" (CD), MAA, Washington, DC, 2005.
→ I. Gerst, Some series for Euler's constant, *Amer. Math. Monthly* **76** (1969) 273–275. doi:10.2307/2316370
21. J. W. L. Glaisher, On the history of Euler's constant, *Mess. of Math.* **1** (1872) 25–30.
22. I. S. Gradshteyn and I. M. Ryzhik (eds.), *Tables of Integrals, Series, and Products*, Academic Press, NY, 1980, pp. 928, 946.
23. Julian Havil, *Gamma: Exploring Euler's Constant*, Princeton Univ. Press, 2003, pp. 47–48, 57–60, 79–89, 97, 105–111, 127–134.
24. F. B. Hildebrand, *Advanced Calculus for Engineers*, Prentice-Hall, Englewood Cliffs, NJ, 1949, pp. 154–157.
25. L. B. W. Jolley, *Summation of Series*, 2nd ed., Dover Publications, NY, 1961, pp. 14–17, 182–183, 206–207.
26. Vito Lampret, The Euler-Maclaurin and Taylor formulas: twin, elementary derivations, this MAGAZINE **74** (2001) 109–122.
27. Jeffrey Nunemacher, On computing Euler's constant, this MAGAZINE **65** (1992) 313–322.
28. J. Ser, Sur une expression de la function $\zeta(s)$ de Riemann, *C. R. Acad. Sci. Paris*, Ser. 1, Math, **182** (1926) 1075–1077.
29. Jonathan Sondow, An antisymmetric formula for Euler's constant, this MAGAZINE **71** (1998) 219–220.
→ Jonathan Sondow, Criteria for irrationality of Euler's constant, *Proc. Amer. Math. Soc.* **131** (2003) 3335–3344. doi:10.1090/S0002-9939-03-07081-3
31. Jonathan Sondow, Double integrals for Euler's constant and $\ln(4/\pi)$ and an analog of Hadjicostas's formula, *Amer. Math. Monthly* **112** (2005) 61–65.
→ Jonathan Sondow, A faster product for π and a new integral for $\ln(\pi/2)$, *Amer. Math. Monthly* **112** (2005) 729–734.
33. Stan Wagon, It's only natural, *Math Horizons* (Sept. 2005) 26–28.
34. Eric W. Weisstein, "Euler-Mascheroni constant," from MathWorld—A Wolfram Web Resource. http://mathworld.wolfram.com/Euler-MascheroniConstant.html.
→ R. M. Young, Euler's constant, *Math. Gaz.* **75** No. 472, (1991) 187–190. doi:10.2307/3620251
36. Dennis Zill, *A First Course in Differential Equations*, 5th ed., PWS-Kent, Boston, 1993, pp. 310–311.

Keeping Dry: The Mathematics of Running in the Rain

DANK HAILMAN
Jamaicarain University

BRUCE TORRENTS
Raindrop-Macon College*

doi:10.4169/193009809X468698

It's really coming down! As you ease into the nearest free parking spot in front of the supermarket the feeling of dread builds—you're going to get wet. But your mathematical mind wonders, what can I do, caught as I am without an umbrella, to stay as dry as possible as I make a run across the parking lot to the supermarket door? Will I stay drier if I run faster? Is there an optimal speed that will minimize my exposure?

These questions are not new, but they do have a knack for stumping people. Indeed, back in 1972, the MAGAZINE published a piece on this subject [**3**], and the following year issued a corrected version [**10**], as the original was deeply flawed. This pattern was repeated in the meteorological community in the mid 1990s, when the journal *Weather* published one piece [**7**] in 1995, and offered a corrected version (by different authors) in 1997 [**9**]. Even the television program *Mythbusters* got it wrong (episode 1) and later offered the revised conclusion that running usually trumps walking (episode 38). The results of these studies were summarized neatly in limerick by Matthew Wright [**11**] in 1995:

When caught in the rain without mac,
Walk as fast as the wind at your back,
But when the wind's in your face
The optimal pace
Is fast as your legs will make track.

In 2002, however, Herb Bailey [**2**] pointed out that the limerick above is only partially correct. It is true that in the case of a head-wind one should travel as quickly as possible. But although one does indeed stay driest by traveling "as fast as the wind at your back" in the case of a *strong* tail-wind, if the tail-wind is sufficiently weak, running "as fast as your legs will make track" is better. In fact, Bailey's argument is simply a restatement of the corrected MAGAZINE piece of 1973, where the same observation was made.

All of these analyses use a rectangular solid to model our damp traveler. In this paper we will study the prospects for more well rounded individuals. Our results for ellipsoidal travelers, for example, show that indeed, shape matters! For such travelers we take further issue with Wright's limerick. Our model suggests that in the presence of a tail-wind, however weak, it is *always* beneficial to move faster than the "wind at your back." In fact, we feel compelled to offer the following advice:

When you find yourself caught in the rain,
while walking exposed on a plane,
for greatest protection
move in the direction
revealed by a fair weather vane.

Moving swift as the wind we'll concede,
for a box shape is just the right speed.
But a soul who's more rounded
will end up less drownded
if the wind's pace he aims to exceed.

*Please direct correspondence to Dan Kalman (kalman@american.edu), American University, Washington, DC 20016, and Bruce Torrence (btorrenc@rmc.edu), Randolph-Macon College, Ashland, VA 23005.

Getting wet

We begin with some assumptions. First, since the subject has forgotten to bring an umbrella, it is reasonable to assume masculine gender. Thus we sacrifice gender neutrality and refer to him accordingly. Next we ask, how exactly is our wandering mathematician going to get wet? We will assume that rain is falling uniformly with constant velocity (no gusts). We represent this velocity by the vector $\mathbf{v}_r$ whose vertical component is negative. The key idea is this: focus on the region occupied by all the raindrops that will strike the traveler during his trip. We call this the *rain region*. The amount of water striking the traveler will be in proportion to the measure of the rain region (area in two dimensions, volume in three). Accordingly, we adopt this geometric measure as an index of total wetness.

Suppose first that our hero is standing still. Then regardless of the shape of our traveler, for each point P on his body that will be hit by a raindrop, we can draw a line segment of fixed length into space from P in the direction $-\mathbf{v}_r$ and conclude that every raindrop that will strike P in a certain time frame lies precisely on this segment. Hence the rain region is the generalized cylinder composed of the union of all such line segments (one for each point on his body that is exposed to the elements).

Now assume that our traveler moves at a constant speed $s > 0$ along a horizontal line, and adopt a distance measure so that he travels a total distance of one unit. We orient a Cartesian coordinate system in such a way that a reference point on our traveler starts at the origin and moves in the positive x direction. Thus the mathematician's velocity vector is $\mathbf{v}_m = \langle s, 0\rangle$ in a two-dimensional model, or $\mathbf{v}_m = \langle s, 0, 0\rangle$ in three. He is exposed to the elements for a finite amount of time, specifically $1/s$. The rain region consists of all initial locations from which a raindrop can land on the mathematician. Let Q be such a location, corresponding to a raindrop that will land at time t. Then it will strike the mathematician at the point $Q + \mathbf{v}_r t$. That point in turn has traveled with the mathematician from its original location $P = Q + \mathbf{v}_r t - \mathbf{v}_m t$. Thus for every exposed point P on the mathematician at time 0, the point $P + (\mathbf{v}_m - \mathbf{v}_r)t$ is in the rain region for $0 \le t \le 1/s$. This shows that the rain region is made up of line segments parallel to the *apparent* rain vector $\mathbf{v} = \mathbf{v}_r - \mathbf{v}_m$, each terminating at an exposed point on the mathematician at time 0, and each of length $\|\mathbf{v}\|/s$. A two-dimensional rendering of this scenario is shown in FIGURE 1 for two different bodies (one rectangular, one elliptical) and three different walking speeds, all in the case of a moderate head-wind. The figures correctly suggest that in these conditions, regardless of the precise shape of his body, the faster he moves, the smaller the area of the rain region, and hence the drier he stays.

Rectangular bodies In the case of a two-dimensional rectangular body as shown in the first row of FIGURE 1, the total wetness measure is simply the sum of the areas of two parallelograms, and analysis is straightforward. In the case of a three-dimensional rectangular solid body, the total wetness measure is the sum of the volumes of three parallelepipeds. That is, for each of the three exposed faces of the body, one finds the volume of the parallelepiped containing the rain that will strike this face. This volume is the product of area of the face with the magnitude of the projection of the vector $\mathbf{v}/s$ onto a line orthogonal to the face. For details, we refer the reader to Bailey **[2]** (who gives an equivalent, although less geometric analysis).

One of the advantages of our geometric approach is that it is easy to visualize the extreme cases. For a two-dimensional traveler whose speed precisely matches that of a tail-wind, the apparent rain vector is vertical. That is, all the rain that will strike our traveler lies directly above his initial position. At the other extreme, if we imagine that our mathematician is moving infinitely fast, the apparent rain vector is horizontal, for

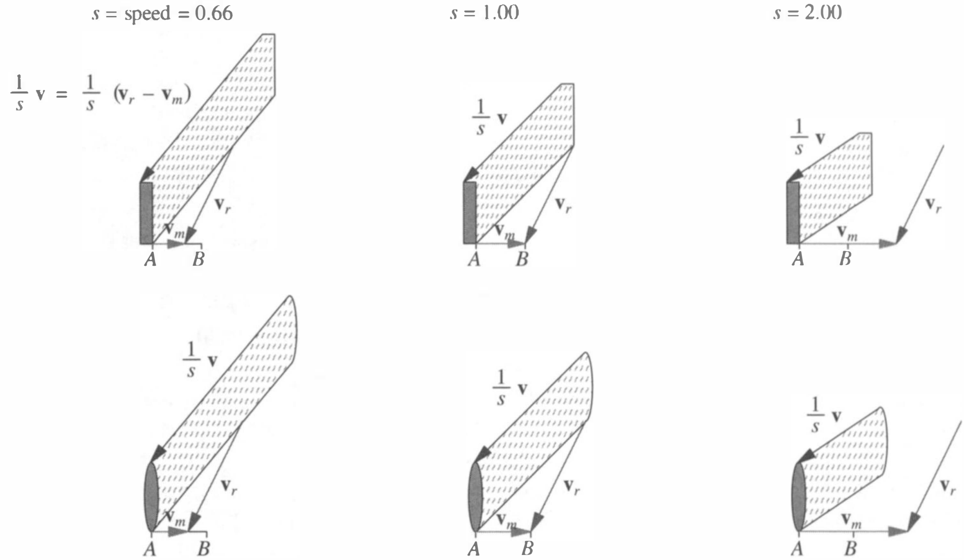


Figure 1 Two different 2-dimensional bodies (yes, 2-dimensional—do not be misled into interpreting these as 3-dimensional figures) each moving at three different speeds. One is rectangular, one elliptical, and they travel under the same rain conditions (a moderate head-wind) from point *A* to point *B*, a total distance of one unit. Total wetness is measured as the area of the *rain region*, the region containing the rain that will strike the body.

all the rain that will strike him (or, more accurately, that he will strike) lies directly in front of him. Between these extremes, where our rectangular hero's speed is finite but exceeds that of any tail-wind, the area of the parallelogram containing the rain that will strike his *front* side is exactly that of the rectangle holding the rain that would strike his front side were he moving infinitely fast. But in this case the rain region also includes the parallelogram holding the rain that will strike his top side (and the area of this parallelogram diminishes as his travel speed increases). A practical conclusion is that in the absence of a tail-wind, a body stays driest by running as fast as possible.

However, in the case of a strong tail-wind (strong here is a relative term—in three dimensions it must be at about human walking speed in the absence of a cross-wind, but stronger if there is a cross-wind—see Bailey [2]), the optimal speed of travel for a cereal-box-shaped mathematician is *precisely* the speed of the tail-wind. It is easy to show the total wetness measure T as a function of the speed s of travel, has a critical point, though not necessarily a local minimum, at the speed w of the tail-wind. FIGURE 2 illustrates this, showing the graphs of T for various cross-wind values. Note that for all cross-wind values, the limit of T as the travel speed $s \to \infty$ is simply the area of the front face of the body. This shows that all T graphs share a common horizontal asymptote. When the cross-wind is sufficiently strong, the limiting value is a lower bound for the total wetness: the faster you go, the drier you stay. But for weaker cross-winds the T curve approaches the asymptote from below. In this case, going too fast actually makes you wetter. A dynamic version appears at the MAGAZINE website.

Other body shapes Similar methods apply to bodies that are more complex polyhedral solids. For each exposed face of the body, one finds the volume of a generalized cylinder with the face as its base. Its volume is the area of this face times the magnitude of the projection of the vector $\mathbf{v}/s$ onto a line orthogonal to the face. Summing these volumes over all exposed faces gives a measure of the total amount of rain to strike

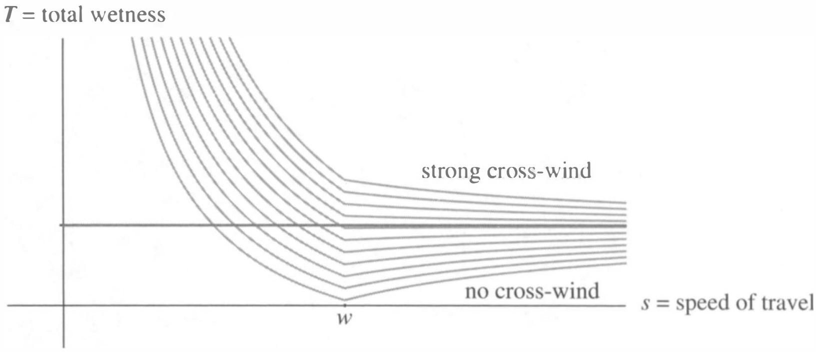


Figure 2 Total wetness as a function of s = speed of travel when traveling one distance unit. The speed w of the tail-wind is constant. Several wetness functions are shown for various cross-wind speeds. All are asymptotic to the horizontal line $y = A$, where A is the area of the front face of the traveler.

the body. For smooth surfaces, one could refine this approach into a surface integral. However, even for such simple surfaces as ellipsoids the resulting integral is difficult.

An alternate method that we pursue in the remainder of this work is to calculate the area of the projection of the body along the apparent rain vector $\mathbf{v}$ onto a plane orthogonal to $\mathbf{v}$. The volume of the rain region will be equivalent to the volume of the right cylinder whose base is this projection, and whose height is the magnitude $\|\mathbf{v}\|/s$, as illustrated in FIGURE 3 for an ellipsoidal body.

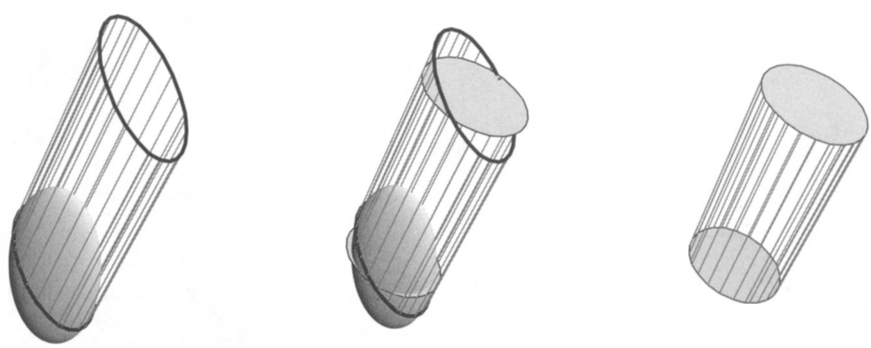


Figure 3 Orthogonal projection of the body along the apparent rain vector to obtain a right cylinder of volume equal to the rain region

Spherical bodies Few people are spherical in shape (although this is a widely accepted model for cows [**6**], and they may also wish to keep dry). Be that as it may, spherical bodies provide an irresistible temptation for modeling due to their symmetry. In the case of a spherical body, its orthogonal projection onto *any* plane through its center is always a disk of the same radius. In particular, variations in the apparent rain direction due to changes in the body's speed do not change the area of the projection. Take a spherical body of radius r, and once and for all let us fix the rain vector $\mathbf{v}_r = \langle w_t, w_c, -l\rangle$, so that a tail-wind is represented by a positive value for w_t, the cross-wind is represented by w_c, and $l > 0$ represents the downward speed of the rain. Using the projection approach outlined in FIGURE 3, our measure of total wetness is πr^2 times the magnitude of the vector $\mathbf{v}/s = (\mathbf{v}_r - \mathbf{v}_m)/s = \langle w_t - s, w_c, -l\rangle/s$. Thus we may write the total wetness function T as

$$T(s) = \frac{\pi r^2 \sqrt{(w_t - s)^2 + w_c^2 + l^2}}{s}.$$

It is easily verified that this function has a limiting value of πr^2 as $s \to \infty$, is strictly decreasing on $(0, \infty)$ when $w_t < 0$ (head-wind present), and that it has an absolute minimum at its lone critical point

$$s = \frac{w_t^2 + w_c^2 + l^2}{w_t} = \frac{\|\mathbf{v}_r\|^2}{w_t}$$

when $w_t > 0$ (tail-wind present). Thus, in contrast to the situation where the body was modeled with a rectangular solid, whenever there is a tail-wind, however weak, there is a particular speed s at which a spherical body stays driest. Moreover, since the vertical component l of the rain vector is nonzero, this speed is strictly *greater* than the speed w_t of the tail-wind. This surprising state of affairs manifests itself not only for the sphere but for capsules and ellipsoids.

R2-D2 It is a reasonably simple matter (at least in theory) to apply this analysis to a body that is a union of solids. Some shapes are particularly simple. Imagine, for instance, a capsule-shaped body composed of a right circular cylinder whose axis is parallel with the z axis, capped above and below by a hemisphere of the same radius—a bit like the Star Wars droid R2-D2. An analysis like that for the sphere shows that if there is a tail-wind, however slight, there is a definite speed s at which the body should move which minimizes the amount of rain to strike the body. As in the case of the sphere, this optimal speed is strictly greater than the speed of the tail-wind.

Ellipsoidal projections

In order to generalize the above result to ellipsoidal bodies, it is necessary to calculate the area of the orthogonal projection of the ellipsoid along the apparent rain vector. This entire section is devoted solely to this endeavor; those readers whose primary interest is in the results for ellipsoidal travelers may safely skip ahead to the next section.

For nonspherical ellipsoids, the area of the orthogonal projection of the ellipsoid along the apparent rain vector will vary as the speed of travel (and hence the apparent rain vector) changes, and so things are considerably more complicated than with the case of the sphere. Nevertheless, there turns out to be a simple formula for the area of the projection. In fact, it generalizes beautifully to n dimensions. Pursuing this more general approach, we unify the analysis for two and three-dimensional models. As a side benefit, we easily obtain results for an n-dimensional ellipsoidal mathematician dashing through the rain.

We consider an n-dimensional ellipsoid in $\mathbb{R}^n$, and will compute the $n - 1$ dimensional measure of its projection on a hyperplane orthogonal to a given vector. Not surprisingly, matrices and determinants play a central role in the derivation. To begin, we prove a computational lemma that will be useful in the main argument. This result is a special case of a more general identity for the determinant of $A + B$ when B is a rank one matrix. The general version, which is derived with a simple partitioned matrix argument in Meyer [**8**, p. 475], is equivalent to the Cauchy expansion of the determinant [**1**, pp. 74–75].

LEMMA. *For any collection $p_1, p_2, \dots, p_n$ of nonzero real numbers, and any collection $r_1, r_2, \dots, r_n$ of real numbers, the $n \times n$ matrix*

$$M = \begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{pmatrix} + \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} \begin{pmatrix} r_1 & r_2 & \cdots & r_n \end{pmatrix}$$

has determinant $p_1 p_2 \cdots p_n \left(1 + \frac{r_1^2}{p_1} + \frac{r_2^2}{p_2} + \cdots + \frac{r_n^2}{p_n}\right)$.

Proof. Let $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ denote the standard basis vectors for $\mathbb{R}^n$, and let $\mathbf{r} = \langle r_1, r_2, \dots, r_n \rangle$. Note that row j of the matrix M is the vector $p_j \mathbf{e}_j + r_j \mathbf{r}$. Since the determinant function is n-linear in the rows of M, and since each row of M is itself a sum, we may express $\mathrm{Det}(M)$ as an expansion, where each row in each matrix whose determinant appears in this expansion is either a multiple of a standard basis vector, or a multiple of $\mathbf{r}$. Note that in this expansion many terms are zero. In particular, the determinant of any matrix with two or more rows that are multiples of $\mathbf{r}$ is zero. Removing these terms, the expansion of the determinant of M has the following form:

$$\mathrm{Det}\begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{pmatrix} + r_1 \mathrm{Det}\begin{pmatrix} r_1 & r_2 & \cdots & r_n \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{pmatrix}$$

$$+ r_2 \mathrm{Det}\begin{pmatrix} p_1 & 0 & \cdots & 0 \\ r_1 & r_2 & \cdots & r_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{pmatrix} + \cdots + r_n \mathrm{Det}\begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ r_1 & r_2 & \cdots & r_n \end{pmatrix}$$

and the result follows. ∎

Proceeding with our analysis, we characterize an ellipsoid in $\mathbb{R}^n$ using matrices. Given any real symmetric $n \times n$ matrix M with positive determinant, we define the *generalized ellipsoid* for M to be the set of (row) vectors $\mathbf{x} = \langle x_1, x_2, \dots, x_n \rangle \in \mathbb{R}^n$ that satisfy the quadratic form equation $\mathbf{x} M \mathbf{x}^T = 1$. Note that there exists a real orthogonal matrix P with the property that PMP^{-1} is a diagonal matrix (with the same determinant as M). Note also that for any collection $a_1, \dots, a_n$ of positive real numbers, the diagonal matrix

$$D = \begin{pmatrix} a_1^{-2} & 0 & \cdots & 0 \\ 0 & a_2^{-2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^{-2} \end{pmatrix}$$

yields the standard generalized ellipsoid with equation $x_1^2/a_1^2 + x_2^2/a_2^2 + \cdots + x_n^2/a_n^2 = 1$. This ellipsoid has volume $U_n a_1 \cdots a_n$, where U_n is the volume of the unit sphere in $\mathbb{R}^n$ (as can be seen from the fact that the linear transformation $x \mapsto \langle a_1 x_1, \dots, a_n x_n \rangle$ with determinant $a_1 \cdots a_n$ maps the unit sphere onto this ellipsoid). Noting that for the diagonal matrix D above we have $1/\sqrt{\mathrm{Det} D} = a_1 \cdots a_n$, and that the ellipsoid associated with the matrix $P^{-1}DP$ has the same volume, we see that the ellipsoid associated with the matrix M has volume $U_n/\sqrt{\mathrm{Det} M}$.

Note that the quantity U_n can be worked out explicity, and in fact is given by the equation $U_n = 2\pi^{n/2}/(n\Gamma(n/2))$. Although this is moderately involved, things simplify nicely by considering the even and odd cases separately. For example, in even dimensions we have $U_{2n} = \pi^n/n!$. Zatzkis [**12**] gives a complete derivation of this. Also, Fraser [**5**] and Dahlka [**4**] give clever derivations that do not make use of the gamma function.

Now let us turn to projections. For any nonzero n-dimensional vector $\mathbf{w}$, let $\pi_{\mathbf{w}} : \mathbb{R}^n \to \mathbb{R}^n$ be orthogonal projection along the vector $\mathbf{w}$ onto a hyperplane of dimension $(n-1)$.

THEOREM. *Let $\mathcal{E} \subset \mathbb{R}^n$ be the generalized ellipsoid defined by*

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \cdots + \frac{x_n^2}{a_n^2} = 1,$$

and let the vector $\mathbf{v} = \langle v_1, v_2, \ldots, v_n \rangle$ belong to $\mathcal{E}$. Then the projection $\pi_{\mathbf{v}}(\mathcal{E})$ of this ellipsoid has volume

$$\frac{U_{n-1}}{\|\mathbf{v}\|} a_1 a_2 \cdots a_n.$$

Proof. Note that $\mathcal{E}$ is a level surface for the real-valued function $f(x_1, x_2, \ldots, x_n) = x_1^2/a_1^2 + x_2^2/a_2^2 + \cdots + x_n^2/a_n^2$. Hence the gradient ∇f when evaluated at a point of $\mathcal{E}$ is normal to $\mathcal{E}$ at that point. So the points of $\mathcal{E}$ satisfying the equation $\nabla f \cdot \mathbf{v} = 0$ are precisely those points for which $\mathbf{v}$ is a tangent vector. Collectively, these points of tangency form what we will call the *terminator ellipsoid* E_T. We note that E_T is indeed an ellipsoid of dimension $n-1$, since the equation $\nabla f \cdot \mathbf{v} = (v_1/a_1^2)x_1 + (v_2/a_2^2)x_2 + \cdots + (v_n/a_n^2)x_n = 0$ is that of a hyperplane, and the intersection of any generalized ellipsoid in $\mathbb{R}^n$ and any $n-1$ dimensional subspace of $\mathbb{R}^n$ is always an ellipsoid of dimension $n-1$. In FIGURE 3 the terminator ellipse is drawn with a heavy black line in each of the first two frames.

We seek the volume of the projection $\pi_{\mathbf{v}}(\mathcal{E})$. Our first observation is that the boundary of this projection is $\pi_{\mathbf{v}}(E_T)$, and hence the volume we seek is that of the ellipsoid $\pi_{\mathbf{v}}(E_T)$. We denote this ellipsoid by E, and note that in FIGURE 3 the ellipse E is drawn with a thin black line in the middle frame.

Consider again the hyperplane containing the terminator ellipsoid E_T. This hyperplane has normal vector $\mathbf{N} = \langle v_1/a_1^2, v_2/a_2^2, \ldots, v_n/a_n^2 \rangle$. Since $\mathbf{v}$ belongs to $\mathcal{E}$, at least one $v_i \neq 0$, so we may suppose without loss of generality that $v_n \neq 0$. We solve for x_n in the hyperplane equation:

$$x_n = -\frac{a_n^2}{v_n}\left(\frac{v_1}{a_1^2}x_1 + \frac{v_2}{a_2^2}x_2 + \cdots + \frac{v_{n-1}}{a_{n-1}^2}x_{n-1}\right).$$

One more ellipsoid is needed for our calculation: the vertical projection of the terminator ellipsoid, $\pi_{\langle 0,\ldots,0,1\rangle}(E_T)$, which we will call the *horizontal ellipsoid* E_H. It can be obtained by substituting the above expression for x_n into the equation of the ellipsoid $\mathcal{E}$. Essentially we just eliminate the variable x_n and obtain

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \cdots + \frac{x_{n-1}^2}{a_{n-1}^2} + \left(\frac{a_n}{v_n}\left(\frac{v_1}{a_1^2}x_1 + \frac{v_2}{a_2^2}x_2 + \cdots + \frac{v_{n-1}}{a_{n-1}^2}x_{n-1}\right)\right)^2 = 1 \quad (1)$$

Writing $\mathbf{r} = (a_n/v_n)\langle v_1/a_1^2, \ldots, v_{n-1}/a_{n-1}^2 \rangle$ and $\mathbf{x} = \langle x_1, \ldots, x_{n-1} \rangle$, and regarding vectors as matrices with one row, we have

$$\left(\frac{a_n}{v_n}\left(\frac{v_1}{a_1^2}x_1+\frac{v_2}{a_2^2}x_2+\cdots+\frac{v_{n-1}}{a_{n-1}^2}x_{n-1}\right)\right)^2 = (\mathbf{x}\cdot\mathbf{r})^2 = (\mathbf{x}\,\mathbf{r}^T)(\mathbf{x}\,\mathbf{r}^T)^T$$
$$= \mathbf{x}(\mathbf{r}^T\mathbf{r})\mathbf{x}^T.$$

Thus we may express equation 1 for E_H as the quadratic form equation

$$\mathbf{x}\left(\begin{pmatrix}\frac{1}{a_1^2} & 0 & \cdots & 0\\ 0 & \frac{1}{a_2^2} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \frac{1}{a_{n-1}^2}\end{pmatrix}+\mathbf{r}^T\mathbf{r}\right)\mathbf{x}^T = 1.$$

Note that this matrix, M, has the form of the matrix in the lemma, and hence we have

$$\begin{aligned}\mathrm{Det}(M) &= \frac{1}{a_1^2\cdots a_{n-1}^2}\left(1+\left(\frac{a_n}{v_n}\right)^2\left(\frac{v_1^2}{a_1^2}+\cdots+\frac{v_{n-1}^2}{a_{n-1}^2}\right)\right)\\ &= \frac{1}{a_1^2\cdots a_{n-1}^2}\left(1+\left(\frac{a_n}{v_n}\right)^2\left(1-\frac{v_n^2}{a_n^2}\right)\right)\\ &= \frac{1}{a_1^2\cdots a_{n-1}^2}\left(\frac{a_n}{v_n}\right)^2.\end{aligned}$$

We conclude that the volume of the horizontal ellipsoid E_H is

$$\frac{U_{n-1}}{\sqrt{\mathrm{Det}(M)}} = U_{n-1}a_1a_2\cdots a_{n-1}\frac{|v_n|}{a_n}.$$

Knowledge of the volume of $E_H = \pi_{\langle 0,\ldots,0,1\rangle}(E_T)$ allows us to find our ultimate goal, the volume of ellipsoid $E = \pi_{\mathbf{v}}(E_T)$, for both are projections of E_T. The idea is that projecting a figure scales its measure by the cosine of a certain angle. This is most easily seen in $\mathbb{R}^3$, where projecting a figure in one plane orthogonally onto a second plane scales the area by the cosine of the dihedral angle between the two planes. This same idea works in $\mathbb{R}^n$. In particular we can relate the volumes of E_T, E_H, and E, using the known volume of E_H to find the other two.

Since the projection $\pi_{\langle 0,\ldots,0,1\rangle}$ maps the terminator ellipsoid E_T onto the horizontal ellipsoid E_H, we consider the angle between their respective hyperplanes. The acute angle between the horizontal hyperplane $x_n = 0$ and the hyperplane of the terminator ellipsoid is the same as the angle between their normal vectors, provided this angle is acute. The normal vectors can be chosen as $\langle 0,\ldots,0,\pm 1\rangle$ and $\mathbf{N} = \langle v_1/a_1^2, v_2/a_2^2,\ldots,v_n/a_n^2\rangle$, where the sign in the first vector is chosen to match the sign of v_n, thus making the angle between them acute. Using the dot product, we find the cosine of this angle is

$$\frac{|v_n|}{a_n^2\,\|\mathbf{N}\|}.$$

Similarly, the projection $\pi_{\mathbf{v}}$ maps the terminator ellipsoid E_T to the ellipsoid E. The acute angle between their respective hyperplanes is the angle between $\mathbf{N}$ and $\mathbf{v}$. Its

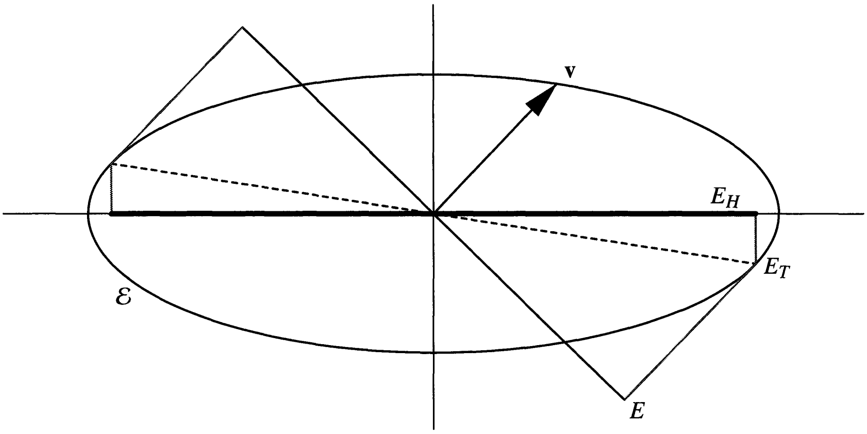


Figure 4 In two dimensions, the ellipses E_H and E shown as projections of the terminator ellipse E_T.

cosine is given by

$$\frac{v_1^2/a_1^2 + v_2^2/a_2^2 + \cdots + v_n^2/a_n^2}{\|\mathbf{v}\| \, \|\mathbf{N}\|} = \frac{1}{\|\mathbf{v}\| \, \|\mathbf{N}\|}$$

since $\mathbf{v}$ is a member of $\mathcal{E}$. Putting this all together, we see that the volume of the ellipsoid E is the volume of E_H multiplied by the ratio of the cosines, which is

$$\left(U_{n-1} a_1 a_2 \cdots a_{n-1} \frac{|v_n|}{a_n} \right) \left(\frac{a_n^2}{|v_n| \, \|\mathbf{v}\|} \right) = \frac{U_{n-1}}{\|\mathbf{v}\|} a_1 a_2 \cdots a_n.$$

Ellipsoidal bodies

Consider the ellipsoidal body $\mathcal{E}$ with equation $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$, moving as before in the positive x direction a distance of one unit with speed s, and with rain vector $\mathbf{v}_r = \langle w_t, w_c, -l \rangle$. The apparent rain vector is as before: $\mathbf{v} = \mathbf{v}_r - \mathbf{v}_m = \langle w_t - s, w_c, -l \rangle$. The measure of total wetness T as a function of s is the volume of the rain region, that is, the volume of the region containing the rain that will strike our ellipsoidal hero in the course of his journey. Specifically, it is the volume of the right cylinder whose base is the the projection $\pi_{\mathbf{v}}(\mathcal{E})$ and whose height is $\|\mathbf{v}\| / s$. To find the area of the base, we use our THEOREM, but take into account the fact that $\mathbf{v}$ may not lie on $\mathcal{E}$. Toward this end, choose $k > 0$ so that

$$k^2 = \frac{(w_t - s)^2}{a^2} + \frac{w_c^2}{b^2} + \frac{l^2}{c^2}.$$

Then $\mathbf{v}/k$ lies on $\mathcal{E}$. So by the theorem, the area of the projection $\pi_{\mathbf{v}}(\mathcal{E})$ is

$$\frac{U_2}{\|\mathbf{v}/k\|} a\,b\,c = \frac{k\pi}{\|\mathbf{v}\|} a\,b\,c$$

We multiply this area by the height of the cylinder to get volume

$$T(s) = \left(\frac{k\pi}{\|\mathbf{v}\|} a\, b\, c\right)\left(\frac{\|\mathbf{v}\|}{s}\right) = \frac{\pi}{s} k\, a\, b\, c \;=\; \frac{\pi}{s}\sqrt{k^2a^2b^2c^2}$$
$$= \frac{\pi}{s}\sqrt{b^2c^2(w_t - s)^2 + a^2c^2w_c^2 + a^2b^2l^2}.$$

This formula reduces to that found earlier for spherical bodies when $a = b = c$. It is readily verified that this total wetness function T has a limiting value of πbc (the area of the projection of the ellipsoid onto the yz plane) as $s \to \infty$, is strictly decreasing on $(0, \infty)$ when $w_t \le 0$ (no tail-wind), and that it has an absolute minimum at its lone critical point

$$s_{\text{opt}} = \frac{b^2c^2w_t^2 + a^2c^2w_c^2 + a^2b^2l^2}{b^2c^2w_t}$$

when $w_t > 0$ (tail-wind present). This optimal speed is again strictly *greater* than the speed w_t of the tail-wind. Moreover, if the traveler becomes very skinny from back to front ($a \to 0$), the optimal speed approaches w_t.

For example, consider an ellipsoid of roughly human proportions, with $a = 1$, $b = 2$, and $c = 6$ (units are not important; it is only the relative dimensions that are relevant). And imagine rain conditions where the vertical downward rainfall speed is $l = 12$ mph, with a tail-wind $w_t = 5$ mph and a cross-wind $w_c = 5$ mph. In this case the total wetness measure $T(s)$ is minimized when the body moves at a speed of $s = 7.05$ mph, well above the speed of the tail-wind! (See the FIGURE 5, where the wetness at speeds $s = 5$ and 7.05 are highlighted).

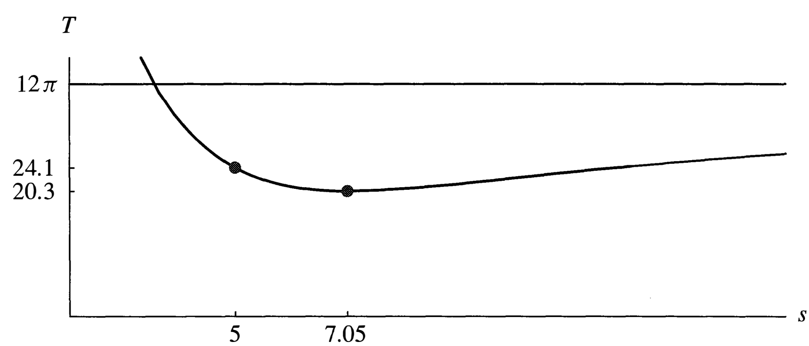


Figure 5 With a 5 mph tail-wind, this elliptical body stays driest by traveling at approximately 7 mph. The MAGAZINE website hosts a dynamic version of the figure.

Moreover, it is definitely advantageous to the body to move at this speed rather than the speed of the tail-wind; in this example the body gets roughly 19% wetter when moving at the speed of the tail-wind instead of the optimal speed. This unexpected result is contrary to that predicted by the rectangular solid model, where total wetness is minimized when the body moves precisely at the speed of the tail-wind (provided the tail-wind is sufficiently strong, as it is in this example).

For an ellipsoidal traveler moving in tail-wind conditions, then, there are three travel speeds worthy of our attention: the speed w_t of the tail-wind, the optimal speed s_{opt}, and the traveler's top running speed, which we will denote by $s_{\max}$. Suppose that conditions are such that $0 < w_t < s_{\max}$. We now investigate how much wetter a wandering ellipsoid can possibly get by traveling at the less-than-ideal speeds w_t or $s_{\max}$ than he would get by proceeding at the optimal pace.

The example above suggests that one can get roughly 20% wetter by traveling at the speed w_t of the tail-wind. Moreover, these are in conditions where the rectangular solid model recommends that the traveler move at precisely the speed of the tail-wind. Under the body dimensions considered above, and under the constraints that our hero walk no more slowly than human walking speed (say 3 mph) and that the weather conditions favor having a box-shaped body optimally travel at the speed of the tail-wind, this value of 20% is near the upper limit on how much wetter an ellipsoid will get by slowing to the speed of the tail wind versus traveling at the faster optimal pace.

Under these conditions, however, our mathematician would get only slightly wetter by running flat-out than he would by hitting the optimal tempo. In the previous example, for instance, running at 9 mph (a brisk pace to sustain in slippery conditions) gets him a little over 5% wetter than moving at the best pace. Will this always be the case, or are there atmospheric conditions when moving at the optimal pace keeps an ellipsoid *significantly* drier than running flat-out? To investigate this, consider the ratio

$$R = \frac{T(s_{\max})}{T(s_{\mathrm{opt}})}.$$

This ratio measures how much wetter a running body will get than a body traveling at the optimal pace. In the case of either ellipsoidal or cereal-box-shaped travelers, it is a simple matter to deduce that this ratio is maximized when the cross-wind $w_c = 0$. In other words, the traveler is moving precisely in the direction of the wind. Equivalently, we need only consider a two-dimensional model. In the case of an elliptical traveler with semi-axes a and c moving when there is a tail-wind at speed w, it is straightforward to calculate

$$R = \frac{\sqrt{(a^2l^2 + c^2w^2)(a^2l^2 + c^2(s_{\max} - w)^2)}}{a\,c\,l\,s_{\max}}.$$

Noting that the numerator is symmetric in w and $s_{\max} - w$, it is a simple matter to deduce that R attains its maximum precisely when

$$w = \frac{s_{\max}}{2}$$

where it attains a maximum value of

$$R_{\max} = \frac{al}{cs_{\max}} + \frac{cs_{\max}}{4al}.$$

In other words, the wind conditions under which an elliptical traveler will pay the highest price for running flat-out as opposed to moving at the optimal pace is when the tail-wind speed is exactly *half* the traveler's top running speed. A similar calculation for rectangular travelers was carried out by Schwartz and Deakin [**10**]. If one then substitutes the human-like values $a = 1$ and $c = 6$, and uses a top running speed $s_{\max} =$ 9 mph and a vertical rainfall velocity $l = 12$ mph, one finds that in the worst case (a 4.5 mph tail-wind), R is approximately 1.34. That is, our elliptical traveler cannot get more than 34% wetter when running as opposed to traveling at the optimal pace. Traveling at the optimal pace, therefore, *can* keep him much drier than running. We note that the ratio R is sensitive to changes in both $s_{\max}$ and l. The ratio will be even higher in light rain conditions (small l), and will be also be higher for faster runners.

Using these same relative dimensions in the case of a rectangular traveler, however, and using the formula for this ratio provided by Schwartz and Deakin [**10**], the maximal ratio is approximately 1.8. In both cases the body stays significantly drier by moving at the optimal pace appropriate for his body shape. But the penalty for running flat-out is greater for boxes than for ellipsoids!

Keeping dry

Before addressing the subtleties suggested by these models, it is important to understand that all of this only applies to relatively short stints in the rain. One can only get so wet before being saturated, so that additional rain just runs off. (In this instance, the mathematician is said to be all wet.) Furthermore, if one walks in the rain when winds are light, or moves near the speed of a tail-wind, one's head takes most of the water (unless there is a calculus book on top of it), and eventually it will drip down to the face and body. Our models don't take this redistribution of water into account—again: short stints in the rain.

In the absence of a tail-wind, regardless of body shape, one stays driest by traveling as quickly as possible. With a tail-wind, the ellipsoidal model suggests traveling slightly *faster* than the speed of the tail-wind, as if to outrun the rain. This differs from the rectangular solid analysis, which suggests that the body move at the *same* speed as the tail-wind (or move as fast as possible if the cross-wind is sufficiently high or the tail-wind sufficiently weak). Our first conclusion, then, is simply that *shape matters.* Moreover, in the case of an ellipsoidal traveler, small perturbations to the lengths of the axes of the ellipsoid will change the optimal speed of travel (for rectangular solids in strong tail-wind conditions, the optimal pace is simply the speed of the tail-wind, and so is immune to any changes in the dimensions of the traveler). Again, the ellipsoidal model reveals the simple truth that shape matters.

But in practical terms, the models considered here suggest that conditions where traveling at some optimal speed (related to the speed of a tail-wind) will keep one *significantly* drier than running at full speed are rare. In particular, a tail-wind must exist but be about half of one's top running speed, and the cross-wind must be minimal for this effect to be apparent. However, in these ideal conditions both the rectangular and ellipsoidal models suggest that a traveler will stay significantly drier by moving at the optimal pace, especially in a light rain.

Our recommendation, therefore, is to *RUN* in the rain unless you find yourself traveling in the perfect storm—where the tail-wind is half your top running speed, the cross-wind is minimal, and the rainfall is light. In such conditions, given the rounded features of the human body, it might make sense to dampen your pace (so to speak) from a run down to a speed that is just a bit *faster* than that of the tail-wind.

REFERENCES

1. A.C. Aitken, *Determinants and Matrices*, 9th ed., Wiley, New York, 1956.
2. Herb Bailey, On running in the rain, *College Math. J.* **33** (2002) 88–92. doi:10.2307/1558989
3. Michael A.B. Deakin, Walking in the rain, this MAGAZINE **45** (1972) 246–253.
4. Karl Dahlke, Integral calculus, The volume of the hypersphere, The Math Reference Project, http://www.mathreference.com/ca-int,hsp.html.
5. Marshall Fraser, The grazing goat in n dimensions, *College Math. J.* **15** (1984) 126–134. doi:10.2307/2686517
6. John Harte, *Consider a Spherical Cow*, W. Kaufmann, US, 1995.
7. J. J. Holden, S. E. Belcher, A. Horvath, and I. Pytharoulis, Raindrops keep falling on my head, *Weather* **50** (1995) 367–370.
8. Carl D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.
9. T. C. Peterson, and T. W. R. Wallis, Running in the rain, *Weather* **52** (1997) 93–96.
10. B. L. Schwartz and M. A. B. Deakin, Walking in the rain, reconsidered, this MAGAZINE **46** (1973) 272–276.
11. Matthew Wright, *New Scientist* **1960** (1995) 57.
12. Henry Zatzkis, Volume and surface of a sphere in an n-dimensional Euclidean space, this MAGAZINE **30** (1957) 155–157.

NOTES

Starting with Two Matrices

GILBERT STRANG
Massachusetts Institute of Technology
Cambridge, MA 02139
gs@math.mit.edu

doi:10.4169/193009809X468742

Imagine that you have never seen matrices. On the principle that examples are amazingly powerful, we study two matrices A and C. The reader is requested to be exceptionally patient, suspending all prior experience—and suspending also any hunger for precision and proof. Please allow a partial understanding to be established first.

The first sections of this paper represent an imaginary lecture, very near the beginning of a linear algebra course. That lecture shows by example where the course is going. The key ideas of linear algebra (and the key words) come very early, to point the way. My own course now includes this lecture, and Notes 1-6 below are addressed to teachers.

A first example Linear algebra can begin with three specific vectors a_1, a_2, a_3:

$$a_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad a_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \qquad a_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The fundamental operation on vectors is to take *linear combinations*. Multiply these vectors a_1, a_2, a_3 by numbers x_1, x_2, x_3 and add. This produces the linear combination $x_1a_1 + x_2a_2 + x_3a_3 = b$:

$$x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{1}$$

The next step is to rewrite that vector equation as a matrix equation $Ax = b$. Put a_1, a_2, a_3 into the columns of a matrix and put x_1, x_2, x_3 into a vector:

$$\text{Matrix } A = \begin{bmatrix} & & \\ a_1 & a_2 & a_3 \\ & & \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad \text{Vector } x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Key point A times x is exactly $x_1a_1 + x_2a_2 + x_3a_3$, *a combination of the columns.* This definition of Ax brings a crucial change in viewpoint. At first, the xs were multiplying the as. Now, the matrix A is multiplying x. The matrix acts on the vector x to produce a vector b:

$$\boldsymbol{Ax = b} \quad Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{2}$$

When the xs are known, the matrix A takes their differences. We could imagine an unwritten $x_0 = 0$, and put in $x_1 - x_0$ to complete the pattern. A is a *difference matrix.*

Note 1 If students have seen Ax before, it was probably *row times column*. In examples they are free to compute that way (as I do). "Dot product with rows" gives the same answer as "combination of columns": When the combination $x_1 a_1 + x_2 a_2 + x_3 a_3$ is computed one component at a time, we are using the rows. The relation of the rows to the columns is truly at the heart of linear algebra.

Note 2 Three basic questions in linear algebra, and their answers, show why the column description of Ax is so essential:

- When does a linear system $Ax = b$ have a solution? This system asks us to express b as a combination of the columns of A. So there is a solution exactly when b is in the *column space* of A.
- When are vectors $a_1, \dots, a_n$ linearly independent? The combinations of $a_1, \dots, a_n$ are the vectors Ax. For independence, $Ax = 0$ must have only the zero solution. The *nullspace* of A must contain only the vector $x = 0$.
- How do you express b as a combination of basis vectors? Put those basis vectors into the columns of A. Solve $Ax = b$.

Note 3 The reader may object that we have answered questions only by introducing new words. My response is that these new words are crucial definitions in this subject and the student moves to a higher level—a subspace level—by working with the column space and the nullspace in examples.

I don't accept that inevitably "The fog rolls in" when linear independence is defined [**1**]. The concrete way to dependence vs. independence is through $Ax = 0$: many solutions or only the solution $x = 0$. This comes immediately in returning to the example.

Suppose the numbers x_1, x_2, x_3 are not known but b_1, b_2, b_3 are known. Then $Ax = b$ becomes an equation for x, not an equation for b. We start with the differences (the bs) and ask which xs have those differences. This is a new viewpoint of $Ax = b$, and linear algebra is always interested first in $b = 0$:

$$\boldsymbol{Ax = 0} \quad Ax = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{leads to} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{3}$$

For this matrix, the only solution to $Ax = 0$ is $x = 0$. That may seem automatic but it's not. A key word in linear algebra (we are foreshadowing its importance) describes this situation. These column vectors a_1, a_2, a_3 are *independent.* Their combination $x_1a_1 + x_2a_2 + x_3a_3$ is $Ax = 0$ only when all the xs are zero.

Move now to nonzero differences $b_1 = 1$, $b_2 = 3$, $b_3 = 5$. Is there a choice of x_1, x_2, x_3 that produces those differences $1, 3, 5$? Solving the three equations in forward order, the xs are 1, 4, 9:

$$\boldsymbol{Ax = b} \quad \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \quad \text{leads to} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix}. \tag{4}$$

This case $x = 1, 4, 9$ has special interest. When the bs are the odd numbers in order, the xs are the perfect squares in order. But linear algebra is not number theory—forget that special case! For any b_1, b_2, b_3 there is a neat formula for x_1, x_2, x_3:

$$\begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{leads to} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_1 + b_2 \\ b_1 + b_2 + b_3 \end{bmatrix}. \tag{5}$$

This general solution includes the examples $b = 0, 0, 0$ (when $x = 0, 0, 0$) and $b = 1, 3, 5$ (when $x = 1, 4, 9$). One more insight will complete the example.

We started with a linear combination of a_1, a_2, a_3 to get b. Now b is given and equation (5) goes backward to find x. Write that solution with three new vectors whose combination gives x. Then write it using a matrix:

$$x = b_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + b_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = Sb. \tag{6}$$

This is beautiful, to see a *sum matrix* S in the formula for x. The equation $Ax = b$ is solved by $x = Sb$. We call the matrix S the *inverse* of the matrix A. The difference matrix is inverted by the sum matrix. *Where A took differences of x_1, x_2, x_3, the new matrix $S = A^{-1}$ takes sums of b_1, b_2, b_3.*

Note 4 I believe there is value in *naming* these matrices. The words "difference matrix" and "sum matrix" tell how they act. It is the action of matrices, when we form Ax and Cx and Sb, that makes linear algebra such a dynamic and beautiful subject.

The second example This example begins with almost the same three vectors—only one component is changed:

$$c_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad c_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \quad c_3 = \begin{bmatrix} \mathbf{-1} \\ 0 \\ 1 \end{bmatrix}.$$

The combination $x_1c_1 + x_2c_2 + x_3c_3$ is again a matrix multiplication Cx:

$$Cx = \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \mathbf{-1} \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{7}$$

With the new vector in the third column, C is a cyclic difference matrix. Instead of $x_1 - 0$ we have $x_1 - x_3$. The differences of xs "wrap around" to give the new bs. The inverse direction begins with b_1, b_2, b_3 and asks for x_1, x_2, x_3.

We always start with 0, 0, 0 as the bs. You will see the change: Nonzero xs can have zero differences. As long as the xs are equal, all their differences will be zero:

$$\mathbf{Cx = 0} \quad \begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{is solved by} \quad x = \begin{bmatrix} x_1 \\ x_1 \\ x_1 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{8}$$

The zero solution $x = 0$ is included (when $x_1 = 0$). But $1, 1, 1$ and $2, 2, 2$ and π, π, π are also solutions—all these constant vectors have zero differences and solve $Cx = 0$. The columns c_1, c_2, c_3 are *dependent* and not independent.

In the row-column description of Ax, we have found a vector $x = (1, 1, 1)$ that is perpendicular to every row of A. The columns combine to give $Ax = 0$ when x is perpendicular to every row.

This misfortune produces a new difficulty, when we try to solve $Cx = b$:

$$\begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{cannot be solved unless} \quad b_1 + b_2 + b_3 = 0.$$

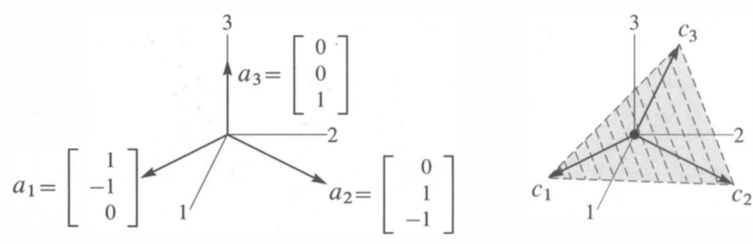
The three left sides add to zero, because x_3 is now canceled by $-x_3$. So the bs on the right side must add to zero. There is no solution like equation (5) for every b_1, b_2, b_3. There is no inverse matrix like S to give $x = Sb$. The cyclic matrix C is *not invertible.*

Summary Both examples began by putting vectors into the columns of a matrix. Combinations of the columns (with multipliers x) became Ax and Cx. Difference matrices A and C (noncyclic and cyclic) multiplied x—that was an important switch in thinking. The details of those column vectors made $Ax = b$ solvable for all b, while $Cx = b$ is not always solvable. The words that express the contrast between A and C are a crucial part of the language of linear algebra:

The vectors a_1, a_2, a_3 are independent.
The nullspace of A (solutions of $Ax = 0$) contains only $x = 0$.
The equation $Ax = b$ is solved by $x = Sb$.
The square matrix A has the inverse matrix $S = A^{-1}$.

The vectors c_1, c_2, c_3 are dependent.
The nullspace of C contains every "constant vector" x_1, x_1, x_1.
The equation $Cx = b$ cannot be solved unless $b_1 + b_2 + b_3 = 0$.
C has no inverse matrix.

A picture of the three vectors, a_1, a_2, a_3 on the left and c_1, c_2, c_3 on the right, explains the difference in a useful way. On the left, the three directions are *independent.* The three arrows don't lie in a plane. The combinations $x_1a_1 + x_2a_2 + x_3a_3$ produce every three-dimensional vector b. The multipliers x_1, x_2, x_3 are given by $x = Sb$.



On the right, the three arrows do lie in a plane. The vectors c_1, c_2, c_3 are *dependent.* Each vector has components adding to $1 - 1 = 0$, so all combinations of these vectors will have $b_1 + b_2 + b_3 = 0$ (this is the equation for the plane). The differences $x_1 - x_3$ and $x_2 - x_1$ and $x_3 - x_2$ can never be 1, 1, 1 because $1 + 1 + 1$ is not 0.

Note 5 These examples illustrate one way to teach a new subject: *The ideas and the words are used before they are fully defined.* I believe we learn our own language this way—by hearing words, trying to use them, making mistakes, and eventually getting it right. A proper definition is certainly needed, it is not at all an afterthought. But maybe it is an afterword.

Note 6 Allow me to close by returning to Note 1: Ax is a combination of the columns of A. Extend that matrix-vector multiplication to *matrix-matrix*: If the columns of B are b_1, b_2, b_3 then the columns of AB are Ab_1, Ab_2, Ab_3.

The crucial fact about matrix multiplication is that $\boldsymbol{(AB)C = A(BC)}$. By the previous sentence we may prove this fact by considering one column vector c.

$$\textit{Left side} \qquad (AB)c = [Ab_1 \;\; Ab_2 \;\; Ab_3]\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = c_1\, Ab_1 + c_2\, Ab_2 + c_3\, Ab_3 \tag{9}$$

$$\textit{Right side} \qquad A(Bc) = A(c_1 b_1 + c_2 b_2 + c_3 b_3). \tag{10}$$

In this way, $(AB)C = A(BC)$ brings out the even more fundamental fact that matrix multiplication is linear: (9) = (10).

Expressed differently, the multiplication AB has been defined to produce the composition rule: AB acting on c is equal to A acting on B acting on c.

Time after time, this associative law is the heart of short proofs. I will admit that these "proofs by parenthesis" are almost the only ones I present in class. Here are examples of $(AB)C = A(BC)$ at three key points in the course. (I don't always use the ominous word *proof* in the video lectures [**2**] on `ocw.mit.edu`, but the reader will see through this loss of courage.)

- If $AB = I$ and $BC = I$ then $C = A$.
  Right inverse = *Left inverse*, because $C = (AB)C = A(BC) = A$.
- If $y^{\mathrm{T}}A = 0$ then y is perpendicular to every Ax in the column space.
  Nullspace of A^{T} $\perp$ *column space of* A, because $y^{\mathrm{T}}(Ax) = (y^{\mathrm{T}}A)x = 0$.
- If an invertible B contains eigenvectors b_1, b_2, b_3 of A, then $B^{-1}AB$ is diagonal.
  Multiply AB by columns, $A[b_1 \;\; b_2 \;\; b_3] = [Ab_1 \;\; Ab_2 \;\; Ab_3] = [\lambda_1 b_1 \;\; \lambda_2 b_2 \;\; \lambda_3 b_3]$.
  Then separate this AB into B times the eigenvalue matrix Λ:

$$AB = [\lambda_1 b_1 \;\; \lambda_2 b_2 \;\; \lambda_3 b_3] = [b_1 \;\; b_2 \;\; b_3]\begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix} \qquad \text{(again by columns!)}.$$

Since $AB = B\Lambda$, we get the diagonalization $B^{-1}AB = \Lambda$ and the factorization $A = B\Lambda B^{-1}$. Parentheses are not necessary in any of these triple factorizations:

Spectral theorem for a symmetric matrix	$A = Q\Lambda Q^{\mathrm{T}}$
Elimination on a symmetric matrix	$A = LDL^{\mathrm{T}}$
Singular Value Decomposition (SVD) of any matrix	$A = U\Sigma V^{\mathrm{T}}$

One final comment: Factorizations express the central ideas of linear algebra in a very effective way. The eigenvectors of a symmetric matrix can be chosen orthonormal: $Q^{\mathrm{T}}Q = I$ in the spectral theorem $A = Q\Lambda Q^{\mathrm{T}}$. For all matrices, eigenvectors of AA^{T} and $A^{\mathrm{T}}A$ are the columns of U and V in the SVD. And our favorite rule $(AA^{\mathrm{T}})A = A(A^{\mathrm{T}}A)$ is the key step in establishing that factorization $U\Sigma V^{\mathrm{T}}$, long after this early lecture . . . .

These orthonormal vectors $u_1, \ldots, u_m$ and $v_1, \ldots, v_n$ are perfect bases for the *Four Fundamental Subspaces*: the column space and nullspace of A and A^{T}. Those subspaces become the organizing principle of the course [**2**]. The Fundamental Theorem of Linear Algebra connects their dimensions to the rank of A. The flow of ideas is from numbers to vectors to subspaces. Each level comes naturally, and everyone can get it—by seeing examples.

REFERENCES

1. David Carlson, Teaching linear algebra: must the fog always roll in?, *College Math. J.* **24** (1993) 29–40. doi: 10.2307/2686429
2. Gilbert Strang, *Introduction to Linear Algebra*, 4th ed., Wellesley-Cambridge Press (2009). Video lectures on http://web.mit.edu/18.06 and on http://ocw.mit.edu.

To Buy or Not to Buy: The Screamin' Demon Ticket Game

FREDERICK CHEN
Department of Economics
Wake Forest University
Winston-Salem, NC 27109-7505

ALEXANDER KAYS
7831 Montreal Court
Cincinnati, OH 45241

doi:10.4169/193009809X468733

In this note, we use game theory to study the incentives that college students have for purchasing season tickets to their schools' football games if the schools implement a lottery system whereby students can try to win free tickets by entering an online drawing.

Picked by the media in the preseason to finish last in its division, the Wake Forest University football team, the Demon Deacons, had one of the best seasons in school history in 2006, winning the Atlantic Coast Conference championship and earning a berth in the Orange Bowl. Not surprisingly, attendance at home games increased throughout the season as the number of victories kept climbing. Up until the 2007 season, Wake students did not have to purchase tickets to attend football games—they only needed to show a student ID to get into the football stadium. However, as interest in the football team rose, demand for seats at Wake home games started exceeding stadium capacity, and in the latter part of the season scores of students were turned away.

In response to this, the Wake Forest athletic department implemented a new mechanism for allocating football tickets to students for the 2007 season. Under the new system, a student may purchase Screamin' Demon season tickets for \$25, which guaranteed a seat for each home game. Students who forego this option may enter an online lottery to try to win free tickets. Under the lottery system, if students wish to go to a home game, they enter the lottery on the Sunday night before a game. The lottery operates on a first-come first-served basis. If students get their name into the drawing quickly enough, they are sent a confirmation e-mail that serves as an automatic ticket reservation.

Assuming that all students without season tickets who wish to go to a game sign up for the online lottery as soon as it is open, every Wake student is faced with a trade-off under the new ticket allocation system. On the one hand, by shelling out \$25, a student can guarantee a seat at every home game. On the other hand, a student can opt for the lottery system and try to attend games for free, although, with the lottery, entry to a game is subject to chance. Matters are complicated by the fact that the probability of getting a free ticket to a game under the lottery system depends on how many students choose to buy season tickets and how many choose the lottery, since these decisions affect the number of tickets available for distribution under the lottery system. Therefore, at the beginning of the football season when students have to choose which option to go with, they need to take other students' decisions into account.

To illustrate this point, consider a simple example: Suppose there are 10 students and the student section seats 6. If 4 students buy season tickets, then there are 2 tickets available for the lottery drawing for every home game. Assuming that all students without season tickets enter the lottery for every a home game, the probability that a student who enters the drawing gets a free ticket to a given home game is $2/6$. On the other hand, if only one student buys season tickets, then the probability of obtaining a free ticket to a given home game from the online lottery is $5/9$. Thus the actions of other students should factor into a student's thinking. The new ticket allocation system essentially places all students in a game-like situation in which students need to guess and anticipate what others will do.

Naturally, this raises the question of how students should behave under the new system, if they seek to maximize their individual well-being or net benefit. Should a student buy season tickets or not? Obviously, for students whose willingness-to-pay (the maximum amount of money one is willing to pay) for attending football games is less than the price of season tickets, the decision problem is trivial since they should not buy tickets and should opt for the online drawing if they desire going to games. But what about students whose willingness-to-pay exceeds the price of season tickets? What should they do?

Fortunately, we can apply to this problem the tools of game theory, a mathematical framework that has been developed to predict how rational actors with well-defined objectives will behave in game-like situations. (Readers may consult in-depth expositions of game theory [**1, 2, 3**] or elementary introductions [**4, 5**].) In what follows, we model the Screamin' Demon scenario using game theory to predict how the students will behave under the new ticket allocation system.

The model To keep things as simple as possible, assume that there is only one home game during the season and that everyone would prefer attending the game for free over not attending. [EDITOR'S NOTE: This means that students who would rather stay home reading the MAGAZINE are construed not to belong to the student body.] Before the season's start, every student must choose one of the following two options: buy a ticket for the home game at price $c > 0$ or enter an online lottery in hopes of winning a free ticket to the home game. Student preference for attending the game is characterized by the parameter $v > 0$, which gives a student's willingness to pay (WTP) to see the game. To allow for differences in how much students value going to the game, let $F(v)$ denote the fraction of students whose WTP is less than or equal to v, where $F : \mathbb{R}_+ \to [0, 1]$ is an increasing function. Let us assume that there is a large number of students and that $F(\cdot)$ is continuous, with $F(0) = 0$, $\lim_{v\to\infty} F(v) = 1$, and $F(v) > 0$ if $v > 0$. The mapping $F(\cdot)$ is the cumulative distribution function for students' WTP to see the game.

A student who would be willing to pay v but buys a ticket at price c gets a net benefit of $v - c$. The same student would receive a net benefit of v by being able to attend the game for free. The net benefit from not attending the game is 0. It is assumed throughout that students seek to maximize (the expected value of) their net benefit.

Suppose the student section of the football stadium can accommodate at most the fraction $T \in (0, 1)$ of students. Since we are assuming that everyone has a positive WTP to see the game, students who do not buy a ticket will choose to enter the online drawing. This means that if the fraction of students who buy a ticket in the student section is $x \leq T$, then the fraction $T - x$ of students would receive a free ticket through the online lottery. We assume that every student in the online lottery has an equal chance of winning a free ticket.

Since the student section cannot accommodate all students ($T < 1$), we must specify what would happen if the fraction of students who choose to buy a ticket x exceeds T (the problem would be trivial if $T \geq 1$, since everyone can get a free ticket by entering the online drawing if the number of tickets available exceeds the total number of students). In that case, let us assume that everyone who wishes to buy a ticket has probability T/x of being able to do so, leaving no tickets for distribution through the online lottery.

Given these assumptions, let us consider in detail the decision facing a student whose WTP to see the game is v. Since we assume that the number of students is large, the fraction of all students who choose to buy a ticket is the same whether this student buys a ticket or not. If $x > T$, meaning that there are more would-be buyers than there are seats, then our student's expected net benefit from buying a ticket is $T(v - c)/x$, while the expected net benefit from not buying is 0. Therefore the student's optimal action is to buy if $v > c$ and not to buy otherwise.

On the other hand, if seats are less in demand, with $x \leq T$, then the student's expected net benefit is $v - c$ from buying a ticket and $(T - x)v/(1 - x)$ from not buying. Thus, to maximize expected net benefit, the student should buy when

$$v - c > \left(\frac{T - x}{1 - x}\right) v \quad \text{or, after rearranging terms,} \quad v > \left(\frac{1 - x}{1 - T}\right) c$$

and not buy otherwise. Note that every student is faced with the same decision problem and that the fraction x is determined by the actions of all students.

How do we actually go about determining x, the fraction of all students who choose to buy a ticket? Notice that, given any $x \in [0, 1]$, our calculations imply that the fraction of students who choose to buy a ticket is

$$y(x) := \begin{cases} 1 - F\left(\left(\dfrac{1 - x}{1 - T}\right) c\right) & \text{if } x \in [0, T] \\ 1 - F(c) & \text{if } x \in (T, 1]. \end{cases} \tag{1}$$

Therefore, x must satisfy the condition $y(x) = x$, that is, the *equilibrium* fraction of students who choose to buy a ticket, which we denote henceforth by x^*, is a fixed point of the function $y(\cdot)$.

Let us consider some properties of this equilibrium fraction x^* and how this equilibrium fraction changes as we vary the price of a ticket c and the parameter T, which is proportional to the total number of student tickets available.

Existence of an equilibrium Given the continuity assumption on $F(\cdot)$, an equilibrium fraction of ticket buyers x^* always exists. In addition, x^* must be less than T if $1 - F(c) < T$. This claim follows since $y(x)$ is increasing so that $1 - F(c) < T$ implies that $y(x) < T$ for all $x \in [0, 1]$. Thus, the model tells us that if the cost of a ticket is high enough, then not all tickets will be sold and some students will be able to attend the game for free. On the other hand, notice from (1) that if $1 - F(c) > T$, then $1 - F(c)$ is an equilibrium fraction of students who choose to buy a ticket, that is, when c is sufficiently low, there is an equilibrium in which all students with WTP exceeding c choose to buy a ticket and all tickets are sold out. At this equilibrium, no student gets to go to a game for free.

An important prediction of the model is that, in general, there can be more than one equilibrium since a fixed point of the mapping $y(\cdot)$ may not be unique. For an example, suppose the cumulative distribution function $F(\cdot)$ is given as follows:

$$F(v) = \begin{cases} 2v^2 & \text{if } v \le \frac{1}{2} \\ -2v^2 + 4v - 1 & \text{if } \frac{1}{2} < v \le 1 \\ 1 & \text{if } v > 1. \end{cases} \tag{2}$$

FIGURE 1 suggests that few students have a very low or very high willingness to pay, with no one willing to pay more than a price of 1 on our scale.

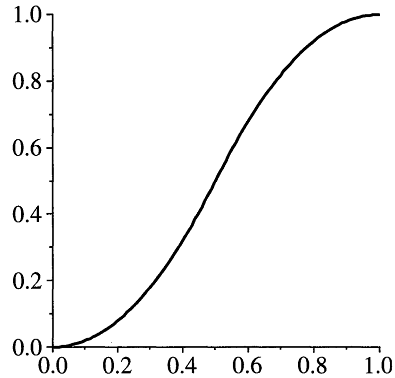


Figure 1 The cumulative distribution of students' willingness to pay

Given the cumulative distribution function (2) we plot as FIGURE 2 the mapping $y(\cdot)$, with parameter values $c = 0.45$ and $T = 0.55$. The function $y(\cdot)$ in this case is given by

$$y(x) = \begin{cases} 1 - \left(-2(1-x)^2 + 4(1-x) - 1\right) & \text{if } 0 \le x \le 0.5 \\ 1 - 2(1-x)^2 & \text{if } 0.5 < x \le 0.55 \\ 1 - 2(0.45)^2 & \text{if } x > 0.55. \end{cases}$$

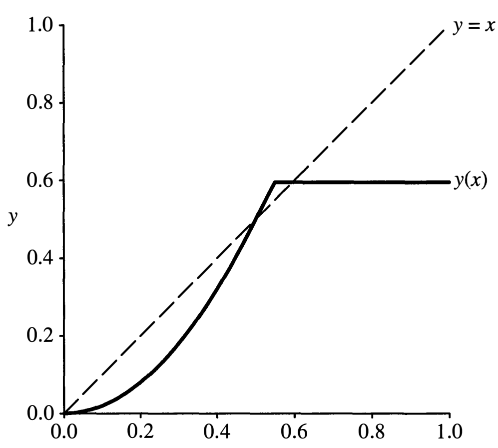


Figure 2 Using the cumulative distribution from (2) and parameter values $c = 0.45$, $T = 0.55$, a plot of the function $y(\cdot)$ shows three fixed points: 0, 0.5, and 0.595

As one can see from the figure, there are three equilibria in this example: a high-demand equilibrium in which all students whose WTP to see the game exceeds the price of the ticket choose to buy a ticket and the game is sold out; a medium-demand equilibrium in which half of the students buy a ticket and the game is not sold out so that some of the remaining students get a free ticket through the online lottery; and a low-demand equilibrium in which no student buys a ticket and all available tickets are distributed using the online drawing system.

The coexistence of multiple equilibria means that, in general, even if we know students' preferences and how much money they are willing to pay to see a game, we may not be able to predict with any degree of accuracy how many students will choose to buy a ticket and how many will opt for the online lottery instead. Why can there be multiple equilibria? What accounts for the fact that, for suitable parameter values and model specifications, there can be equilibria for which very few students buy a ticket as well as equilibria for which many students choose to buy a ticket and all tickets are sold? Intuitively, the multiplicity of equilibria arises because the probability of obtaining a free ticket through the online lottery is decreasing in the proportion of students who choose to buy a ticket. In other words, the higher the number of students who choose to purchase a ticket, the less likely one is able to obtain a free ticket through the online drawing. To see this, note that the probability of obtaining a free ticket for a student through the lottery is

$$g(x) := \begin{cases} \dfrac{T - x}{1 - x} & \text{if } x \leq T \\ 0 & \text{if } x > T, \end{cases}$$

where x is the fraction of students who choose to buy a ticket, and that $g(\cdot)$ is decreasing in x. This means that, before the season begins, if a student expects many other students to buy a ticket, then the student will have a greater incentive to buy a ticket as well since the likelihood of getting a free ticket though the lottery would be low in that case. Because all students face the same decision problem, the expectation that many students will choose to buy a ticket can be self-fulfilling and cause the number of ticket buyers to be high. On the other hand, if students believe that very few people will buy a ticket and that, as a consequence, the probability of winning a free ticket from the online drawing is high, then the number of students who choose to purchase a ticket will tend to be low. Note from this discussion that, when multiple equilibria coexist, which equilibrium the students will end up in depends on their expectations and beliefs concerning what other people will do: An equilibrium with a high number of buyers will obtain if students expect the number of buyers to be high; an equilibrium with a low number of buyers will obtain if students expect the number of buyers to be low.

Changing the price of tickets What is the effect of an increase in the ticket price c? At the outset, one might expect that higher ticket prices would cause the number of students who buy a ticket in equilibrium to decrease. However, as FIGURE 3 shows, that is not necessarily the case—a rise in the cost of tickets can actually lead to an *increase* in the number of ticket buyers. This counter-intuitive result can occur, however, only if multiple equilibria coexist. When an equilibrium is unique, increasing ticket price unambiguously leads to a lower proportion of students who choose to buy a ticket. To see this, suppose x^* is the unique equilibrium fraction of students who buy tickets before the price increase. This means that $y(x)$ must be less than x for all $x > x^*$. Increasing the price of tickets decreases the net benefit of buying tickets, all else being

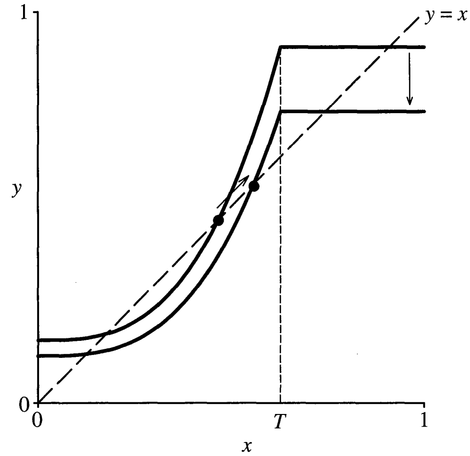


Figure 3 An increase in ticket price c shifts the function $y(\cdot)$ downwards and can cause the equilibrium number of buyers to rise

equal. This decreases $y(x)$ at all levels of x, so that the new equilibrium fraction of students who buy tickets cannot be higher than x^*.

Changing the number of tickets Due to the possible existence of multiple equilibria, the equilibrium fraction of students who choose to buy a ticket can increase, decrease, or stay the same in response to a change in T, as illustrated in FIGURE 4. However, just as in the case of changes in ticket price c, the effect of a change in T on the fraction of students who choose to buy a ticket in equilibrium is unambiguous when there is a unique equilibrium. In particular, if uniqueness is assured, then there is an inverse relationship between the number of tickets and the number of people who choose to buy tickets. To see this, suppose x^* is the unique equilibrium fraction of students who buy tickets before the change in number of tickets. This means that $y(x)$ is less than x for all $x > x^*$. Now, an increase in T improves the chance of winning a free ticket through the lottery, which decreases $y(x)$ at all levels of x. Therefore, the new equilibrium fraction of students who buy tickets cannot be greater than x^*.

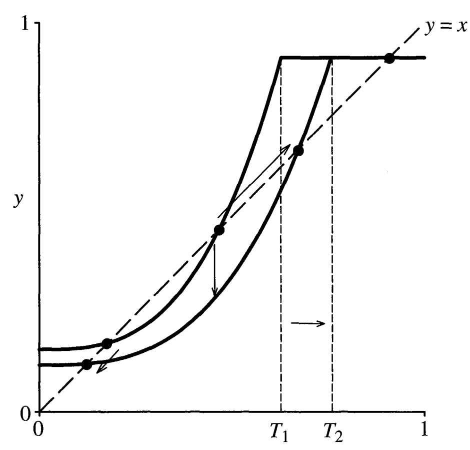


Figure 4 When multiple equilibria coexist, an increase in the number of tickets can cause the number of students who choose to buy a ticket to increase, decrease, or stay the same

Concluding remarks We have used game theory to predict the ticket-buying behavior of students faced with a lottery system that allows them to win free tickets. Such a lottery system can give rise to counter-intuitive effects. For example, a decrease in ticket prices can actually cause the number of ticket buyers to drop. Thus, when designing a mechanism to allocate tickets, it is vital to carefully consider buyers' incentives.

Acknowledgment. Comments and suggestions from a referee are gratefully acknowledged.

REFERENCES

1. D. Fudenberg and J. Tirole, *Game Theory*, MIT Press, Cambridge, MA, 1991
2. R. Myerson, *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, MA, 1997
3. M. Osborne and A. Rubinstein, *A Course in Game Theory*, MIT Press, Cambridge, MA, 1994
4. P. Dutta, *Strategy and Games: Theory and Practice*, MIT Press, Cambridge, MA, 1999
5. J. Watson, *Strategy: An Introduction to Game Theory*, W.W. Norton, New York, NY, 2001

What Do You Get When You Cross a Power Sum with an Iraqi Bank Note?

GREGORY MARK BOUDREAUX
University of North Carolina at Asheville
Asheville, NC 28804-8511
gboudrea@unca.edu

doi:10.4169/193009809X468715

I was completely captivated by my high school calculus course. It wove together all of the loose threads that had accumulated in all previous math courses, including the "summing of number sequences" thread introduced in my first algebra class. Now, as a college professor, I still know of no better way to justify the word "Fundamental" in the Fundamental Theorem of Calculus than to compute a definite integral of a polynomial using limits, which necessarily involves knowing the closed-forms of the power sums $1^k + 2^k + \cdots + n^k$. Back in high school, our calculus text gave the required formulas that induction was employed to verify, but this *a posteriori* verification only served to fuel my interest in the derivation of these mysterious, closed-form, power sum formulas.

Over the years, I have concocted several methods to derive these power sum formulas and I have studied many methods devised by others [**2, 3, 5, 6**]. These methods always seem to fall into one of two categories:

1. Intuitive and well-motivated, but not generalizable to all k;
2. Clever and often mystifying, yet generalizable.

Although there is nothing wrong with clever (and some of these clever methods are quite spectacular [**1**]), I would prefer to introduce my own students to these gems in an intuitive way that is also generalizable. This short note introduces a method I found in 2001 that satisfies these pedagogical criteria, but with a thousand year old twist involving a recently released Iraqi bank note.

Before deriving a formula for the general case $1^k + 2^k + \cdots + n^k$, I demonstrate the technique on some simpler cases, achieving familiar results.

When $k = 1$,

$$\sum_{i=1}^{n} i^k = 1 + 2 + \cdots + n = \left.\begin{array}{r} \overbrace{1 + 1 + \cdots + 1}^{n} \\ +1 + \cdots + 1 \\ \ddots \\ +1 \end{array}\right\} n$$

since the entries in the ith column in the addition array sum to i. On computing the sums of the n rows in the array, we find that

$$\begin{aligned} \sum_{i=1}^{n} i &= \sum_{i=1}^{n} 1 + \sum_{i=2}^{n} 1 + \sum_{i=3}^{n} 1 + \cdots + \sum_{i=n}^{n} 1 = \sum_{j=1}^{n}\sum_{i=j}^{n} 1 \\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} 1 - \sum_{i=1}^{j-1} 1\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} 1 + \left(1 - \sum_{i=1}^{j} 1\right)\right) \\ &= \sum_{j=1}^{n} (n + 1 - j) = n^2 + n - \sum_{i=1}^{n} i. \end{aligned}$$

Solving for $\sum_{i=1}^{n} i$, we obtain the expected formula. Note that the key to the this derivation is rewriting

$$\sum_{j=1}^{n}\sum_{i=j}^{n} 1 \quad \text{as} \quad \sum_{j=1}^{n}\left(\sum_{i=1}^{n} 1 - \sum_{i=1}^{j-1} 1\right) \qquad \text{and} \qquad -\sum_{i=1}^{j-1} 1 \quad \text{as} \quad 1 - \sum_{i=1}^{j} 1,$$

a maneuver we repeat throughout this note. Also note that only basic summation rules and simple algebra are required for this result, which may be applied to the $k = 2$ case as follows:

$$\sum_{i=1}^{n} i^2 = 1^2 + 2^2 + 3^2 + \cdots + n^2 = \left.\begin{array}{r} 1 + 2 + 3 + \cdots + n \\ + 2 + 3 + \cdots + n \\ + 3 + \cdots + n \\ \ddots \\ + n \end{array}\right\} n.$$

Since the entries in the ith column sum to $i \cdot i = i^2$,

$$\begin{aligned} \sum_{i=1}^{n} i^2 &= \sum_{j=1}^{n}\sum_{i=j}^{n} i = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} i + \left(j - \sum_{i=1}^{j} i\right)\right) \\ &= \sum_{j=1}^{n}\left(\frac{n(n+1)}{2} + j - \frac{j(j+1)}{2}\right) \\ &= \frac{1}{2}\sum_{j=1}^{n}\left(n^2 + n + j - j^2\right) = \frac{1}{2}\left(n^3 + n^2 + \frac{n(n+1)}{2} - \sum_{i=1}^{n} i^2\right). \end{aligned}$$

Thus,

$$\frac{3}{2}\sum_{i=1}^{n} i^2 = \frac{2n^3 + 2n^2 + n^2 + n}{4},$$

yielding the familiar formula

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$$

In general, when $k \geq 1$,

$$\sum_{i=1}^{n} i^k = 1^k + 2^k + 3^k + \cdots + n^k = \left.\begin{array}{r} 1 + 2^{k-1} + 3^{k-1} + \cdots + n^{k-1} \\ + 2^{k-1} + 3^{k-1} + \cdots + n^{k-1} \\ + 3^{k-1} + \cdots + n^{k-1} \\ \ddots \\ + n^{k-1} \end{array}\right\} n.$$

Since the entries in the ith column sum to $i \cdot i^{k-1} = i^k$,

$$\sum_{i=1}^{n} i^k = \sum_{j=1}^{n} \sum_{i=j}^{n} i^{k-1} = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} i^{k-1} + \left(j^{k-1} - \sum_{i=1}^{j} i^{k-1} \right) \right)$$

$$= n \sum_{i=1}^{n} i^{k-1} + \sum_{j=1}^{n} j^{k-1} - \sum_{j=1}^{n} \sum_{i=1}^{j} i^{k-1}.$$

Therefore, for $k \geq 1$,

$$\sum_{i=1}^{n} i^k = (n+1) \sum_{i=1}^{n} i^{k-1} - \sum_{j=1}^{n} \sum_{i=1}^{j} i^{k-1}, \tag{1}$$

which is a reduction-type formula for the general case. When paired with your favorite computer algebra system, this formula can be used to generate, rather painlessly, the closed form power sum formulas for any value of k. It is also arguably the easiest path to a general formula for $1^k + 2^k + \cdots + n^k$.

It is time to reveal the riddle posed in the title of this article: This involves the Muslim scientist Abū ‘Alī al-Hasan ibn al-Hasan ibn al-Haytham, known in Europe as Alhazen. He was born in 965 AD in Basra, now in Iraq, and appears on the new Iraqi 10,000-dinar note, launched in October 2003, honoring his reputation as one of the most influential of Islamic scientists. (FIGURE 1 shows the note, which was recently featured in the April 2006 issue of *National Geographic*.) His connection to power sums provides the twist to this tale, which was uncovered during my literature review to search for previously published general formulas for $1^k + 2^k + \cdots + n^k$. In the "Mathematics of Islam" section of Katz' history of mathematics text [**4**, p. 240] appears the formula

$$(n+1) \sum_{i=1}^{n} i^k = \sum_{i=1}^{n} i^{k+1} + \sum_{p=1}^{n} \sum_{i=1}^{p} i^k, \tag{2}$$

which is obviously (1) with the slightest twist.

Alas. It seems this gem had been unearthed almost a millennium before I was born. But aside from a brief mention in a math history or two, it appears never to be used in modern times for the purpose ibn al-Haytham had intended and for which it excels: to generate closed forms for power sums. (Apparently, he needed it to generate the power sum formulas only for the second and fourth powers for his computation of the volume of a paraboloid, according to Katz.)

Figure 1 Ibn al-Haytham on the 10,000-dinar Iraqi bank note

I greatly enjoy reading about the history of mathematics. Some results seem destined never to be obscured by time, like the Pythagorean Theorem. Others are superseded by more general or more elegant discoveries made possible by our deepening understanding over time, like Apollonius' treatment of constructing tangents to conics. But there is a third type of result—one that simply needs to be dusted off and polished a bit to reveal its inherent brilliance. This type of result should be rescued from obscurity, for it has a rightful place in the mathematics classroom. I contend that ibn al-Haytham's formula is of this last type. Katz dusted off this result by presenting it using modern notation and stating it in its general form in his book, although he points out that ibn al-Haytham only stated (2) for the particular integers $n = 3$ and $k = 1, 2, 3$. Katz [**4**, p. 240] then discloses ibn al-Haytham's original method of deriving the special cases of (2), which differs significantly from what has been presented here. Namely, for the case $n = 3$ and $k = 3$, ibn al-Haytham reasons (more or less) as follows:

$$\begin{aligned} 2 \cdot 1^3 &= 1^4 + 1^3 \\ 3\left(1^3 + 2^3\right) &= (2+1)\left(1^3 + 2^3\right) = 2 \cdot 1^3 + 2^4 + 1^3 + 2^3 \\ &= 1^4 + 1^3 + 2^4 + 1^3 + 2^3 = \left(1^4 + 2^4\right) + 1^3 + \left(1^3 + 2^3\right) \\ 4\left(1^3 + 2^3 + 3^3\right) &= (3+1)\left(1^3 + 2^3 + 3^3\right) = 3\left(1^3 + 2^3\right) + 3^4 + 1^3 + 2^3 + 3^3 \\ &= \left(1^4 + 2^4 + 3^4\right) + 1^3 + \left(1^3 + 2^3\right) + \left(1^3 + 2^3 + 3^3\right). \end{aligned}$$

The first step is true. Each subsequent step follows from the previous one. Induction on n is then needed to rigorously prove (2), as Katz suggests, for the $k = 3$ case. However, it is not immediately apparent that the steps presented still remain valid for other values of k. And it is not clear how ibn al-Haytham discovered this amazing pattern, encapsulated in (2), in the first place.

The first part of this note, I hope, polishes out the roughness of ibn al-Haytham's original treatment, providing a well-motivated, intuitive, and generalizable approach to these particular power sums. I hope that the second part, the unexpected twist, reveals that history has much more to offer the modern mathematics classroom, both content-wise and pedagogically speaking, than just a good story.

REFERENCES

1. A. B. Ayoub, A note on the sums of squares of natural numbers, *Math. Comput. Ed.* **26** (1992) 246–247.
2. A. W. F. Edwards, Sums of powers of integers: a little history, *Math. Gaz.* **66** (1982) 22–28. doi:10.2307/3617302

3. A. W. F. Edwards, A quick route to sums of powers, *Amer. Math. Monthly* **93** (1986) 451–455. doi:10.2307/2323466
4. V. J. Katz, *A History of Mathematics: An Introduction*, HarperCollins College Publishers, New York, 1993.
5. H. K. Krishnapriyan, Eulerian polynomials and Faulhaber's results on sums of powers of integers, *College Math. J.* **26** (1995) 118–123.
6. B. Turner, Sums of powers of integers via the binomial theorem, this MAGAZINE **53** (1980) 92–96.

Math Bite: $\mathbb{Q}$ Is Not Complete

JOSÉ VILLA MORALES
Departamento de Matemáticas y Física
Universidad Autónoma de Aguascalientes
Aguascalientes, Ags., 20130 MEXICO
jvilla@correo.uaa.mx

doi:10.4169/193009809X468751

Define the sequence $\{x_n\}$ in $\mathbb{Q}$ by

$$x_n = \begin{cases} \dfrac{1}{10^n}, & \text{if } n = k^2 \text{ for } k \in \mathbb{N}, \\ -\dfrac{1}{10^n}, & \text{otherwise.} \end{cases}$$

Note that the series $\sum_{n=1}^{\infty} x_n$ converges absolutely:

$$\sum_{n=1}^{\infty} |x_n| = \sum_{n=1}^{\infty} \frac{1}{10^n} = \frac{1}{1-\frac{1}{10}} - 1 = \frac{1}{9} \in \mathbb{Q}. \tag{1}$$

Consider the partial sums of $\{x_n\}$, $s_n = \sum_{k=1}^{n} x_k$. For each $n, p \in \mathbb{N}$ we have

$$|s_{n+p} - s_n| \leq |x_{n+1}| + \cdots + |x_{n+p}|,$$

from which we can see that $\{s_n\}$ is a Cauchy sequence of rational numbers. Since $\sum_{n=1}^{\infty} x_n$ converges absolutely, its sum does not depend on the way the terms of the series are grouped. Therefore

$$\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty} \left(\sum_{i=n^2}^{(n+1)^2-1} x_i \right).$$

Some algebraic magic, combining terms in geometric progression, tells us that

$$\begin{aligned}
\sum_{i=n^2}^{(n+1)^2-1} x_i &= \frac{1}{10^{n^2}} - \frac{1}{10^{n^2+1}} - \cdots - \frac{1}{10^{(n+1)^2-1}} \\
&= \frac{10^{(n+1)^2-1-n^2} - (10^{(n+1)^2-1-n^2-1} + \cdots + 10 + 1)}{10^{(n+1)^2-1}} \\
&= \frac{10^{(n+1)^2-1-n^2} - \left(\frac{10^{(n+1)^2-1-n^2}-1}{9} \right)}{10^{(n+1)^2-1}} = \frac{\frac{8\times 10^{(n+1)^2-1-n^2}+1}{9}}{10^{(n+1)^2-1}} \\
&= \frac{\frac{8(10^{(n+1)^2-1-n^2}-1)+9}{9}}{10^{(n+1)^2-1}} = \frac{8\left(\frac{10^{(n+1)^2-1-n^2}-1}{9} \right) + 1}{10^{(n+1)^2-1}}
\end{aligned}$$

$$= \frac{8 \times 10^{(n+1)^2-1-n^2-1} + 8 \times 10^{(n+1)^2-1-n^2-2} + \cdots + 8 \times 10 + 9}{10^{(n+1)^2-1}}.$$

Hence

$$\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty} \left(\frac{0}{10^{n^2}} + \frac{8}{10^{n^2+1}} + \frac{8}{10^{n^2+2}} + \cdots + \frac{8}{10^{(n+1)^2-2}} + \frac{9}{10^{(n+1)^2-1}} \right)$$
$$= 0.0890888908888890888888889.... \qquad (2)$$

Since a real number is rational if and only if its decimal expansion is periodic **[1]**, we see that $\lim_{n\to\infty} s_n = \sum_{n=1}^{\infty} x_n \notin \mathbb{Q}$. We have a Cauchy sequence in $\mathbb{Q}$ that converges to a number not in $\mathbb{Q}$, and this means that $\mathbb{Q}$ is not complete.

Also it is interesting to note that this series of rational numbers converge absolutely to a rational number (1), but it converges to an irrational number (2).

Acknowledgment. The author is grateful to an anonymous referee for suggestions. He was partially supported by grant PIM 08-2 of UAA.

REFERENCES

1. R. G. Bartle and D. R. Sherbert, *Introduction to Real Analysis*, Wiley, 1999.

A Low-Level Proof of Chebyshev's Pre-Prime Number Theorem

M. SCOTT OSBORNE
Department of Mathematics, Box 354350
University of Washington
Seattle, WA 98195–4350
sosborne@math.washington.edu

doi:10.4169/193009809X468706

The Prime Number Theorem (PNT) is a favorite of many mathematicians, for several reasons. At the top of the list is probably the nature of the proof: Heavy Complex Analysis yields Combinatorial Information about an Algebraic Object. The complex analysis is applied to the Riemann zeta function (and its Euler product), and is fairly intricate. Furthermore, there are typically several number theoretic reductions (involving several associated number theoretic functions) along the road.

The impression all this gives is that you can't deduce much from the Euler product for Riemann's zeta function without using complex analysis, and you can't get close to the PNT without using a lot of auxiliary prime-oriented functions. It therefore came as a surprise to me that a "proto" version of the PNT could be derived using little more than sophomore-level real analysis of infinite series and one auxiliary prime function. The purpose of this paper is to present this approach. It is not particularly complicated. (In 1948, Selberg **[10]** gave a calculus-level derivation of the PNT itself. His proof is *not* simple!)

PNT or bust The Prime Number Theorem was first proved by Hadamard and de la Vallée Poussin (independently) in 1896. However, the PNT was conjectured by Legendre long before and, in two papers in the early 1850s, Chebyshev **[2, 3]** essentially showed that the PNT as we know it was the only game in town.

As usual, let $\pi(x)$ denote the number of primes less than or equal to x. The prime number theorem states that $\pi(x)$ is asymptotic to $x/\ln x$. In Chebyshev's first paper, he showed (among other things) that

$$\text{If } \pi(x)\Big/\frac{x}{\ln x} \text{ approaches a limit } L, \text{ then } L = 1. \tag{1}$$

In the second paper, Chebyshev established some fairly tight estimates of the form

$$a \cdot \frac{x}{\ln x} \le \pi(x) \le b\frac{x}{\ln x} \text{ for large } x \tag{2}$$

with $a \approx .92$ and $b \approx 1.105$. Both arguments were fairly intricate. The second result strongly favors examining the ratio between $\pi(x)$ and $x/\ln x$, while the first tells us what the PNT should say.

These days, the inequalities (2) are done in two possible ways. Elementary methods yield the inequalities with fairly coarse constants; typically, $a = \frac{1}{4}\ln 2$, while b ranges from $9\ln 2$ to $32\ln 2$. The argument is purely number theoretic, based on the behavior of the binomial coefficient $\binom{2n}{n}$, with nary an integral in sight. Andrews [**1**], Dudley [**5**], and Niven and Zuckerman [**7**] use this approach. Robertson and Staton [**9**] recently gave an elegant argument reducing b to $3\ln 2$. Substantially more complicated arguments appear in LeVeque [**6**] and in Niven, Zuckerman, and Montgomery [**8**]; after substantial reductions (and a fair amount of real analysis), constants closer to Chebyshev's are obtained. LeVeque's presentation eventually leads to (1) as well.

The purpose here is to establish (1) fairly simply, using one auxiliary prime function and real analysis of the Euler product for the Riemann zeta function.

Listing primes The number theoretic reduction involves the kth *prime function*, p_k, which simply denotes the kth prime. In particular, the PNT is equivalent to the statement that p_k is asymptotic to $k\ln k$. This form is often used to explain the PNT to nonmathematicians. The proof is based on part of inequality (2), with any constant $a > 0$ at all. As this version of (2) is considered "elementary," it will be taken as a given. Since $\pi(p_k) = k$,

$$\pi(p_k)\Big/\frac{p_k}{\ln p_k} = \frac{k\ln p_k}{p_k} = \frac{k\ln k}{p_k}\cdot\frac{\ln p_k}{\ln k}.$$

We can see that $\ln p_k/\ln k$ approaches 1 as $k \to \infty$ by noting that $p_k \ge k$ and using (2) and the squeezing principle:

$$\begin{aligned} 1 \ge \frac{\ln k}{\ln p_k} &= \frac{\ln \pi(p_k)}{\ln p_k} \ge \frac{\ln(ap_k/\ln p_k)}{\ln p_k} \\ &= \frac{\ln a + \ln p_k - \ln\ln p_k}{\ln p_k} \to 1. \end{aligned}$$

It follows immediately that $k\ln k/p_k$ and $\pi(x)\big/\frac{x}{\ln x}$ have the same limit, as long as x is restricted to primes. However, for real x, with $p_k \le x < p_{k+1}$ (and $k \ge 2$, so that $p_k > e$):

$$\begin{aligned} \pi(p_k)\Big/\frac{p_k}{\ln p_k} &\ge \pi(x)\Big/\frac{x}{\ln x} > (\pi(p_{k+1}) - 1)\Big/\frac{p_{k+1}}{\ln p_{k+1}} \\ &= \left[\pi(p_{k+1})\Big/\frac{p_{k+1}}{\ln p_{k+1}}\right] - \frac{\ln p_{k+1}}{p_{k+1}} \end{aligned}$$

since $x/\ln x$ is increasing for $x \geq e$. That is, the quotient $\pi(x)\Big/\frac{x}{\ln x}$ is controlled by its values at the primes.

We move on to the analytic result that yields Chebyshev's first result (1). Roughly speaking, it says that p_k and $k \ln k$ have a similar "harmonic average," and qualifies as a "pre-Prime Number Theorem."

THEOREM 1. *Although both series tend to ∞ as $s \to 1^+$, the difference*

$$\sum_{k=1}^{\infty} \frac{1}{p_k^s} - \sum_{k=2}^{\infty} \frac{1}{(k \ln k)^s}$$

stays bounded as $s \to 1^+$. Specifically, the absolute value of the difference is less than 6 for $1 < s \leq 2$.

To see quickly how this implies Chebyshev's result (1), suppose for a moment that $\pi(x)\Big/\frac{x}{\ln x} \to L < 1$. Then $k \ln k/p_k \to L$, too. Pick some α with $L < \alpha < 1$. Then beyond some point (say for $k \geq N$), $k \ln k/p_k < \alpha$, giving $p_k^{-1} < \alpha(k \ln k)^{-1}$. So:

$$\sum_{k=N}^{\infty} \frac{1}{(k \ln k)^s} - \sum_{k=N}^{\infty} \frac{1}{p_k^s} \geq (1 - \alpha^s) \sum_{k=N}^{\infty} \frac{1}{(k \ln k)^s}$$

and the right-hand side goes off to $+\infty$ as $s \to 1^+$. The finite number of terms left off the series in Theorem 1 can't prevent divergence, and this contradicts Theorem 1. The proof that $L > 1$ is untenable is similar.

Riemann's zeta function The proof of Theorem 1 is based on the Euler product for the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \; s > 1.$$

The Euler product is given by

$$\zeta(s) = \prod_{k=1}^{\infty} (1 - p_k^{-s})^{-1}, \; s > 1,$$

and taking logarithms produces

$$\ln \zeta(s) = \sum_{k=1}^{\infty} - \ln(1 - p_k^{-s}), \; s > 1.$$

A derivation of the Euler product formula can be found in many places, including Derbyshire's expository account [**4**], so we take it as a given here. That's all that is needed: no uniform convergence, no complex s, nothing else. (By the way, Derbyshire's description of Euler's proof is spectacularly transparent, and is highly recommended. Derbyshire calls it the "Golden Key," an apt name.)

The four inequalities The proof of Theorem 1 comes from summing four inequalities, all true for $1 < s \leq 2$:

$$0 \leq \sum_{k=1}^{\infty} \frac{1}{p_k^s} - \ln \zeta(s) \leq \frac{\pi^2}{3}, \tag{i}$$

$$-\ln 2 \leq \ln \zeta(s) + \ln(s - 1) \leq \ln 2, \tag{ii}$$

$$-e^{-1} \leq -\ln(s-1) - \int_e^\infty \frac{dx}{(x\ln x)^s} < 2, \text{ and} \tag{iii}$$

$$-1.1 < \int_e^\infty \frac{dx}{(x\ln x)^s} - \sum_{k=2}^\infty \frac{1}{(k\ln k)^s} \leq 0. \tag{iv}$$

Note that (i) and (ii) establish that $\sum p_k^{-s}$ is asymptotic to $-\ln(s-1)$ as $s \to 1^+$, so the two series in Theorem 1 do "blow up" as $s \to 1^+$. All that's left is to establish these inequalities.

Inequality (i) comes from Taylor's theorem with remainder. For $x < 1$, there is an η between 0 and x such that

$$-\ln(1-x) = x + \frac{1}{2(1-\eta)^2}x^2.$$

For $0 \leq x \leq 1/2$, we have $1-\eta > 1/2$, which means $2(1-\eta)^2 > 1/2$, so that

$$x \leq -\ln(1-x) \leq x + 2x^2.$$

Setting $x = p_k^{-s} \leq 1/2$:

$$\sum_{k=1}^\infty \frac{1}{p_k^s} \leq \sum_{k=1}^\infty -\ln(1-p_k^{-s}) \leq \sum_{k=1}^\infty \frac{1}{p_k^s} + 2\sum_{k=1}^\infty \frac{1}{p_k^{2s}}.$$

But $\sum p_k^{-2s} \leq \sum k^{-2} = \pi^2/6$, giving (i), thanks to Euler.

The proofs for (ii) and (iv) come from the integral test for infinite series. For (ii),

$$\int_1^\infty \frac{dt}{t^s} \geq \sum_{n=2}^\infty \frac{1}{n^s} \geq \int_2^\infty \frac{dt}{t^s} \text{ becomes } \frac{1}{s-1} \geq \zeta(s) - 1 \geq \frac{2^{1-s}}{s-1},$$

producing

$$\frac{s}{s-1} = \frac{1}{s-1} + \frac{s-1}{s-1} \geq \zeta(s) \geq \frac{2^{1-s}}{s-1} + \frac{s-1}{s-1}.$$

Taking logarithms,

$$\ln s - \ln(s-1) \geq \ln\zeta(s) \geq \ln(2^{1-s}+s-1) - \ln(s-1),$$

which gives

$$\ln s \geq \ln\zeta(s) + \ln(s-1) \geq \ln(2^{1-s}+s-1).$$

But $\ln(2^{1-s}+s-1) \geq \ln 2^{1-s} = (1-s)\ln 2 \geq -\ln 2$ for $1 \leq s \leq 2$. Similarly, for (iv),

$$\sum_{k=2}^\infty \frac{1}{(k\ln k)^s} \geq \int_2^\infty \frac{dx}{(x\ln x)^s} \geq \int_e^\infty \frac{dx}{(x\ln x)^s} \geq \sum_{k=4}^\infty \frac{1}{(k\ln k)^s}$$

so that

$$-\frac{1}{(2\ln 2)^s} - \frac{1}{(3\ln 3)^s} \leq \int_e^\infty \frac{dx}{(x\ln x)^s} - \sum_{k=2}^\infty \frac{1}{(k\ln k)^s} \leq 0.$$

But for $s > 1$,

$$\frac{1}{(2\ln 2)^s} + \frac{1}{(3\ln 3)^s} < \frac{1}{2\ln 2} + \frac{1}{3\ln 3} = 1.02476\ldots.$$

All that remains is (iii), which is the messiest.

First, change the sign in (iii), and make two changes of variables in the integral, first $u = \ln x$ and then $t = (s-1)u$:

$$\int_e^\infty \frac{dx}{(x\ln x)^s} = \int_1^\infty \frac{e^u du}{(ue^u)^s} = \int_1^\infty u^{-s}e^{(1-s)u}du$$
$$= \int_{s-1}^\infty \left(\frac{t}{s-1}\right)^{-s} e^{-t}\frac{dt}{s-1} = (s-1)^{s-1}\int_{s-1}^\infty t^{-s}e^{-t}dt.$$

(The simplicity of the lower limit now provides the rationale for using e as the lower limit originally.) Now $\ln(s-1)$ is the integral of t^{-1} from 1 to $s-1$, so

$$\int_e^\infty \frac{dx}{(x\ln x)^s} + \ln(s-1) = \int_{s-1}^1 [(s-1)^{s-1}t^{-s}e^{-t} - t^{-1}]\,dt$$
$$+ \int_1^\infty (s-1)^{s-1}t^{-s}e^{-t}\,dt.$$

However,

$$0 \le \int_1^\infty (s-1)^{s-1}t^{-s}e^{-t}dt \le \int_1^\infty e^{-t}\,dt = e^{-1},$$

so it suffices to show that

$$0 \ge \int_{s-1}^1 ((s-1)^{s-1}t^{-s}e^{-t} - t^{-1})\,dt > -2.$$

This is done by bounding the integrand, showing that, for $s-1 \le t \le 1$ and $s \le 2$,

$$0 \le t^{-1} - (s-1)^{s-1}t^{-s}e^{-t} < 2:$$

$$t^{-1} - (s-1)^{s-1}t^{-s}e^{-t} = \frac{1-\left(\frac{s-1}{t}\right)^{s-1}e^{-t}}{t} = \frac{1-e^{-t}}{t} + e^{-t}\frac{1-\left(\frac{s-1}{t}\right)^{s-1}}{t}$$
$$= \frac{1}{t}\left[-e^{-y}\right]_{y=0}^{y=t} + \frac{e^{-t}}{t}\left[-\left(\frac{s-1}{y}\right)^{s-1}\right]_{y=s-1}^{y=t}$$
$$= \frac{1}{t}\int_0^t e^{-y}dy + \frac{e^{-t}}{t}\int_{s-1}^t \left(\frac{s-1}{y}\right)^s dy$$
$$\le \frac{1}{t}\cdot\int_0^t dy + \frac{e^{-t}}{t}\int_{s-1}^t dy = 1 + e^{-t}\frac{t-(s-1)}{t} < 2.$$

Final notes Chebyshev's result (1) appears, for example, in LeVeque [**6**] and in Shapiro [**11**]. Their proofs do not resemble this one (nor do they resemble each other). Chebyshev's original approach, however, does bear a vague resemblance to the one here. He established the existence of

$$\lim_{s\to 1^+}\left(\sum_{k=1}^\infty \frac{(\ln p_k)^n}{p_k^s} - \sum_{k=2}^\infty \frac{(\ln k)^{n-1}}{k^s}\right)$$

regardless of n. (Chebyshev grouped his terms differently, but these were his series. Note that the $n = 0$ case is similar to Theorem 1.) The argument is longer but more definitive, leading to the conclusion that the logarithmic integral

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\ln t}$$

approximates $\pi(x)$ better than $x/\ln x$.

This kind of thing is possible because a result like Theorem 1 is quite a bit stronger than result (1), justifying the name "Pre-Prime Number Theorem." For example, it follows easily from Theorem 1 that

$$\lim_{s\to 1^+} \frac{\sum_{k=1}^{\infty} p_k^{-s}}{\sum_{k=2}^{\infty}(k \ln k)^{-s}} = 1,$$

and this weaker result is enough to establish (1), thanks to the following:

THEOREM 2. *Suppose* $a_k > 0$ *and* $b_k > 0$, *and* $\sum a_k^{-s}$ *and* $\sum b_k^{-s}$ *converge for* $s > 1$, *with* $\lim_{s\to 1^+} \sum a_k^{-s} = \lim_{s\to 1^+} \sum b_k^{-s} = \infty$. *Then*

$$\lim_{k\to\infty} \frac{a_k}{b_k} = L > 0 \Rightarrow \lim_{s\to 1^+} \frac{\sum b_k^{-s}}{\sum a_k^{-s}} = L.$$

This is a fairly straightforward, though complicated, $\epsilon/3$ argument: Given $\epsilon > 0$, replace b_k with

$$\widetilde{b}_k = \begin{cases} b_k & \text{if } |a_k/b_k - L| < \epsilon/3 \\ a_k/L & \text{if not.} \end{cases}$$

Note that $\left(L - \frac{\epsilon}{3}\right)\widetilde{b}_k < a_k < \left(L + \frac{\epsilon}{3}\right)\widetilde{b}_k$ for all k, so

$$\left(L - \frac{\epsilon}{3}\right)^{-s} \sum \widetilde{b}_k^{-s} > \sum a_k^{-s} > \left(L + \frac{\epsilon}{3}\right)^{-s} \sum \widetilde{b}_k^{-s}, \text{ that is,}$$

$$\left(L - \frac{\epsilon}{3}\right)^{s} < \frac{\sum \widetilde{b}_k^{-s}}{\sum a_k^{-s}} < \left(L + \frac{\epsilon}{3}\right)^{s}.$$

But for s close to 1, $\left(L + \frac{\epsilon}{3}\right)^s < L + \frac{2\epsilon}{3}$, $\left(L - \frac{\epsilon}{3}\right)^s > \left(L - \frac{2\epsilon}{3}\right)$, and

$$\left|\frac{\sum \widetilde{b}_k^{-s}}{\sum a_k^{-s}} - \frac{\sum b_k^{-s}}{\sum a_k^{-s}}\right| < \frac{\epsilon}{3},$$

since $\widetilde{b}_k = b_k$ except for finitely many k. (By the way, the theorem is also true when $L = 0$: Replace b_k with $4a_k/\epsilon$ when $|a_k/b_k| \geq \epsilon/3$ and, with suitable modifications, the argument goes through.)

One last thing, for those familiar with the concepts. Result (1) can be strengthened to

$$\liminf \pi(x)\Big/\frac{x}{\ln x} \leq 1 \leq \limsup \pi(x)\Big/\frac{x}{\ln x}$$

with only minor changes above.

REFERENCES

1. George Andrews, *Number Theory*, Philadelphia, W. B. Saunders, 1971.
2. P. L. Chebyshev, *Sur la Fonction qui Détermine la Totalité des Nombres Premiuers Inférieurs à une limite Donnée, Journal de Mathématiques Pures et Appliquées*, I serie, XVII, 1852, 341–365.
3. P.L. Chebyshev, *Memoire sur les Nombres Premiers, Journal de Mathématiques Pures et Appliques*, I serie, XVII, 1852, 366–390.
4. John Derbyshire, *Prime Obsession*, London, Penguin Books, 2003.
5. Underwood Dudley, *Elementary Number Theory*, 2nd ed., San Francisco, W. H. Freeman, 1978.
6. William LeVeque, *Fundamentals of Number Theory*, Reading, MA, Addison-Wesley, 1977.
7. Ivan Niven and Herbert Zuckerman, *An Introduction to the Theory of Numbers*, New York, John Wiley & Sons, 1960.
8. Ivan Niven, Herbert Zuckerman and Hugh Montgomery, *An Introduction to the Theory of Numbers*, New York, John Wiley & Sons, 1991.
9. Kimberly Robertson and William Staton, A Short Proof of Chebychev's Upper Bound, this MAGAZINE **78** (2005) 385–387.
10. Atle Selberg, An Elementary Proof of the Prime-number Theorem, *Ann. of Math.* **50**, (1949) 305–313. doi: 10.2307/1969455
11. Harold Shapiro, *Introduction to the Theory of Numbers*, New York, John Wiley & Sons, 1983.

Note: [**2**] and [**3**] appear in P. L. Chebyshev, *Oeuvres* Tome 1 (collected works), New York, Chelsea, 1962.

Tile in a Corner

RICHARD P. JERRARD
University of Illinois at Urbana-Champaign
Urbana, IL 61801
jerrard@illinois.edu

JOHN E. WETZEL
University of Illinois at Urbana-Champaign
Urbana, IL 61801
j-wetzel@illinois.edu

doi:10.4169/193009809X471081

In this note we consider the following somewhat awkward locus problem in elementary geometry:

TILE IN A CORNER. Find the locus of the center of a rectangular tile with given sides s and t that moves in the first octant with all four of its corners on coordinate planes, at least one corner on each.

This problem is closely related to the well-known "penny in a corner" problem from the 1948 William Lowell Putnam Mathematical Competition (Mackey [**9**]): *Find the locus of the center of a disk of given radius r that moves in the first octant staying tangent to each of the three coordinate planes.* This locus problem played a significant role in our recent investigation of the longest right circular cylinder of given small radius that fits diagonally in a given $a \times b \times c$ box (see Jerrard et al. [**8**]).

In 1923 Garnett [**4**] asked for the longest "plank" of given rectangular cross section that can fit in a given box. Two years later Carver [**2**] supplied a solution based on the reasonable but unproven assumption that "When the longest possible timber of given cross section is placed in the box, all of its vertices must lie in faces of the box," an assumption that he took "as axiomatic." We have not been able to prove this claim, and our hope is that the solution of the "tile in a corner" problem might provide some insight into this somewhat unsatisfactory situation, just as the solution of the "penny in a corner" aided in the solution of the cylinder problem.

A few preliminaries If the four corners of a tile $\mathbb{T}$ in the first octant all lie on the coordinate planes, with at least one corner on each plane, then one of the coordinate planes contains two of the corners. Label the corners of the tile A, B, Q, P in cyclic order, with $AB = PQ = s$ and $AP = BQ = t$ (FIGURE 1). When $s = t$, the tile is square, and writing, for example, $\mathcal{S}_z(s, s)$ for the surface that is the locus of its center C when one side lies in the plane $z = 0$, we see that the locus $\mathcal{S}(s)$ we seek is the union of three such surfaces:

$$\mathcal{S}(s) = \mathcal{S}_x(s, s) \cup \mathcal{S}_y(s, s) \cup \mathcal{S}_z(s, s).$$

When $s \neq t$, the tile is rectangular, and the locus of its center when the longer side lies in a coordinate plane is different from that traced when the shorter side lies in the same coordinate plane. Write, for example, $\mathcal{S}_z(s, t)$ for the locus of the center C of the tile when a side of the first-named length s lies in the plane $z = 0$. Then the locus $\mathcal{S}(s, t)$ we seek is the union of six such surfaces:

$$\mathcal{S}(s, t) = (\mathcal{S}_x(s, t) \cup \mathcal{S}_x(t, s)) \cup (\mathcal{S}_y(s, t) \cup \mathcal{S}_y(t, s)) \cup (\mathcal{S}_z(s, t) \cup \mathcal{S}_z(t, s)).$$

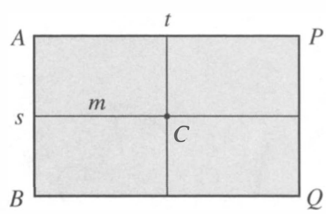


Figure 1 Tile $\mathbb{T} = ABQP$

Our intention is first to study the surface $\mathcal{S}_z(s, t)$ that is the locus of C when the two corners A and B of $\mathbb{T}$ lie in the plane $z = 0$, without regard to the relation between s and t. As every calculus student knows, one useful way to study an unfamiliar surface is to investigate its cross sections. Although the cross sections of $\mathcal{S}_z(s, t)$ in planes parallel to the xz-plane and to the yz-plane are higher-order curves and not very helpful, the cross sections in the horizontal planes turn out to be elliptic arcs! And as we shall see, certain other sections turn out to be quarter circles. From this information we shall be able to gain a reasonably good picture of the surface.

Some familiar loci We begin rather far afield with a familiar elementary locus problem. Suppose a line segment AB of length s moves with its endpoints on the coordinate axes, A on the x-axis and B on the y-axis. Let M be the midpoint of AB, and suppose that C is the point on the perpendicular bisector of AB at (signed) distance d from M. What is the locus of C?

Finding the locus of the midpoint of the segment AB (the case $C = M$) is a familiar exercise in analytic geometry; it is easy to see that the locus is the circle centered at the origin with radius $s/2$. For $C \neq M$ the locus isn't so familiar, although some experimentation, say with a program like *Geometer's Sketchpad* or *Cinderella*, certainly suggests that the locus is an oval of some kind, perhaps an ellipse. In fact, deriving an equation for the locus is quite easy.

Write $A = (u, 0)$ and $B = (0, v)$, so that $u^2 + v^2 = s^2$. The midpoint M has coordinates $\left(\frac{1}{2}u, \frac{1}{2}v\right)$, the vector $\overrightarrow{MC}$ is

$$\frac{d}{s}\langle v, u\rangle,$$

and the vector equation $\overrightarrow{OC} = \overrightarrow{OM} + \overrightarrow{MC}$ gives the coordinates (x, y) of C:

$$\begin{cases} x = \frac{1}{2}u + \frac{d}{s}v \\ y = \frac{d}{s}u + \frac{1}{2}v \end{cases} \tag{1}$$

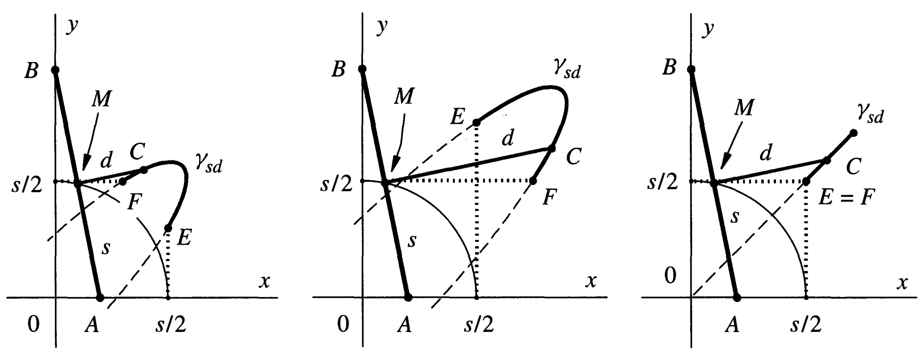
(FIGURE 2). Parameterize by writing $u = s\cos\theta$ and $v = s\sin\theta$ for $0 \le \theta \le 360°$ to obtain the parametric equations

$$\begin{cases} x = \frac{1}{2}s\cos\theta + d\sin\theta \\ y = d\cos\theta + \frac{1}{2}s\sin\theta, \end{cases} \tag{2}$$

equations we shall need again. Eliminating the parameter θ leads to the equation

$$4(s^2 + 4d^2)x^2 - 32sdxy + 4(s^2 + 4d^2)y^2 = (s^2 - 4d^2)^2, \tag{3}$$

whose graph is a (possibly degenerate) ellipse centered at the origin having axes along the 45° lines $y = \pm x$ and whose semi-axes have lengths $\left|d \pm \frac{1}{2}s\right|$. The graph collapses to a line segment when the left side is a square, i.e., when $d = \pm\frac{1}{2}s$, in which case the locus lies along one of the 45° lines $y = \pm x$.



a. $d < \frac{1}{2}s$ b. $d > \frac{1}{2}s$ c. $d = \frac{1}{2}s$

Figure 2 Three loci

This result is a special case of a locus investigated in 1643 by the Dutch mathematician Frans van Schooten the younger (see Dörrie [**3**, pp. 214–216]):

> If two vertices of a given triangular tile are restricted to move along two fixed intersecting lines, the locus of the third vertex is a (possibly degenerate) ellipse whose center lies at the point of intersection of the two lines.

Pedoe [**11**, p. 97] (see also [**10**]) claims that the locus was known more than a century earlier by Leonardo da Vinci, but he cites no source. The locus has many interesting special cases, including the well-known *elliptic compass* or *trammel of Archimedes*, also called the *ellipsograph*. See, for example, Taimina [**12**], Apostol and Mnatsakanian [**1**], Honsberger [**7**, pp. 173–177], and Wetzel [**13**].

In the cases that concern us, the segment AB lies in the first quadrant with A on the positive x-axis and B on the positive y-axis, and $d \ge 0$. So $u \ge 0$, $v \ge 0$, and $0 \le \theta \le 90°$, and the locus of C is an arc γ_{sd} of the ellipse (3).

Three special situations are of interest to us:

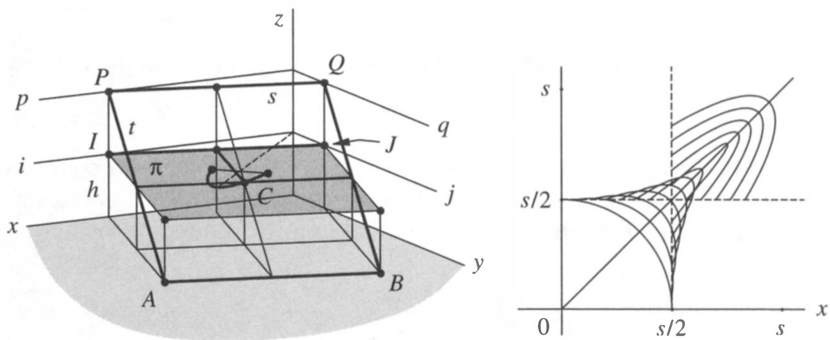
LOCUS 1. When $d = 0$, the point C is the midpoint M of the segment AB, and it traces the quarter circle $x^2 + y^2 = \frac{1}{4}s^2$ with endpoints $\left(\frac{1}{2}s, 0\right)$ and $\left(0, \frac{1}{2}s\right)$ and center at the origin (FIGURE 2).

LOCUS 2. When $0 < d \neq \frac{1}{2}s$, the point C traces the arc γ_{sd} of the ellipse (3) from the point $E\left(\frac{1}{2}s, d\right)$ to the point $F\left(d, \frac{1}{2}s\right)$ as θ increases from 0 to 90° (see FIGURES 2a,b).

LOCUS 3. When $d = \frac{1}{2}s$, as noted above, the point C traces the line segment with endpoints $\left(\frac{1}{2}s, \frac{1}{2}s\right)$ and $\left(\frac{1}{2}s\sqrt{2}, \frac{1}{2}s\sqrt{2}\right)$ twice (FIGURE 2c).

The surface via cross sections

We return to our examination of the surface $\mathcal{S}_z(s, t)$. Recall that we place the tile $\mathbb{T} = ABQP$ in such a way that its corners A and B lie in the xy-plane, P lies in the xz-plane, and Q lies in the yz-plane (FIGURES 3a, 4). Then the surface $\mathcal{S}_z(s, t)$ is the locus of the center C of $\mathbb{T}$ that we seek to describe.



a. Plane $z = h$ b. Possible traces

Figure 3 Horizontal cross section

Horizontal sections We begin our investigation by determining the horizontal cross sections of $\mathcal{S}_z(s, t)$.

For each h in $[0, t/2]$ let π be the plane with equation $z = h$ (FIGURE 3a). Then the point C lies on the perpendicular bisector of projection $\overline{IJ}$ of $\overline{PQ}$ in π and lies at distance

$$d = \frac{1}{2}\sqrt{t^2 - 4h^2} \tag{4}$$

from $\overline{IJ}$, and it generates the desired cross section in π as I moves in the xz-plane on the ray i midway between p and the x axis and J moves in the yz-plane on the ray j midway between q and the y-axis. Thus from Loci 1, 2, and 3 we have:

LEMMA 1. *For each h, $0 \leq h \leq \frac{1}{2}t$, the cross section of $\mathcal{S}_z(s, t)$ in the plane $z = h$ is the arc in the first octant with endpoints $\left(d, \frac{1}{2}s, h\right)$ and $\left(\frac{1}{2}s, d, h\right)$ of the (possibly degenerate) ellipse*

$$4(s^2 + t^2 - 4h^2)(x^2 + y^2) - 16s\sqrt{t^2 - 4h^2}xy = (4h^2 + s^2 - t^2)^2. \tag{5}$$

The center of this ellipse is $(0, 0, h)$*, its axes lie along the* $45°$ *lines, its semiaxes have lengths* $\left|d \pm \frac{1}{2}s\right|$*, and its vertex in the first octant is the point*

$$\left(\tfrac{1}{2}\sqrt{2}\left(d + \tfrac{1}{2}s\right), \tfrac{1}{2}\sqrt{2}\left(d + \tfrac{1}{2}s\right), h\right),$$

where d *is given by (4).*

We note a few consequences. To find the top of the surface we set $h = \frac{1}{2}t$, so that $d = 0$, and the top of the surface is the horizontal quarter circle with endpoints $\left(\frac{1}{2}s, 0, h\right)$ and $\left(0, \frac{1}{2}s, h\right)$, according to Locus 1. To determine the trace of the surface in the xy-plane we set $h = 0$, so that $d = \frac{1}{2}t$, and the trace in the xy-plane is the elliptic arc γ_0 whose equation is

$$4(s^2 + t^2)x^2 - 16stxy + 4(s^2 + t^2)y^2 = (s^2 - t^2)^2 \tag{6}$$

and whose endpoints are $\left(\frac{1}{2}t, \frac{1}{2}s, 0\right)$ and $\left(\frac{1}{2}s, \frac{1}{2}t, 0\right)$.

Further, if $s \leq t$, then the discriminant

$$-64(4h^2 + s^2 - t^2)^2$$

of (5) vanishes when $h = h_0 = \frac{1}{2}\sqrt{t^2 - s^2}$, so the cross section collapses to a line segment at the height h_0.

Replacing h by z in (5), we see that $\mathcal{S}_z(s, t)$ satisfies a polynomial equation $\mathcal{P}_{s,t}(x, y, z) = 0$ of degree eight.

Quarter-circular sections Somewhat surprisingly, $\mathcal{S}_z(s, t)$ can also be regarded as an assemblage of congruent quarter circles, although they are not cross sections lying in parallel planes. Let m be the line segment that joins the midpoints of $\overline{AB}$ and $\overline{PQ}$ in $\mathbb{T}$ (FIGURE 1).

LEMMA 2. *For each* u*,* $0 \leq u \leq s$*, let* $v = \sqrt{s^2 - u^2}$*. Then the plane*

$$ux - vy = \frac{1}{2}(u^2 - v^2) \tag{7}$$

perpendicular to $\mathbb{T}$ *through its midline* m *passes through the point* $\left(\frac{1}{2}u, \frac{1}{2}v, \frac{1}{2}t\right)$ *and meets the surface* $\mathcal{S}_z(s, t)$ *in a quarter circle* $\mathbb{Q}_{st}(u)$ *whose radius is* $\frac{1}{2}t$*, whose center is* $\left(\frac{1}{2}u, \frac{1}{2}v, 0\right)$*, whose upper endpoint* $\left(\frac{1}{2}u, \frac{1}{2}v, \frac{1}{2}t\right)$ *lies on the quarter circle* $u^2 + v^2 = s^2$ *in the plane* $z = \frac{1}{2}t$*, and whose lower endpoint*

$$\left(\frac{1}{2}u + \frac{t}{2s}v, \frac{t}{2s}u + \frac{1}{2}v, 0\right)$$

lies on the elliptic arc γ_0 *with equation* (6).

Suppose that initially the tile $\mathbb{T}$ is parallel to the z-axis, and has its corners A, B, P, and Q at the points $A'(u, 0, 0)$ on the positive x-axis, $B'(0, v, 0)$ on the positive y-axis, $P'(u, 0, t)$, and $Q'(0, v, t)$. The center C of $\mathbb{T}$ is initially at the point $C'\left(\frac{1}{2}u, \frac{1}{2}v, \frac{1}{2}t\right)$, a point that lies on the circular arc $u^2 + v^2 = s^2$ in the plane $z = \frac{1}{2}t$. Now permit the tile $\mathbb{T}$ to slide downward, keeping P on $A'P'$, Q on $B'Q'$, and the lower edge AB in the xy-plane parallel to $A'B'$. (FIGURE 4 shows $\mathbb{T}$ partway down.) Then the medial segment m moves with its endpoints on two perpendicular lines, so (according to Locus 1) its midpoint C traces a quarter circle that lies in a vertical plane and joins the center C' of the vertical tile $B'A'P'Q'$ to the point in which the plane (7) meets the trace γ_0 in the xy-plane.

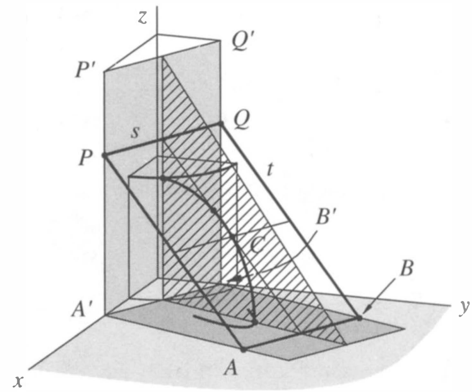


Figure 4 Quarter circle sections

Hence the surface $\mathcal{S}_z(s,t)$ can be regarded as the union of the quarter circles $\mathbb{Q}_{st}(u)$ for $0 \le u \le s$. Perhaps one could say that the surface is "ruled" by this family of quarter circles, in the same way that a ruled surface is formed by the union of a family of lines.

The surface We have found two quite different descriptions of the locus surface $\mathcal{S}_z(s,t)$, which, combined, give us a reasonably good idea of its shape.

THEOREM 3. *An $s \times t$ rectangular tile $\mathbb{T}$ moves in the positive octant of Euclidean space with one edge of length s lying in the xy-plane and a corner in each of the remaining two coordinate planes. Then the locus of the center C of $\mathbb{T}$ is the surface $\mathcal{S}_z(s,t)$ given by*

$$\mathcal{S}_z(s,t) = \bigcup_{0\le h\le t/2} \gamma_{sd}(h) = \bigcup_{0\le u\le s} \mathbb{Q}_{st}(u)$$

where $d = \frac{1}{2}\sqrt{t^2-4h^2}$.

The shape of the surface $\mathcal{S}_z(s,t)$ clearly depends on the shape of the tile, i.e., on how the dimensions s and t of $\mathbb{T}$ are related. To learn more we examine the projection of the surface $\mathcal{S}_z(s,t)$ into the xy-plane. The cross section elliptic arc in the plane $z = h$ projects onto the arc γ_{sd} of the ellipse centered at the origin whose axes lie along the lines $y = \pm x$, whose semiaxes have lengths $\frac{1}{2}\left|\sqrt{t^2-4h^2} \pm s\right|$ and whose endpoints are the points $(\frac{1}{2}s, \frac{1}{2}\sqrt{t^2-4h^2}, 0)$ and $(\frac{1}{2}\sqrt{t^2-4h^2}, \frac{1}{2}s, 0)$ (as pictured in FIGURE 2). The circular arcs project to line segments that connect points on the quarter circle $x^2+y^2=\frac{1}{4}s^2$ to the appropriate points on γ_0.

Preliminary sketches suggest that some of the projected ellipses have an envelope. Parametric equations for this envelope can be obtained by adding the Jacobian condition

$$\frac{\partial(x,y)}{\partial(\theta,d)} = \begin{vmatrix} -\frac{1}{2}s\sin\theta + d\cos\theta & \sin\theta \\ -d\sin\theta + \frac{1}{2}s\cos\theta & \cos\theta \end{vmatrix} = d - s\sin\theta\cos\theta = 0 \tag{8}$$

to the system (2) (see Green [6] and the classical analysis sources cited there). The resulting parametric equations for the envelope are

$$\begin{cases} x = \frac{1}{2}s(\cos\theta + 2\sin^2\theta\cos\theta) = \frac{1}{2}s(\cos^3\theta + 3\sin^2\theta\cos\theta) \\ y = \frac{1}{2}s(\sin\theta + 2\cos^2\theta\sin\theta) = \frac{1}{2}s(\sin^3\theta + 3\cos^2\theta\sin\theta) \end{cases} \tag{9}$$

for appropriate values of θ, formulas that can be rewritten in the form

$$\begin{cases} x + y = \frac{1}{2}s(\cos\theta + \sin\theta)^3 = s\sqrt{2}\sin^3(\theta + 45°) \\ x - y = \frac{1}{2}s(\cos\theta - \sin\theta)^3 = s\sqrt{2}\cos^3(\theta + 45°). \end{cases}$$

It follows that if the family of projected cross section elliptic arcs has an envelope, it is an arc of the astroid

$$\left(\frac{x+y}{\sqrt{2}}\right)^{2/3} + \left(\frac{x-y}{\sqrt{2}}\right)^{2/3} = s^{2/3}, \tag{10}$$

whose graph in the first quadrant is pictured in FIGURE 5.

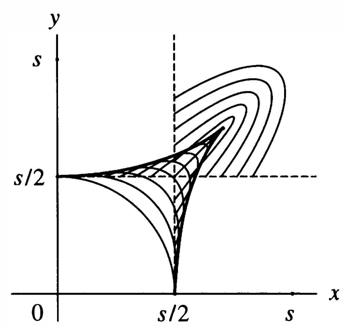
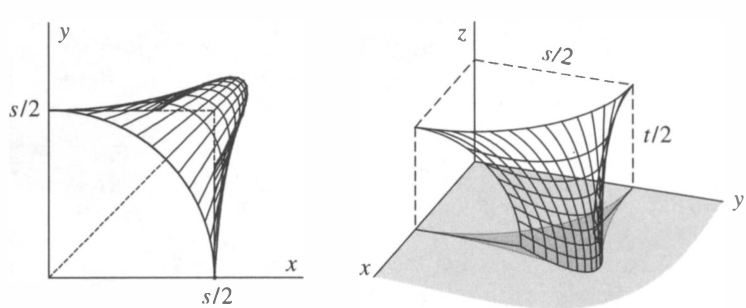


Figure 5 Envelope

The projected elliptic arcs γ_{st} lie out and beyond the astroid when $s < t$, and they fall within the curve when $s > t$. We consider the cases separately.

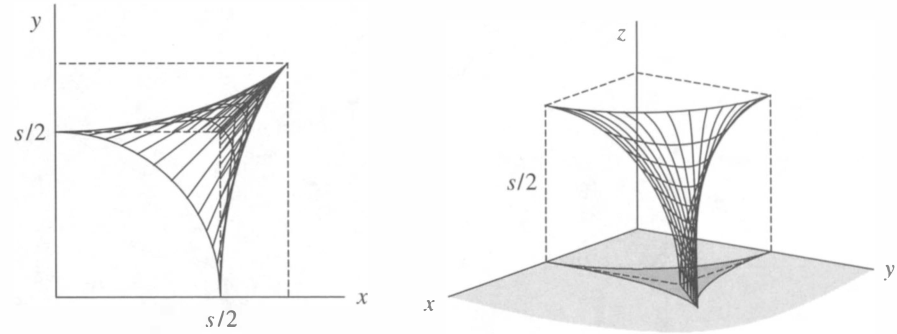
Case 1. $s > t$. In this case the projection of the surface $\mathcal{S}_z(s, t)$ is a region bounded by the cross section for $h = 0$, whose endpoints are $\left(\frac{1}{2}s, \frac{1}{2}t, 0\right)$ and $\left(\frac{1}{2}t, \frac{1}{2}s, 0\right)$, the quarter circle of radius $\frac{1}{2}s$ centered at the origin, and the two arcs of the astroid (10) that join the ends (FIGURE 6a). FIGURE 6b is a wireframe drawing of the surface showing both the cross section elliptic arcs and the generating circles.

Case 2. $s = t$. The projection of the surface $\mathcal{S}_z(s, s)$ in this case is the region bounded by the quarter circle of radius $s/2$ and the portion of the astroid that lies



a. Projection b. Surface

Figure 6 The case $s > t$

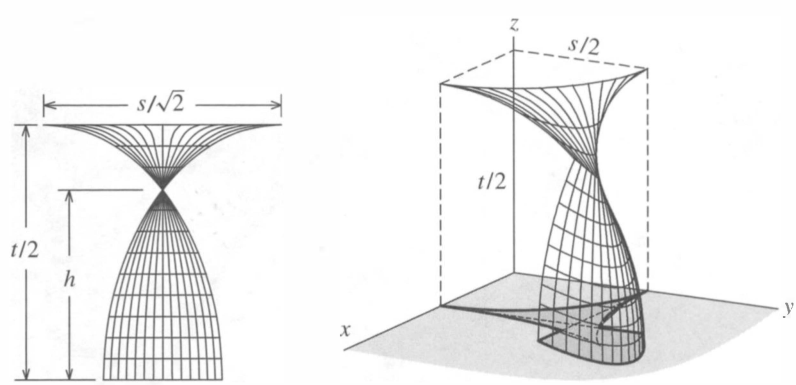


a. Projection b. Surface

Figure 7 The case $s = t$

in the first quadrant (FIGURE 7a). FIGURE 7b is a wireframe drawing of the surface showing both the cross section elliptic arcs and the generating circles. The quarter-circular arcs $\mathbb{Q}_{ss}$ join points on the upper quarter circle of radius $\frac{1}{2}s$ to points on the line segment with endpoints $\left(\frac{1}{2}s, \frac{1}{2}s\right)$ and $\left(\frac{1}{2}s\sqrt{2}, \frac{1}{2}s\sqrt{2}\right)$, so the surface is *prow*-shaped, tailing down to the line segment at the base.

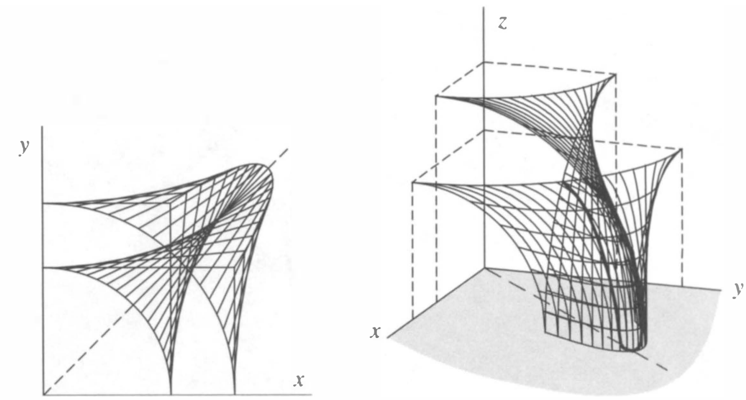
Case 3. $s < t$. In this case, $0 < h_0 = \frac{1}{2}\sqrt{t^2 - s^2} < \frac{1}{2}t$, so (as noted above) the elliptic cross section (5) collapses to the line segment with endpoints $\left(\frac{1}{4}s\sqrt{2}, \frac{1}{4}s\sqrt{2}, h_0\right)$ and $\left(\frac{1}{2}s\sqrt{2}, \frac{1}{2}s\sqrt{2}, h_0\right)$ through which all of the quarter circles $\mathbb{Q}_{st}$ pass. The projection of this surface into the xy-plane is shown in FIGURE 5. FIGURE 8b is a wireframe drawing of the surface, and a view showing the pinch appears in FIGURE 8a.



a. Pinch b. Surface

Figure 8 The case $s < t$

In the nonsquare case $s \neq t$, the surface we seek is the union of the two surfaces shown in FIGURES 6 and 8, so the locus of the center C is the combined surface $\mathcal{T}_z(s, t) = \mathcal{S}_z(s, t) \cup \mathcal{S}_z(t, s)$ (FIGURE 9). The surfaces $\mathcal{S}_z(s, t)$ and $\mathcal{S}_z(t, s)$ meet the xy-plane in the *same* elliptic arc, and although it is difficult to see in FIGURE 9, the taller surface bulges through the shorter.



a. Projection b. Combined surface

Figure 9 The surface $T_z(s, t) = S_z(s, t) \cup S_z(t, s)$

The locus Finally we arrive at the solution to our problem.

THEOREM 4. *The locus of the center of an $s \times t$ rectangular tile that moves in the first octant with two corners in one coordinate plane and one corner in each of the other two coordinate planes is*

a. *When $s = t$, $\mathcal{S}(s) = \mathcal{S}_x(s, s) \cup \mathcal{S}_y(s, s) \cup \mathcal{S}_z(s, s)$, where $\mathcal{S}_z(s, s)$ is the surface shown in* FIGURE 7.

b. *When $s \neq t$, $\mathcal{S}(s, t) = \mathcal{T}_x(s, t) \cup \mathcal{T}_y(s, t) \cup \mathcal{T}_z(s, t)$, where $\mathcal{T}_z(s, t) = \mathcal{S}_z(s, t) \cup \mathcal{S}_z(t, s)$ is the combined surface shown in* FIGURE 9.

FIGURE 10a shows the full surface $\mathcal{S}(s)$, and FIGURE 10b shows the full surface $\mathcal{S}(s, t)$ in the case $s/t = 2/3$.

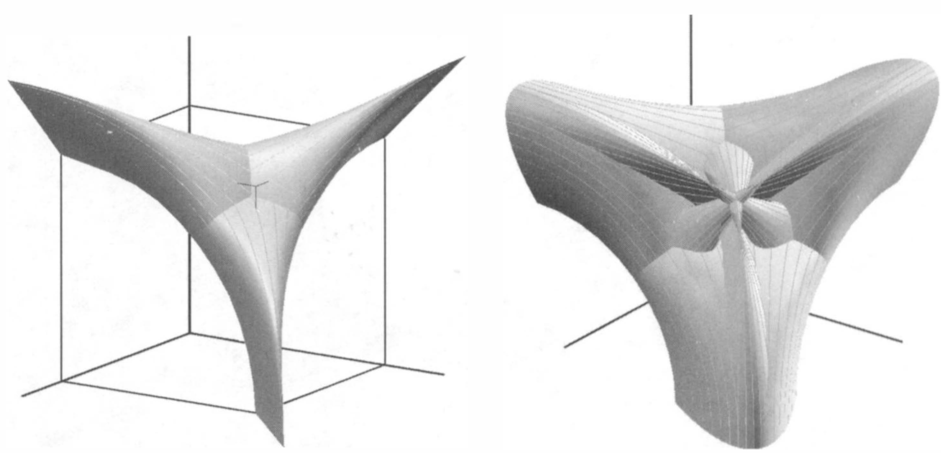


Figure 10 a. $\mathcal{S}(s)$ b. $\mathcal{S}(s, t)$, with $s/t = 2/3$

Related questions Some closely related problems are of interest. What is the locus of the center of an equilateral triangle of unit side that moves with one vertex on each of the coordinate planes?

What is the locus of the center of a cube of side 1 that moves in the first quadrant with one corner on each of the three bounding quarter-planes? What about the center of an $r \times s \times t$ block (a rectangular parallelopiped)?

Acknowledgment. We acknowledge with pleasure the insightful contributions of a referee, who in addition to correcting several not so minor errors made several suggestions and observations that significantly improved the article.

REFERENCES

1. T. M. Apostol and M. A. Mnatsakanian, A new look at the so-called trammel of Archimedes, *Amer. Math. Monthly* **116** (2009) 115–133.
2. W. B. Carver, Solution to Problem 3036, *Amer. Math. Monthly* **32** (1925) 47–49. doi:10.2307/2300097
3. H. Dörrie, *100 Great Problems of Elementary Mathematics*, Dover, New York, 1965.
4. F. W. Garnett, Problem 3036, *Amer. Math. Monthly* **30** (1923) 337. doi:10.2307/2300282
5. A. M. Gleason, R. E. Greenwood, and L. M. Kelly, *The William Lowell Putnam Mathematical Competition problems and solutions: 1938–1964*, MAA, Washington, DC, 1980.
6. J. W. Green, On the envelope of curves given in parametric form, *Amer. Math. Monthly* **59** (1952) 626–628. doi:10.2307/2306769
7. R. Honsberger, *Ingenuity in Mathematics*, New Mathematical Library #23, MAA, Washington, DC, 1970.
8. R. P. Jerrard, J. Schneider, R. Smallberg, and J. E. Wetzel, Straw in a box, *College Math. J.* **37** (2006) 93–102.
9. G. W. Mackey, The William Lowell Putnam mathematical competition, *Amer. Math. Monthly* **55** (1948) 630–633. doi:10.2307/2305618
10. D. Pedoe, The ellipse as an hypotrochoid, this MAGAZINE **48** (1975) 228–230.
11. ———, *Geometry and the Liberal Arts*, St. Martin's Press, New York, 1976.
12. D. Taimina, Historical mechanisms for drawing curves, in *Hands On History: A Resource for Teaching Mathematics*, MAA Notes, vol. 72, Ann Shell-Gellasch, ed., MAA, Washington, DC, 2007, 89–104.
13. J. E. Wetzel, An ancient elliptic locus, *Amer. Math. Monthly* (to appear).
14. ———, Rectangles in rectangles, this MAGAZINE **73** (2000) 204–211.
15. ———, Fits and covers, this MAGAZINE **76** (2003) 349–363, 398.

PROBLEMS

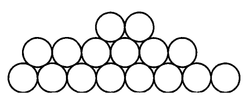
doi:10.4169/193009809X471090

PROPOSALS

To be considered for publication, solutions should be received by March 1, 2010.

1826. *Proposed by Michael Woltermann, Washington & Jefferson College, Washington, PA.*

A block fountain of coins is an arrangement of n identical coins in rows such that the coins in the first row form a contiguous block, and each row above that forms a contiguous block. As an example,



If a_n denotes the number of block fountains with exactly n coins in the base, then $a_n = F_{2n-1}$, where F_k denotes the kth Fibonacci number. (Wilf, generatingfunctionology, 1994.) How many block fountains are there if two fountains that are mirror images of each other are considered to be the same? That is, if two fountains such as

are the same, while two fountains such as

are different?

We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals must, in general, be accompanied by solutions and by any bibliographical information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution.

Solutions should be written in a style appropriate for this MAGAZINE. Each solution should begin on a separate sheet.

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

1827. *Proposed by Christopher Hilliar, Texas A & M University, College Station, TX.*

Let A be an $n \times n$ matrix with integer entries and such that each column of A is a permutation of the first column. Prove that if the entries in the first column do not sum to 0, then this sum divides $\det(A)$.

1828. *Proposed by Stephen J. Herschkorn, Department of Statistics, Rutgers University, New Brunswick, NJ.*

Let α_0 be the smallest value of α for which there exists a positive constant C such that

$$\prod_{k=1}^{n} \frac{2k}{2k-1} \leq Cn^{\alpha}$$

for all positive integer n.

a. Find the value of α_0.

b. Prove that the sequence

$$\left\{ \frac{1}{n^{\alpha_0}} \prod_{k=1}^{n} \frac{2k}{2k-1} \right\}_{n=1}^{\infty}$$

is decreasing and find its limit.

1829. *Proposed by Oleh Faynshteyn, Leipzig, Germany.*

Let ABC be a triangle with $BC = a$, $CA = b$, and $AB = c$. Let r_a denote the radius of the excircle tangent to BC, r_b the radius of the excircle tangent to CA, and r_c the radius of the excircle tangent to AB. Prove that

$$\frac{r_a r_b}{(a+b)^2} + \frac{r_b r_c}{(b+c)^2} + \frac{r_c r_a}{(c+a)^2} \geq \frac{9}{16}.$$

1830. *Proposed by H. A. ShahAli, Tehran, Iran.*

Let α and β be positive real numbers and let r be a positive rational number. Find necessary and sufficient conditions to ensure that there exist infinitely many positive integers m such that

$$\frac{\lfloor m\alpha \rfloor}{\lfloor m\beta \rfloor} = r.$$

Quickies

Answers to the Quickies are on page 316.

Q993. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

Let (X, d) be a metric space and let

$$A = \{f : f \text{ is a real valued nonexpansive function on } X\}.$$

Let x_0 and y_0 be two given points in X. Find

$$\sup\{|f(x_0) - f(y_0)| : f \in A\},$$

and justify your answer. (We say $f : X \to R$ is nonexpansive if $|f(x) - f(y)| \leq d(x, y)$ for all x and y in X.)

Q994. *Proposed by Ovidiu Furdui, Campia Turzii, Cluj, Romania.*

Let k be a positive real number. Find the value of

$$\int_0^1 \int_0^1 \left\{ \left(\frac{1}{x}\right)^k - \left(\frac{1}{y}\right)^k \right\} dx\, dy,$$

where $\{z\} = z - \lfloor z \rfloor$ denotes the fractional part of z.

Solutions

An unmatched tournament **October 2008**

2001. *Proposed by José H. Nieto, Universidad del Zulia, Maracaibo, Venezuela.*

A chess club has n members. Each member of the club has played against all but k of the other members. The club decides to hold a tournament in which each member plays exactly one game against those he/she has not played before. The tournament is played in rounds, with each player playing at most one game each round. Each round is scheduled by randomly selecting pairs who have not previously played against each other (and who are not already scheduled for the round) until no more such pairs are available for the round. Determine the maximum possible number of rounds for such a tournament. (For example, if the club has six members A, B, C, D, E, and F and the pairs that have never played against each other are AB, AC, AD, BC, BE, CF, DE, DF, and EF, then the tournament could consist of the following rounds:

$$1: \ AB,\ CF,\ DE, \qquad 2: \ BC,\ DF \qquad 3: \ AC,\ EF \qquad 4: \ AD,\ BE.)$$

Solution by Allen Schwenk, Western Michigan University, Kalamazoo, MI.

We first show that at most $2k - 1$ rounds are needed. We model the situation with a regular graph of of n vertices and degree k. The vertices represent the players, and two vertices are connected by an edge if and only if the two corresponding players have not played. Now assume $2k - 1$ rounds have been played and that two players, x and y, have not yet played. Because xy is an edge and each vertex is of order k, each of x and y could have been assigned to play in at most $k - 1$ rounds, so there can be at most $2k - 2$ rounds in which the xy pairing was not possible. Thus there must be at least one round of the $2k - 1$ rounds in which neither x nor y is assigned to play. However, this is impossible because for each round, pairings are assigned until there are no remaining pairs who have not played.

We now show that there are situations in which $2k - 1$ rounds are necessary. Suppose we have $n = 2k + 2$ players (or vertices) labeled x, y, u_i, v_i for $1 \le i \le k$ and that the pairs that have not played (e.g., edges) are xu_i, $u_i v_j$, $v_i y$ for $1 \le i,\ j \le k$, and $i \neq j$. It is easy to check that every vertex has degree k as required. Now assume that the first $k - 1$ rounds are scheduled with k matches in each round as follows:

> Reading subscripts modulo k, round m, $1 \le m \le k - 1$ consists of matches $u_i v_{i+m}$, $1 \le i \le k$.

For each round the same two players, x and y, are unmatched, and because they will not play each other, neither can be included in these first $k - 1$ rounds. Thus after these $k - 1$ rounds are completed, x still has to play k games. This means the tournament will last at least $2k - 1$ rounds.

Also solved by Robert B. Eggleton, and the proposer.

An inequality **October 2008**

2002. *Proposed by Dorin Marghidanu, Colegiul National "A. I. Cuza," Corabia, Romania.*

Let $a_1, a_2, \ldots, a_n$ be positive real numbers. Prove that

$$\frac{a_1^2}{a_1+a_2}+\frac{a_2^2}{a_2+a_3}+\cdots+\frac{a_{n-1}^2}{a_{n-1}+a_n}+\frac{a_n^2}{a_n+a_1}\geq\frac{a_1+a_2+\cdots+a_n}{2}.$$

I. *Solution by Robert L. Doucette, McNeese State University, Lake Charles, LA.*
Cauchy's inequality for real numbers may be written

$$(x_1y_1+x_2y_2+\cdots+x_ny_n)^2\leq(x_1^2+x_2^2+\cdots+x_n^2)(y_1^2+y_2^2+\cdots+y_n^2).$$

Setting $x_i=\sqrt{a_i+a_{i+1}}$ and $y_i=a_i/\sqrt{a_i+a_{i+1}}$, for $1\leq i\leq n$, yields (with $a_{n+1}=a_1$)

$$(a_1+a_2+\cdots+a_n)^2$$
$$\leq 2(a_1+a_2+\cdots+a_n)\left(\frac{a_1^2}{a_1+a_2}+\frac{a_2^2}{a_2+a_3}+\cdots+\frac{a_{n-1}^2}{a_{n-1}+a_n}+\frac{a_n^2}{a_n+a_1}\right)$$

This inequality is equivalent to the desired result.

II. *Solution by Northwestern University Math Problem Solving Group, Northwestern University, Evanston, IL.*
Let $a_{n+1}=a_1$ and consider the two sums

$$S=\sum_{k=1}^{n}\frac{a_k^2}{a_k+a_{k+1}}\qquad\text{and}\qquad T=\sum_{k=1}^{n}\frac{a_{k+1}^2}{a_k+a_{k+1}}.$$

Then

$$S-T=\sum_{k=1}^{n}\frac{(a_k^2-a_{k+1}^2)}{a_k+a_{k+1}}=\sum_{k=1}^{n}(a_k-a_{k+1})=0,$$

so $S=T$. On the other hand, by the quadratic mean-arithmetic mean inequality,

$$\frac{a_k^2+a_{k+1}^2}{a_k+a_{k+1}}\geq\frac{a_k+a_{k+1}}{2},$$

so

$$S+T=\sum_{k=1}^{n}\frac{a_k^2+a_{k+1}^2}{a_k+a_{k+1}}\geq\sum_{k=1}^{n}\frac{a_k+a_{k+1}}{2}=\sum_{k=1}^{n}a_k.$$

Because $S=T$, this becomes

$$S\geq\frac{1}{2}\sum_{k=1}^{n}a_k,$$

which is the desired inequality.

Also solved by George Apostolopoulos (Greece), Byoung Tae Bae (Spain), Michel Bataille (France), Brain Bradie, Robert Calcaterra, Minh Can, Hongwei Chen, Chip Curtis, Daniele Degiorgi (Switzerland), Roger B. Eggleton, Fejéntaláltuka Szeged Problem Solving Group (Hungary), E. S. Friedkin, Peter Haggstrom (Australia), Eugene A. Herman, John G. Huever (Canada), Bianca-Teodora Iordache (Romania), D. Kipp Johnson, Lucyna Kabza, Omran Kouba (Syria), Victor Y. Kutsenok, Harris Kwong, Elias Lampakis (Greece), David P. Lang, *Kee-Wai Lau, (China), Peter W. Lindstrom, Graham Lord, Junaid N. Mansuri, Kim McInturff, Edward Omey (Belgium), Paolo Perfetti (Italy), Éric Pité (France), Ángel Plaza (Spain), Gabriel T. Prăjitură, Sebastián García Sáenz (Chile), Edward Schmeichel, Armend Shabani (Republic of Kosova), Joel M. Siegel, John Simons (Netherlands), Nicholas C. Singer, Albert Stadler (Switzerland), John Sumner and Aida Kadic-Galeb, Marian Tetiva (Romania), George Tsapakidis (Greece), Zhexiu Tu, Michael Vowe (Switzerland), Winona State Problem Solvers, and the proposer. There were two incorrect submissions.*

A convex set of fixed points **October 2008**

2003. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

Let $(X, \langle\,\rangle)$ be a real inner product space, and let

$$B = \{x \in X : \|x\| \leq 1\}$$

be the unit ball in X, where $\|x\| = \sqrt{\langle x, x\rangle}$. Let $f : B \to B$ be a function satisfying $\|f(x) - f(y)\| \leq \|x - y\|$ for all $x, y \in B$. Prove that the set of fixed points of f is convex.

Solution by Eugene A. Herman, Grinnell College, Grinnell, IA.

Let a, b be distinct fixed points of f, and let c be a point on the line segment joining a and b different from either endpoint. We show that $f(c) = c$. We use the full version of the triangle inequality:

For all $x, y \in X$, $\|x + y\| \leq \|x\| + \|y\|$ with equality if and only if one of x, y is a nonnegative multiple of the other.

Since $c = ta + (1-t)b$ for some $t \in (0, 1)$, $a - c$ is a positive multiple $((1-t)/t)$ of $c - b$. Hence

$$\begin{aligned}\|a-b\| &= \|(a - f(c)) + (f(c) - b)\| \leq \|a - f(c)\| + \|f(c) - b\| \\ &= \|f(a) - f(c)\| + \|f(c) - f(b)\| \leq \|a - c\| + \|c - b\| = \|a - b\|\end{aligned}$$

Since the second of the above inequalities is an equality, we have

$$\|a - f(c)\| = \|a - c\|, \quad \|f(c) - b\| = \|c - b\| \tag{1}$$

Because the first of the inequalities is an equality, we conclude from the triangle inequality that one of $a - f(c)$, $f(c) - b$ is a nonnegative multiple of the other. Since (1) shows that neither of these vectors is zero, there exists $\alpha > 0$ such that $a - f(c) = \alpha(f(c) - b)$, and so $f(c) = ua + (1-u)b$, where $u = 1/(\alpha + 1) \in (0, 1)$. Substituting this expression for $f(c)$ and $ta + (1-t)b$ for c in the second of equations (1) yields $u = t$ and therefore $f(c) = c$.

Also solved by Michel Bataille (France), Paul Budney, Bruce S. Burdick, Robert Calcaterra, Hongwei Chen, Chip Curtis, Jim Delany, Charles R. Diminnie, Robert L. Doucette, Fejéntaláltuka Szeged Problem Solving Group (Hungary), Dmitry Fleischman, G.R.A.20 Problem Solving Group (Italy), Cody Guinan and Jennifer Pajda, Dan Jurca, Omran Kouba (Syria), Elias Lampakis (Greece), Missouri State University Problem Solving Group, Éric Pité (France), Edward Schmeichel, Nicholas C. Singer, Albert Stadler (Switzerland), John Sumner and Aida Kadic-Galeb, Marian Tetiva (Romania), Bob Tomper, Haohao Wang and Jerzy Wojdylo, and the proposer.

A generator for F_{q^n} **October 2008**

2004. *Proposed by Jody M. Lockhart and William P. Wardlaw, U.S. Naval Academy, Annapolis, MD.*

Let A be an $n \times n$ matrix over the finite field F_q of q elements, and assume that A has multiplicative order $\mathrm{ord}(A) = q^n - 1$. Prove or give a counterexample to the following statement:

A is a cyclic generator for F_{q^n}, that is, $\{0, A, A^2, \ldots, A^{q^n-1}\}$ is the finite field F_{q^n}.

Solution by Jim Delany, Emeritus, California Polytechnic State University, San Luis Obispo, CA.

Let $S = \{0, A, A^2, \dots, A^{q^n-1}\}$. The nonzero elements of S form a multiplicative (abelian) group, and the distributive law is obvious, so we must show that S is an additive (abelian) group.

Let $p(x)$ be the minimal polynomial of A and d be the degree of $p(x)$. Note that $d \leq n$ since $p(x)$ is a factor of $\det(xI - A)$, the characteristic polynomial of A. Let

$$T = \{a_0 I + a_1 A + \cdots + a_{d-1} A^{d-1} : a_i \in F_q\}.$$

Then T is a vector space of dimension d over F_q, and hence is an additive group. Note also that T has q^d elements and that $q^d \leq q^n$.

On the other hand $S \subset T$. To prove this, let $A^k \in S$ and write $x^k = p(x)q(x) + r(x)$ with $\deg(r(x)) < \deg(p(x)) = d$. Setting $x = A$ in this expression yields

$$A^k = p(A)q(A) + r(A) = 0 \cdot q(A) + r(A) = r(A) \in T.$$

Thus $q^n = |S| \leq |T| = q^d \leq q^n$ so $|S| = |T|$ and $S = T$. Hence S is a group under addition.

Also solved by Michel Bataille (France), Robert Calcaterra, Fejéntaláltuka Szeged Problem Solving Group (Hungary), Elias Lampakis (Greece), Éric Pité (France), Nicholas C. Singer, Gregory P. Wene, and the proposers.

A comparison test **October 2008**

2005. *Proposed by Ovidiu Furdui, Campia Turzii, Cluj, Romania.*

Let $f : [0, \infty) \to (0, \infty)$ be an increasing, differentiable function with continuous derivative, and let k be a nonnegative integer. Prove that

$$\int_0^\infty \frac{x^k}{f(x)}\,dx \quad \text{converges if and only if} \quad \int_0^\infty \frac{x^k}{f(x) + f'(x)}\,dx \quad \text{converges.}$$

Solution by Nicholas C. Singer, Annandale, VA.

Since $f'(x) \geq 0$ everywhere, the "only if" part is trivial. Suppose the second integral converges. Because f is a positive continuous function on $[0, 1]$, it attains a positive minimum there. Thus $\int_0^1 (x^k/f(x))\,dx$ is finite and can be neglected. For $X > 1$ and nonnegative integer k,

$$\int_1^X \frac{x^k dx}{f(x)} = \int_1^X \frac{x^k dx}{f(x) + f'(x)} + \int_1^X \frac{x^k f'(x) dx}{f(x)(f(x) + f'(x))}$$

$$\int_1^X \frac{dx}{f(x)} \leq \int_1^X \frac{dx}{f(x) + f'(x)} + \int_1^X \frac{f'(x)\,dx}{f^2(x)}$$

$$\leq \int_1^X \frac{dx}{f(x) + f'(x)} + \frac{1}{f(1)}.$$

The right hand side is bounded as $X \to \infty$, so the claim is true for $k = 0$. Suppose the claim is true for nonnegative integer k and suppose $\int_1^\infty x^{k+1} dx/(f(x) + f'(x))$ converges. The integrand dominates $x^k/(f(x) + f'(x))$, so $\int_1^\infty x^k dx/(f(x) + f'(x))$

converges. Then by the induction hypothesis, $\int_1^\infty x^k dx/f(x)$ converges. Now

$$\begin{aligned}\int_1^X \frac{x^{k+1}dx}{f(x)} &\le \int_1^X \frac{x^{k+1}dx}{f(x)+f'(x)} + \int_1^X \frac{x^{k+1}f'(x)\,dx}{f^2(x)}\\ &= \int_1^X \frac{x^{k+1}dx}{f(x)+f'(x)} - \frac{x^{k+1}}{f(x)}\Big|_1^X + \int_1^X \frac{(k+1)x^k dx}{f(x)}\\ &\le \int_1^X \frac{x^{k+1}dx}{f(x)+f'(x)} + \frac{1}{f(1)} + \int_1^X \frac{(k+1)x^k dx}{f(x)}.\end{aligned}$$

The terms on the right are bounded as $X \to \infty$, thus the claim is true for $k+1$ and we've proven the result by induction.

Also solved by Michel Bataille (France), Hongwei Chen, E. S. Friedkin, Kee-Wai Lau (China), Peter W. Lindstrom, José Miguel Pacheco (Spain) and Ángel Plaza (Spain), Paolo Perfetti (Italy), Raul A. Simon (Chile), and the proposer.

Answers

Solutions to the Quickies from page 311.

A993. Let $L = \sup\{|f(x_0) - f(y_0)| : f \in A\}$. We prove that $L = d(x_0, y_0)$. For any $f \in A$, we have $|f(x_0) - f(y_0)| \le d(x_0, y_0)$. Thus $L \le d(x_0, y_0)$. To prove the reverse inequality, define $g : X \to R$ by $g(x) = d(x, x_0)$. Then for x and y in X,

$$|g(x) - g(y)| = |d(x, x_0) - d(y, x_0)| \le d(x, y),$$

which proves that $g \in A$. Therefore

$$\begin{aligned}L &= \sup\{|f(x_0) - f(y_0)| : f \in A\}\\ &\ge |g(x_0) - g(y_0)| = |d(x_0, x_0) - d(x_0, y_0)| = d(x_0, y_0).\end{aligned}$$

Thus, $L \ge d(x_0, y_0)$, and it follows that $L = d(x_0, y_0)$.

A994. Let I denote the value of the integral. Note that if z is a real number and z is not an integer, then $\{z\} + \{-z\} = 1$. By symmetry

$$I = \int_0^1\int_0^1 \left\{\left(\frac{1}{x}\right)^k - \left(\frac{1}{y}\right)^k\right\} dx\,dy = \int_0^1\int_0^1 \left\{\left(\frac{1}{y}\right)^k - \left(\frac{1}{x}\right)^k\right\} dx\,dy.$$

Hence,

$$\begin{aligned}I &= \frac{1}{2}(I+I)\\ &= \frac{1}{2}\int_0^1\int_0^1 \left(\left\{\left(\frac{1}{x}\right)^k - \left(\frac{1}{y}\right)^k\right\} + \left\{\left(\frac{1}{y}\right)^k - \left(\frac{1}{x}\right)^k\right\}\right) dx\,dy\\ &= \frac{1}{2}\int_0^1\int_0^1 1\,dx\,dy = \frac{1}{2},\end{aligned}$$

because the set on which the integrand is 0 is a set of measure 0.

REVIEWS

PAUL J. CAMPBELL, *Editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles, books, and other materials are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of mathematics literature. Readers are invited to suggest items for review to the editors.

doi:10.4169/002557009X474340

Rosenhouse, Jason, *The Monty Hall Problem: The Remarkable Story of Math's Most Contentious Brain Teaser*, Oxford University Press, 2009; xii + 194 pp, $24.95. ISBN 978-0-19-536789-8. Chapter 1 and Bibliography at `http://www.math.jmu.edu/~rosenhjd/ChapOne.pdf`.

A prize lies behind one of three doors, you choose a door; before it is opened, another door is opened showing no prize there. Should you switch your choice? This "Monty Hall problem" (whose ultimate source was Martin Gardner), with its "perils of intuition," has bedeviled millions since its popularization in a magazine in 1990. Many people remain unconvinced (and even irritated) by arguments in favor of switching. Author Rosenhouse recounts the arguments for switching, points out the importance of how the opened door is chosen ("host behavior"), distinguishes variations of the problem, and discusses the cases of more doors and more than one prize. The book goes beyond the probability involved (a number of equations appear) to consider philosophical aspects ("single-case" vs. long-run probability) as well as the psychology involved in the problem:"The issue is people applying intuitive arguments to unfamiliar problems that can only be solved properly by more complex means." Rosenhouse directs focus to the key mathematical concepts of independence and of conditioning probability on events and knowledge, and he illustrates the ubiquity of the problem in many other contexts. This is a wonderful book, providing enlightenment in an informal and engaging style. (Rosenhouse also points out "an occupational hazard among mathematicians... the desire always to be the smartest person in the room... especially prevalent when

interacting with non-mathematicians," noting the "relish" with which some mathematicians declare others' arguments wrong. Let us all be chastened!) [Editor's note: By coincidence, "The Monty Hall Problem, Reconsidered" by Stephen Lucas, Jason Rosenhouse, and Andrew Schepler, will appear in the December issue of this MAGAZINE.]

Gelman, Andrew, and David Weakliem, Of beauty, sex, and power, *American Scientist* 97 (4) (July–August 2009) 3100–316, `http://www.stat.columbia.edu/~gelman/research/published/power4r.pdf`.

Do big and tall parents have more sons? Do violent men have more sons? Do engineers have more sons, nurses more daughters? Do beautiful parents have more daughters? Yes, to all of these, claims S. Kanazawa in a series of articles in the *Journal of Theoretical Biology*, in a popular article in *Psychology Today*, and in the 2007 book *Why Beautiful People Have More Daughters.* Authors Gelman and Weakliem show that Kanazawa's findings suffer from basic statistical flaws and are not statistically significant. Gelman and Weakliem usefully distinguish two relevant kinds of errors, Type M (magnitude) and Type S (sign). They go further to ask "how to interpret nonsignificant results" that are nevertheless suggestive because they have a large effect size. In fact, it was the claimed effect size (36%) that caught the authors' attention, since such an effect is one to two magnitudes larger than other reported sex-ratio effects. Their reaction: "Wow, this study is underpowered!" They note: "[The problem of implausibly large estimates from small samples] will occur again and again, and is worth thinking about now," particularly since "public discourse can get cluttered with unproven claims."

Lockhart, Paul, *A Mathematician's Lament: How School Cheats Us Out of Our Most Fascinating and Imaginative Art Form*, Bellevue Literary Press, 2009; 140 pp, $12.95. ISBN 978-1-934137-17-8. Part I at `http://www.maa.org/devlin/LockhartsLament.pdf`.

This pocket-sized book adds to the author's original essay, incorporated as "Part I: Lamentation," a follow-up of almost equal length, "Part II: Exultation." Keith Devlin has referred to this critique of current practice in school mathematics as a "recognized landmark in the world of mathematics education that cannot and should not be ignored." Apart from blunt passion and provocative prose, what adds impact to *Lament* is that the author, formerly a university researcher in algebra, for almost a decade has been teaching in a K-12 school. Lockhart laments that mathematics, the art of making patterns with ideas (G. H. Hardy's encapsulation), is "not recognized by our culture as [an art]. . . a meaningful human experience." The "art is not in the 'truth' but in the explanation, the argument." The "glory," and why it is fun, is that it is "completely irrelevant to our lives."

Lockhart's critique is that as instituted in schools, mathematics is a parade of techniques and an escalator of notations, devoid of the "great problems, their history, and the creative process" and of any context, perspective, or invention. School mathematics offers no mathematics ("the art of explanation"), no problems (only "insipid exercises"), and leaves nothing for the student to do or imagine (questions are answered without first being pondered). He laments that school is about training children to perform so that we can sort them, but maintains that "mental acuity comes from solving problems yourself, not from being told how to solve them." He illustrates his ideas with an excellent simple example of discovering the area of a triangle; and he illuminates the essay with Galileo-like dialogues, in which "Simplicio" plays devil's advocate to "Salviati" giving Lockhart's responses.

Lockhart admits that making problems central is "painful and creatively frustrating," that "I can't make you curious," and that "the only thing that I am interested in using mathematics for is to have a good time and help others to do the same." He claims that "the last thing anyone needs is to be trained" and that he need not justify mathematics on the basis of its practical value.

No doubt you have your own reactions, but let me give a brief summary of my take on all this: (1) He has a point, expressed very well, regarding what education should be about. (2) Children—and adults, too—vastly prefer other ways of having a good time (but note the popularity of puzzle books and Sudoku!). (3) Not only American culture but other cultures, too, value school mathematics almost solely for its utilitarian value, as training in useful algorithms applied to familiar settings. (4) There is already enough pain and frustration among students in confronting that training (not to mention their agony in trying to apply it to "word problems").

[Disclosures: I edit an journal devoted to fostering the use of mathematical modeling and applications in undergraduate mathematics instruction; but I have also written about the teaching of calculus in a vein similar to Lockhart's, in "Calculus is crap," *The UMAP Journal* 27 (2006) (4) 415–430. I would be happy to send a copy of that, and Woody Dudley's response, to any who inquire.]

Coppin, Charles A., W. Ted Mahavier, E. Lee May, and G. Edgar Parker, *The Moore Method: A Pathway to Learner-Centered Instruction*, MAA, 2009; xi + 245 pp, $57.50 ($47.50 to MAA members).ISBN 978-0-88385-185-2.

"That student is taught the best who is told the least." —R. L. Moore (1966). Moore established a tradition at the University of Texas of a discovery-based method of learning college mathematics, now known as the *Moore method*: no textbooks, no reading, no lectures, no conferring with others, just students thinking up on their own solutions to a guided sequence of problems and exposing their explanations at the blackboard to the curiosity and critique of other class members. This book, by long-time practitioners of variations of the Moore method, was written to serve as a resource for those who would like to try or adopt such a method themselves. The authors have used such a method in courses in topology, analysis, abstract algebra, introduction to abstract mathematics, and liberal-arts statistics through baseball. This remarkable book pools experience and expertise, offers specifics (syllabi, worksheets, handouts, diary entries, model exams), and conveys motivation and enthusiasm for using the method. Paul Lockhart, of the *Lament* reviewed above, would no doubt approve.

NEWS AND LETTERS

2008 Carl B. Allendoerfer Awards

doi:10.4169/193009809X471108

Vesna Stojanoska and Orlin Stoytchev, Touching the $\mathbb{Z}^2$ in Three-Dimensional Rotations, this MAGAZINE **81**:5 (December 2008) 345–357.

We think that we understand rotations in $\mathbb{R}^3$ entirely, but in fact we don't. Some complete rotations, that is, rigid motions of a body that keep one point fixed and return the body to its original position, cannot be continuously deformed to the null motion. But if we compose two such nontrivial complete rotations, the resulting motion can always be so deformed. This paper gives a mathematical formulation of this nonobvious geometric property.

If we choose a basis in $\mathbb{R}^3$, any rotation about a fixed axis can be expressed by means of a 3×3 orthogonal matrix with determinant 1. The set of all such matrices forms the special orthogonal group, SO(3), a fundamental mathematical object with various algebraic, topological, and analytic properties. A complete rotation can be viewed as a continuous path in SO(3), so the statement about the deformability of complete rotations is just the assertion that the fundamental group $\pi_1(\mathrm{SO}(3))$ is isomorphic to $\mathbb{Z}^2$.

The authors provide an elegant visual proof by relating three-dimensional rotations to braids. The argument is not easy, but it provides a good example of the interaction of algebraic and topological ideas and sheds light on an intriguing and deep physical property of the space in which we live. It offers a stimulating introduction to braids and topology and makes nontrivial mathematics accessible to undergraduates.

Biographical Notes

Vesna Stojanoska received her B.A. from the American University in Bulgaria and is now a Ph.D. student at Northwestern University. Her research is in algebraic topology: she is interested in using algebraic geometry and number theory to better understand various phenomena in stable homotopy theory.

Orlin Stoytchev is a professor at the American University in Bulgaria. He received his Ph.D. from Virginia Tech. His research interests can be summarized as "the different aspects of symmetries in mathematics and physics" and have led to works on Von Neumann algebras, representations of infinite-dimensional Lie groups and algebras, and recently on braid groups.

Response from Vesna Stojanoska and Orlin Stoytchev One of the most gratifying feelings for a mathematician is when, after a long struggle with a problem, he or she suddenly starts seeing the solution. Everything fits in place and becomes embarrassingly simple. Even though for most of us these occasions are rare and we only see little steps further, the satisfaction is not diminished. Perhaps the only greater joy is sharing with others the beauty of this amazing human creation—mathematics. We are extremely honored to receive the Allendoerfer Award. We would like to express our gratitude to the MAA for providing a forum for popularizing mathematics and to the former editor of MATHEMATICS MAGAZINE, Professor Allen Schwenk, for his guidance and encouragement.

Jeff Suzuki, A Brief History of Impossibility, this MAGAZINE **81**:1 (February 2008) 27–38.

When the trisection of the angle with only straightedge and compass is mentioned, we often see words to the effect, "In 1837 Wantzel showed that this is impossible." However, we are never told how Wantzel did it. In this paper Jeff Suzuki tells us, along with considerably more, in prose that is eminently clear and readable.

He starts at the beginning with Euclid, shows how Descartes made it possible to transform geometrical constructions into algebraic equations, mentions the work of Vandermonde and Lagrange on roots of polynomials, and gives a thorough outline of Gauss' proof that a regular 17-gon is constructible with straightedge and compass alone. He concludes with a proof of Wantzel's theorem that if a line segment of length γ can be constructed with only straightedge and compass then γ must be a zero of a polynomial of degree 2^n for some positive integer n.

The reader is drawn along, seemingly effortlessly, through results that are historically and mathematically important. The paper is reminiscent of the works of Euler. For expository writing there can be no higher praise.

Biographical Note

Jeff Suzuki, currently associate professor of mathematics at Brooklyn College, grew up in southern California with an inability to decide what he really wanted to do, so he earned a bachelor's in mathematics (with a physics concentration) and history from California State University, Fullerton. He went on to earn his M.A. and Ph.D. from Boston University, with a dissertation on the history of a topic in mathematical physics. Still unsure of what he really wants to do, he has spent some of the past year developing a course on cryptography, researching interdisciplinary topics in mathematics, and learning sign language. His wife Jacqui and children William X and Dorothy Z (yes, their middle names are X and Z) somehow manage to put up with his idiosyncrasies and have survived his culinary, literary, and musical efforts. His latest book, *Mathematics in Historical Context*, has just been published by the MAA (2009).

Response from Jeff Suzuki I am very honored to have won the Allendoerfer Award for "A Brief History of Impossibility." I wrote the paper to answer a question that had been on my mind for many years; I was quite pleased to discover that MATHEMATICS MAGAZINE found it interesting enough to publish. I'd like to thank the former editor, Allen Schwenk, and editorial assistant, Margo Chapman, for their work shepherding this manuscript through the publication process.

MAGAZINE Author Ellermeyer Wins Foundation Prize

The Kennesaw State University Foundation recognized Sean Ellermeyer's contribution to our April 2008 issue, "A Closer Look at the Crease Length Problem," with its Foundation Prize. Ellermeyer will receive a cash award and a grant to support his ongoing research activity.

The paper sheds new light on a classical calculus problem—that of finding the minimum possible length of a crease formed by folding a rectangular sheet of paper. When studied in its full generality, this problem is much more interesting and difficult than the version that usually appears in calculus textbooks. In particular, the solution of the general crease length problem is found to depend critically on whether the ratio of the side lengths of the paper being folded exceeds the square root of the golden ratio.

Congratulations to Sean Ellermeyer!

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

The Mathematical Association of America
1529 Eighteenth Street, NW
Washington, DC 20036