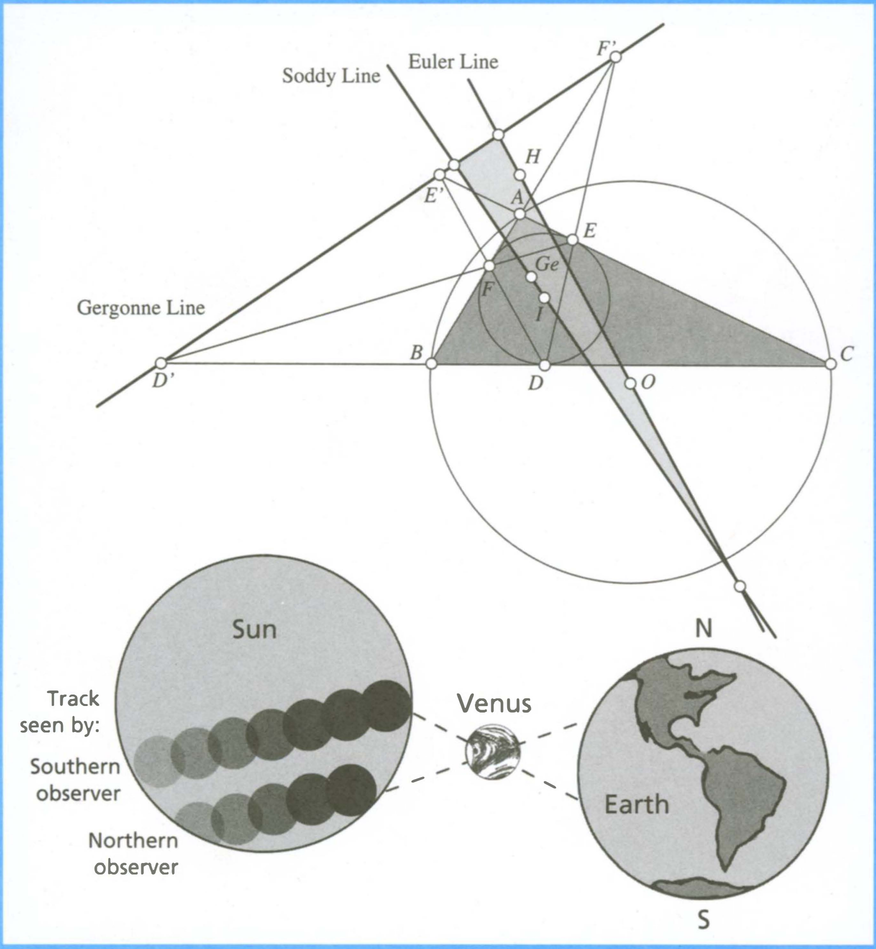
Vol. 76, No. 5, December 2003

MATHEMATICS MAGAZINE



- Transits of Venus and the Astronomical Unit
- Fits and Covers
- The Operator $(x\frac{d}{dx})^n$ and Its Applications to Series
- Symmetric Polynomials in the Work of Newton and Lagrange

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.

Submit new manuscripts to Frank A. Farris, Editor, Mathematics Magazine, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053-0373. Manuscripts should be laser printed, with wide line spacing, and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should mail three copies and keep one copy. In addition, authors should supply the full five-symbol 2000 Mathematics Subject Classification number, as described in *Mathematical Reviews.*

Cover image by Jason Challas, with assistance from Ahn Pham. On June 8, 2004, Venus will pass directly between the Earth and Sun. In this issue, Donald Teets explains how the geometry of this rare phenomenon was used in 1761 to determine the size of the solar system. Does Ray Beauregard's discussion of the Euler-Gergonne-Soddy triangle occur on a higher plane, as shown?

Jason Challas lectures on celestial triangles and computer art at Santa Clara University, where Ahn Pham is a student.

AUTHORS

Donald A. Teets received his B.A. from the University of Colorado, his M.S. from Colorado State University, and his Doctor of Arts from Idaho State University. He has taught at the South Dakota School of Mines and Technology since 1988, and served as chair of the Department of Mathematics and Computer Science from 1997–2000. His research interests lie in the intersection of mathematics, history, and astronomy. He received the Carl B. Allendoerfer writing award from the MAA in 2000 for an article on the astronomical work of Gauss. When he is not doing mathematics, he enjoys backpacking, cross-country skiing, mountain biking, and rock climbing.

John E. Wetzel received his bachelor's degree from Purdue and his Ph.D. (mathematics) from Stanford University, a student of Halsey Royden. He spent his entire academic career at the University of Illinois at Urbana-Champaign, from which he retired in 1999 after 38 years of service. Always interested in classical geometry, he has most recently been studying the ways in which one shape fits in another—questions he regards as "'fitting problems' for retirement."

Peter M. Knopf received his Ph.D. from Cornell University in 1977, and has been at Pace University for more than 19 years. He is an active researcher in classical harmonic analysis and potential theory. His favorite hobby is chess, in which he earned the title of National Master from the United States Chess Federation in 1996. The roots of his article began from his experiences teaching calculus. The author wondered how to sum certain generalizations of geometric series that are encountered in calculus. After solving this problem, it led him naturally to investigate the Maclaurin series of trigonometric functions and related questions.

Greg St. George received his Ph.D in the area of operator theory from the University of Montana in 1989 under the direction of Keith Yale. He is interested in the roots of mathematical ideas, and believes strongly in reading the original works. He is also enjoying rediscovering geometry. He lives in the Bitterroot valley of Montana with his wife Jan and 11-year old son Aaron. He enjoys gardening, baking bread, and making fruit wines.

Vol. 76, No. 5, December 2003

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 Dave Riska (*driska@maa.org*), Advertising Manager, the Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 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

Transits of Venus and the Astronomical Unit

DONALD A. TEETS
South Dakota School of Mines and Technology
Rapid City, SD 57701
donald.teets@sdsmt.edu

Edmond Halley. Leonhard Euler. Mason and Dixon (of "line" fame). Captain James Cook. James Short.

With the probable exception of the last man listed, these names are all very recognizable, yet it seems most unusual for them to appear together. What could they possibly have in common?

The answer, of course, is given away in the title of this article. All were players in the extraordinary story surrounding observations of the transits of Venus—that is, the passages of Venus across the disk of the Sun, as viewed from Earth—that took place in the eighteenth century. Only five transits of Venus are known to have been observed in the history of mankind, in 1639, 1761, 1769, 1874, and 1882. Thus no one alive today has seen one. But this will soon change, for the next transit will take place June 8, 2004, and another will follow on June 6, 2012.

Though transits of Venus are rare and beautiful astronomical events, they could not have earned a significant place in the history of science for aesthetic reasons alone. *The extraordinary attention devoted to these transits, especially in 1761 and 1769, was due to their usefulness in determining the length of the astronomical unit, that is, the mean distance from Earth to the Sun, in terms of terrestrial distance units such as miles.* Indeed, one estimate of the astronomical unit, computed from observations of the 1769 transit and published in 1771, differs from modern radar-based values by a mere eight-tenths of a percent [**5, 9**].

The first purpose of this article is to offer a glimpse into the rich history surrounding observations of the transits of Venus, especially the transit of 1761. But a second and more important purpose is to give a mathematical description of the methods used by Mr. James Short following the 1761 transit to deduce the length of the astronomical unit. As June 8, 2004 draws near, one is sure to read of the upcoming transit in the popular press. This article is intended to augment the popular accounts by providing mathematical insight into the event for those who are able to appreciate it.

Kepler's prediction and the first observed transits

Our story starts with the German astronomer Johannes Kepler in the early part of the seventeenth century. Though Kepler never witnessed a transit himself, his significance in the story is enormous for two reasons.

First, according to Kepler's Third Law, as it is now known, the ratio of the square of a planet's orbital period to the cube of its mean distance from the Sun is the same for all planets. From this law, the *relative* scale of the solar system can be determined simply by observing the orbital periods of the planets. In fact, Kepler's own estimates of the relative distances of the known planets from the Sun do not differ significantly from modern values. But Kepler was unable to translate his discovery of the relative scale of the solar system into absolute terms, for he badly underestimated the length of the astronomical unit. His estimate of 3469 Earth radii (actually the largest of several of his estimates) was roughly seven times too small, and so his understanding of absolute distances within the solar system remained considerably flawed [**13**]. Despite Kepler's naive estimate of the solar distance, his third law remains one of the great achievements in the history of science, and is unquestionably fundamental to understanding the size of the solar system.

The second connection between Kepler and the transit problem is much more direct. It was his prediction in 1629 of the transits of Mercury, in November of 1631, and Venus, in December of 1631, that led to the first-ever observations of such events. Kepler predicted that the Venus transit would not be visible in Europe, nevertheless he asked astronomers to keep watch on the 6th and 7th of December in case his calculations were imperfect. He also "directed his request to observe this transit... to sailors who would be on the high seas, and learned men in America..." [**13**]. Unfortunately, there is no evidence that anyone successfully observed the 1631 Venus transit. On the other hand, at least three people saw the transit of Mercury in 1631 as a result of Kepler's prediction. Of these, Pierre Gassendi wrote a detailed account of the event. Though no attempt was made to use this transit to determine the length of the astronomical unit, Gassendi's observation was significant nonetheless, for it revealed that the apparent diameter of Mercury was far smaller than had been assumed by Kepler and his contemporaries [**7**]. Kepler, unfortunately, died on November 15, 1630, and thus did not live to see his brilliant prediction fulfilled.

Kepler's transit predictions were based on his *Rudolphine Tables* of 1627, which were produced as a result of his work with the great Danish astronomer Tycho Brahe. By the same method, Kepler also predicted the 1761 transit of Venus, but imperfections in the tables led him to believe that no transit would take place in 1639. Following Kepler's death, the Belgian astronomer Philip van Lansberg produced a set of tables, now known to be considerably inferior to Kepler's, but which did in fact predict a transit in 1639. It was in trying to reconcile differences between Lansberg's tables and the *Rudolphine Tables* that a brilliant young Englishman, Mr. Jeremiah Horrox, became convinced that a Venus transit would indeed occur in 1639 [**15**]. Regarding Lansberg's tables, Horrox wrote [**14**], "I pardon, in the meantime, the miserable arrogance of the Belgian astronomer, who has overloaded his useless tables with such unmerited praise... deeming it a sufficient reward that I was thereby led to consider and forsee the appearance of Venus in the Sun."

Horrox was richly rewarded for his labors in correcting the *Rudolphine Tables*, for on December 4, 1639, he became one of the first two people ever to observe a transit of Venus. The event was also observed at a nearby location by his friend William Crabtree, whom Horrox had alerted in the weeks preceding the event. No attempt was made by Horrox to use the Venus transit to determine the solar distance, but as with Gassendi's observations of Mercury in 1631, the event served to show that the angular size of Venus was far smaller than had been assumed [**13**].

Solar parallax

We now pause to present a few technical terms that are essential for the development of the story, using the terminology found in Taff [**12**]. In FIGURE 1, α is the angle between the line through the centers of the Earth and Sun and a line through the center

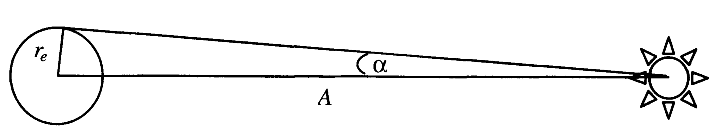


Figure 1 The equatorial horizontal solar parallax α

of the Sun and tangent to the Earth's surface. The angle α is known as the *horizontal solar parallax.* Denoting the Earth's equatorial radius by r_e and the Earth-Sun distance by A, it is clear from FIGURE 1 that $\alpha = \sin^{-1}(r_e/A)$. But of course A is not constant! If A is chosen to be one astronomical unit (au), the mean distance from Earth to the Sun, then α is known as the *mean equatorial horizontal solar parallax.* For brevity, the term *solar parallax* is commonly used in place of mean equatorial horizontal solar parallax, with further distinctions made only when the context requires it. With the preceding definitions and the understanding that r_e is known, it should be clear that *the problem of finding the length of the astronomical unit in terms of terrestrial units is equivalent to determining the solar parallax* α. Finally, to properly prepare the reader for the discussion that follows, we note that a modern value for the solar parallax is $8.794148''$ **[12]**, where a minute ($'$) is the sixtieth part of a degree and a second ($''$) is the sixtieth part of a minute. We also point out that *smaller* estimates of the solar parallax correspond to *larger* estimates of the astronomical unit, as expressed in miles (or kilometers, or Earth radii, or . . .).

Kepler's Earth-Sun distance of 3469 Earth radii corresponds to a solar parallax of about one minute. Through the course of the seventeenth century, estimates of the solar parallax continued to diminish, due in large part to a vast increase in the quality and quantity of telescopic observations of the planets. By the end of the century, leading astronomers had all begun to believe that the solar parallax was considerably less than one minute, though there was little uniformity and often less than compelling reasoning behind the variety of values that continued to appear in scholarly works. The uncertainty that remained early in the eighteenth century is nicely illustrated in van Helden's *Measuring the Universe* **[13]**, where we find that no less an authority than Newton was still undecided about the solar parallax: In the second edition of the *Principia* (1713), he used $10''$; in notes for the third edition he variously used $11''$, $12''$, and $13''$, and in the third edition itself one finds a solar parallax of $10\frac{1}{2}''$.

Edmond Halley's call for action: an international scientific effort

Though the idea of using a transit of Venus or Mercury to determine the solar parallax dates back at least to the Scottish mathematician James Gregory in 1663, it was Edmond Halley who became its greatest advocate. Halley observed a transit of Mercury from the southern hemisphere in 1677, and in his report on the observations, he discussed the possibility of using transits of Mercury or Venus to determine the solar parallax. Of the two, he believed that the geometry of Venus transits was far more likely to produce accurate results. Halley proposed the Venus transit idea in papers presented to the Royal Society in 1691, 1694, and most importantly, in 1716. Because Halley was one of the most influential astronomers of his time (he became the second Astronomer Royal in 1719), his paper of 1716 became "a clarion call for scientists everywhere to prepare for the rare opportunity presented by the forthcoming transits of 1761 and 1769." **[15]**

Halley's 1716 paper **[4]** begins by lamenting the wide variety of solar parallax values in use at the time, some as large as $15''$, and suggests $12\frac{1}{2}''$ as a plausible value. He goes on to describe roughly his method of determining the solar parallax from observations of the transit of Venus that would take place in 1761, even going so far as to describe the proper locations to send observers. "Therefore again and again," writes Halley, "I recommend it to the curious strenuously to apply themselves to this observation. By this means, the Sun's parallax may be discovered, to within its five hundredth part . . . " The essence of his method was to calculate, based on the $12\frac{1}{2}''$ hypothesis, the *expected* difference in the duration of the transit as observed at two widely differing locations. "And if this difference be found to be greater or less by observation, the Sun's parallax will be greater or less nearly in the same ratio." As we shall see, this was exactly the idea behind the methods employed by James Short when the transit actually took place. But despite the claim by Acker and Jaschek [**1**] that "this method was used by Halley in 1761 and 1769," Halley had no illusions that he would personally put his method into practice, for he died in 1742 at the age of 85.

Halley's paper called for observers to be stationed far and wide across the globe, a monumental task in 1761. Despite the obvious difficulties involved in sending observers to distant locations, not to mention the fact that Great Britain and France were in the midst of the Seven Years' War at the time, the response to his call was overwhelming. In all, when the transit took place, there were at least 122 observers at sixty-two separate stations, from Calcutta to the Siberian city of Tobolsk, from the Cape of Good Hope to St. John's in Newfoundland, and of course, at a large number of locations throughout Europe [**15**]. Many had traveled weeks or even months to reach their destinations. Unfortunately, it is impossible to describe in this short article all the adventures of those who set out to observe the 1761 transit: of Charles Mason and Jeremiah Dixon who set out for the East Indies, but hadn't so much as left the English channel when their ship was attacked by a French warship, leaving 11 dead and 37 wounded; of the Frenchman Chappe who traveled 1500 miles across Russia to Tobolsk by horse-drawn sleigh, once having to round up his deserting guides at gunpoint; of the Frenchman Le Gentil who was prevented by the war from reaching his destination in India, and so was forced to observe the transit from the rolling deck of a ship in the Indian Ocean. The interested reader will find excellent descriptions of these and other expeditions in Harry Woolf's book on the eighteenth-century transits of Venus [**15**]. All in all, the efforts to observe the 1761 transit of Venus surely amounted to the greatest international scientific collaboration in history up to that time.

James Short and his computation of the solar parallax

James Short (1710–1768) is not well known in modern mathematical circles for the simple reason that he was not primarily a mathematician. Though Short studied under Colin Maclaurin and displayed some talent in mathematics, he achieved fame and fortune as one of the most skilled telescope makers of the eighteenth century. In his lifetime, Short made some 1,370 telescopes, of which 110 still exist today [**3**]. A "Short biography" might also mention that he was a candidate for the post of Astronomer Royal, a frequent contributor to the *Philosophical Transactions of the Royal Society*, a friend of Benjamin Franklin, and a co-discoverer of a nonexistent moon of Venus [**3, 6**]. Short was a member of a special committee established by the Royal Society to plan the study of the 1769 transit of Venus, but died before the plan could be implemented.

Short observed the 1761 transit of Venus from London, in the company of the Duke of York and other honored guests. In the months following the transit, Short collected a good deal of data from the various observations that had taken place worldwide. These he published in the *Philosophical Transactions* in December 1761, in a paper entitled *The Observations of the internal Contact of Venus with the Sun's Limb, in the late Transit, made in different Places of Europe, compared with the Time of the Same Contact observed at the Cape of Good Hope, and the Parallax of the Sun from thence determined* [**10**]. A second article [**11**], virtually identical in nature but with a great deal more data, appeared a year later in an attempt to strengthen the case for his computed solar parallax value. We shall now examine the methods Short used, as described in the 1761 paper.

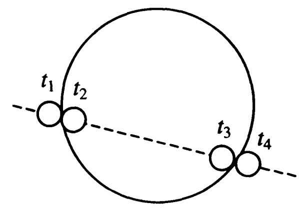


Figure 2 Contacts at ingress and egress

FIGURE 2 illustrates the positions of Venus on the disk of the Sun at four crucial times during the transit. Times t_1, t_2, t_3, and t_4 are the times of *external contact at ingress, internal contact at ingress, internal contact at egress*, and *external contact at egress*, respectively. Next, in FIGURE 3, one can see that the track of Venus across the Sun shifts *upward* as the observer moves further *south* on the surface of the Earth. This upward shift has two consequences that are crucial to Short's computational plans: first, the t_3 time is earlier for northern observers than for southern observers, and second, the total duration of the transit $t_3 - t_2$ is shorter for northern observers than for southern observers. Short's two methods simply amount to quantifying these two ideas.

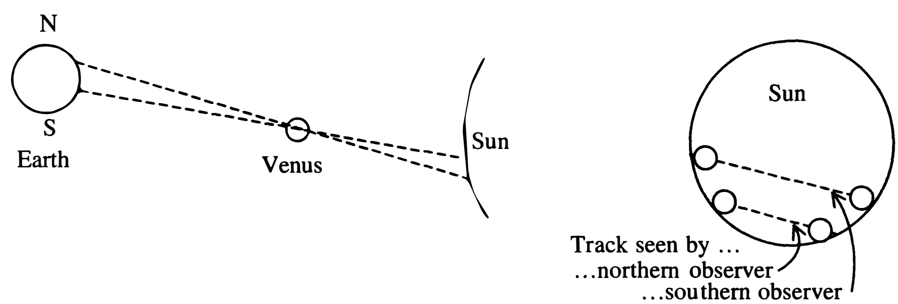


Figure 3 Effect of latitude on the apparent track of Venus

We can easily illustrate Short's first method just as he presents it in his first paper, that is, with virtually no computational details whatsoever! First, we note that the t_3 time observed in Greenwich was 8:19:00 AM local time, whereas the t_3 time as observed at the Cape of Good Hope was 9:39:50 local time. (In fact, it was Mason and Dixon who provided the valuable observations from the Cape, having been prevented from reaching the East Indies by their skirmish with the French warship.) The difference is $1^h\,20'\,50''$. Now most of this difference is due to the difference in local times, which Short determines to be $1^h\,13'\,35''$. Since one hour of local time difference corresponds to 15° of longitude, Short's figure is equivalent to saying that the Cape's longitude is 18°23′45″ east of Greenwich. But after the difference in local times is accounted for, a time difference of 7′ 15″ remains, which must be the difference due to the effect of latitude illustrated in FIGURE 3.

Next, Short asserts a theoretical difference to compare with this observed difference of 7′ 15″. *Assuming a solar parallax of* 8.5″ *on the day of the transit, the t_3 time for an observer at the Cape should be* 6′ 8″ *later than the t_3 time for a hypothetical observer at the center of the Earth, and the t_3 time for an observer at Greenwich should be* 1′ 11″ *earlier.* Thus the 8.5″ hypothesis leads to a difference of 7′ 19″ between the t_3 times predicted for these two stations. "But the difference in absolute time," Short writes, "as found by observation, as above, is only = 7′ 15″, therefore the Sun's parallax, by supposition, viz. 8.5″, is to the parallax of the Sun found by observation, as 7′ 19″ is to 7′ 15″, which gives 8.42″ for the Sun's parallax, on the day of the transit, by this observation..." In other words, after converting times to seconds, Short has solved the proportion

$$\frac{8.5}{\alpha} = \frac{439}{435}, \tag{1}$$

much as Halley had suggested.

In an identical manner, Short compares observations from fourteen other locations to those taken at the Cape, and concludes that "by taking a mean of the results of these fifteen observations, the parallax of the Sun, on the day of the transit, comes out = 8.47″, and by rejecting the 2d, the 8th, the 12th, and the 14th results, which differ the most from the rest, the Sun's parallax, on the day of the transit, by the mean of the eleven remaining ones is = 8.52″." He then uses this value to compute the mean equatorial horizontal solar parallax, which can be accomplished as follows. First, recall from FIGURE 1 that the radius of the Earth, which is of course constant, is $A \sin\alpha$. If A_t is the Earth-Sun distance (in au) on the day of the transit and α_m is the mean equatorial horizontal solar parallax (which corresponds to an Earth-Sun distance of $A = 1$ au), then

$$A_t \sin 8.52'' = \sin \alpha_m.$$

Clearly, Short knew that $A_t \approx 1.015$ au, allowing him to compute α_m, for he writes "The parallax of the Sun being thus found, by the observations of the internal contact at the egress, = 8.52″ on the day of the transit, the mean (equatorial) horizontal parallax of the Sun is = 8.65″." Thus the solar parallax computation is complete. The length of the astronomical unit in miles is now simply $r_e/\sin 8.65''$, where r_e is the radius of the Earth in miles.

But there is a gaping hole in our understanding of Short's method. To complete our understanding, we must develop a way to determine the 6′ 8″ and 1′ 11″ time values noted above (and similar values for other observer locations), which arise from the hypothesis of an 8.5″ solar parallax on the day of the transit. We shall approach the problem in a manner that is undoubtedly different from what Short used in 1761, preferring to use the tools of vector and matrix algebra that are so familiar to us.

Our first task is to develop two coordinate systems and relate them to one another. FIGURE 4 shows the *geocentric equatorial* coordinate system $x'y'z'$ whose origin is at the center of the Earth. The $x'y'$ plane contains the Earth's equator, and the z' axis passes through the north pole. The positive x' axis is oriented so that it passes through the center of the Sun on the first day of spring, and is fixed in space; that is, the Earth's daily motion and annual motion do not change the orientation, but only the location of the origin. Thus the angle θ in FIGURE 4 changes continuously as the Earth rotates.

Now consider an observer at longitude λ and latitude β, measured with the convention that $-180° < \lambda \le 180°$ and $-90° \le \beta \le 90°$, with $\lambda > 0$ east of Greenwich and $\beta > 0$ north of the equator. If θ represents the angular position of Greenwich with respect to the x' axis at a particular instant, then an observer at longitude λ and latitude β will have $x'y'z'$ coordinates

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} r_e \cos\beta \cos(\theta+\lambda) \\ r_e \cos\beta \sin(\theta+\lambda) \\ r_e \sin\beta \end{pmatrix}. \tag{2}$$

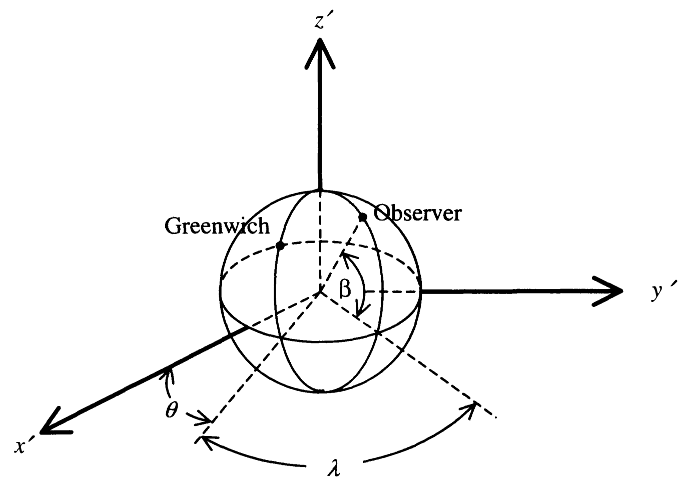


Figure 4 Geocentric equatorial coordinates

This is just the usual spherical-to-rectangular coordinate conversion, with the observation that latitudes are measured up from the equator rather than down from the north pole, as is standard in calculus texts.

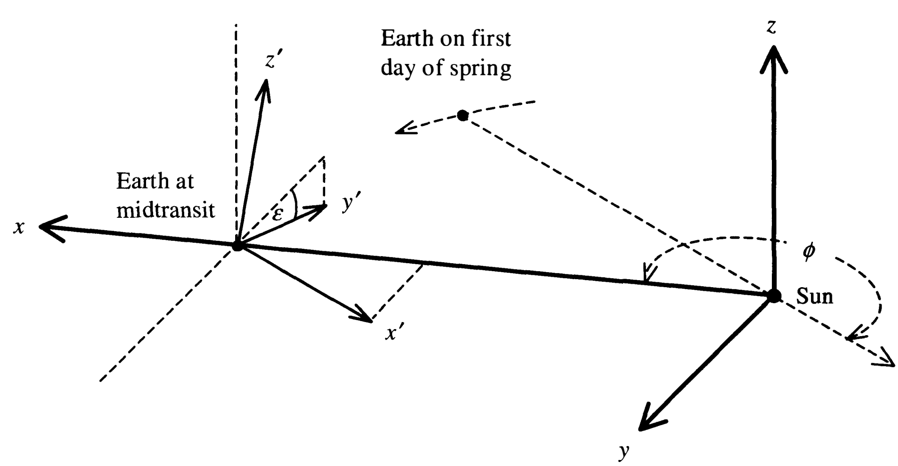


Figure 5 Geocentric equatorial and heliocentric Venus transit coordinates

FIGURE 5 shows the $x'y'z'$ coordinate system along with a second system xyz that we shall call the *heliocentric Venus transit* coordinate system. The xyz coordinate system is chosen with the origin at the center of the Sun, the xy plane containing the Earth's orbit, and the positive x axis directed through the center of the Earth *at midtransit as viewed from the center of the Earth.* The xy plane is known as the ***ecliptic*** plane.

Converting from one coordinate system to the other is accomplished as follows: first, a rotation about the x' axis through the angle ϵ (the tilt of the Earth's axis with respect to the ecliptic plane) makes the $x'y'$ plane coincide with the xy plane; next, a rotation about the z' axis (now pointing in the same direction as the z axis) through an angle of ϕ makes the x' axis coincide with the x axis; finally, the $x'y'z'$ coordinate system is translated one unit in the positive x direction. The angle ϕ represents the Earth's position with respect to the Sun at midtransit June 6, 1761, and placing the center of the Earth exactly one unit from the center of the Sun at midtransit is simply a computational convenience. (This distance unit, which we shall use to measure *all* distances in the following discussion, is approximately one astronomical unit, but not exactly so because the Earth is not at its mean distance from the Sun on June 6.) The rotations and translation are accomplished via

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\epsilon & \sin\epsilon \\ 0 & -\sin\epsilon & \cos\epsilon \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{3}$$

The Earth's daily motion and annual motion must be accounted for as time elapses from $t = 0$, the moment of midtransit (as viewed from the center of the Earth). For the former, we increase θ in (2) according to $\theta = \theta_0 + 15t$ (degrees), where θ_0 is the position of Greenwich (with respect to the positive x' axis) at $t = 0$ and t is measured in hours. Of course the 15 arises from the fact that the Earth rotates 15° per hour. We shall approximate the Earth's annual motion for the short duration of the transit by assuming that it takes place entirely in the positive y direction. Denoting the Earth's angular velocity at the time of the transit by ω_e, the displacement at time t due to the Earth's annual motion is approximated by $[0, \omega_e t, 0]^T$. By adding this displacement to the right side of (3) and by using (2) to determine $[x', y', z']^T$, we can determine the xyz coordinates at time t of an observer at longitude λ and latitude β.

The essence of our method is to derive vector equations based on the simple observation that the center of the Earth, the center of Venus, and the center of Venus's image on the Sun (as viewed from the center of the Earth) must be collinear. We shall then repeat the computation, replacing the center of the Earth with the position of an observer on the surface of the Earth. FIGURE 6 shows the track of Venus's image across the disk of the Sun, as viewed from the center of the Earth. At $t = 0$ (midtransit), the center of Venus's image on the Sun (in xyz coordinates) is at $\mathbf{I}_0 = [0, d\cos u, d\sin u]^T$, where d and u will be computed from observations. The center of the Earth is at $\mathbf{E}_0 = [1, 0, 0]^T$. Assuming that, for the short duration of the transit, the motion of Venus takes place in the plane $x = x_v$, simple vector addition shows that the position of the center of Venus at time $t = 0$ is $\mathbf{V}_0 = \mathbf{I}_0 + x_v(\mathbf{E}_0 - \mathbf{I}_0)$.

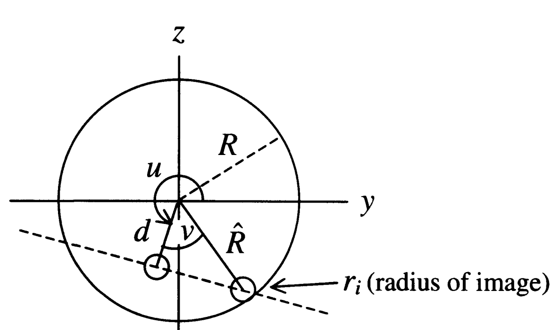


Figure 6 Track of Venus's image on the disk of the Sun

FIGURE 6 shows that at time $t = T$ (the moment of internal contact at egress as viewed from the center of the Earth), the center of Venus's image is at $\mathbf{I}_T = [0, \hat{R}\cos(u + v), \hat{R}\sin(u + v)]^T$. Meanwhile, the center of the Earth has moved to

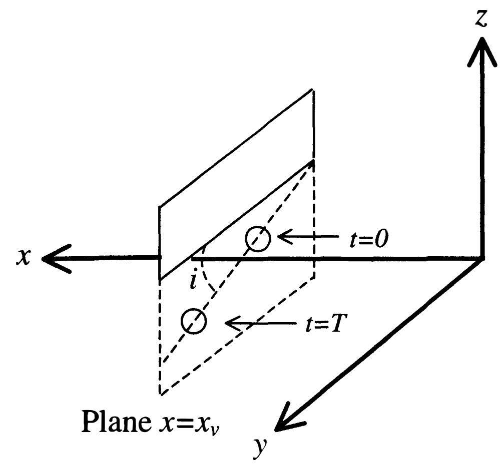


Figure 7 Motion of Venus in the plane $x = x_v$

$\mathbf{E}_T = [1, \omega_e T, 0]^T$. In FIGURE 7, which shows the motion of Venus from $t = 0$ to $t = T$, the angle i is the inclination of Venus's orbit to the xy plane. If ω_v is the angular velocity of Venus, then the distance the planet travels from $t = 0$ to $t = T$ is approximated by $x_v \omega_v T$, so that its location at time T is

$$\mathbf{V}_T = \mathbf{V}_0 + x_v \omega_v T \begin{pmatrix} 0 \\ \cos i \\ -\sin i \end{pmatrix}. \tag{4}$$

The vector $\mathbf{V}_T - \mathbf{I}_T$ must be a constant multiple of the vector $\mathbf{E}_T - \mathbf{I}_T$ in order for the center of the Earth, the center of Venus, and the center of Venus's image on the Sun to be collinear, and the first coordinates tell us that the constant is x_v. By equating the second coordinates in the vector equation $\mathbf{V}_T - \mathbf{I}_T = x_v(\mathbf{E}_T - \mathbf{I}_T)$, expanding $\cos(u + v)$ and simplifying, we are able to obtain

$$\sin u = \frac{x_v(\omega_e - \omega_v \cos i)}{(1 - x_v)\sqrt{\hat{R}^2 - d^2}} T. \tag{5}$$

Likewise, equating the third coordinates, expanding $\sin(u + v)$ and simplifying, we get

$$\cos u = -\frac{x_v \omega_v \sin i}{(1 - x_v)\sqrt{\hat{R}^2 - d^2}} T. \tag{6}$$

The identity $\sin^2 u + \cos^2 u = 1$ allows us to obtain

$$T = \frac{(1 - x_v)}{x_v} \sqrt{\frac{\hat{R}^2 - d^2}{\omega_v^2 + \omega_e^2 - 2\omega_v \omega_e \cos i}}. \tag{7}$$

Once T is known, equations (5) and (6) allow us to determine $\sin u$ and $\cos u$.

Next we repeat the computation for a viewer on the Earth's surface at longitude λ and latitude β. Let $t = T_0$ be the time at which our observer sees the internal contact at egress, where once again $t = 0$ refers to the moment of midtransit as seen from the center of the Earth. Let $\mathbf{I}_{T_0} = [0, y_c, z_c]^T$ designate the center of the image as seen by the observer at time $t = T_0$, and note that $y_c^2 + z_c^2 = \hat{R}^2$. As noted in the paragraph following (3), the observer's position at $t = T_0$ is

$$\mathbf{O}_{T_0} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ \omega_e T_0 \\ 0 \end{pmatrix},$$

where the vector $[x, y, z]^T$ is exactly as given in (3). By changing T to T_0 in (4), we obtain the position of the center of Venus at time $t = T_0$, that is

$$\mathbf{V}_{T_0} = \mathbf{V}_0 + x_v \omega_v T_0 \begin{pmatrix} 0 \\ \cos i \\ -\sin i \end{pmatrix}.$$

As before, the vector $\mathbf{V}_{T_0} - \mathbf{I}_{T_0}$ is a constant multiple of the vector $\mathbf{O}_{T_0} - \mathbf{I}_{T_0}$, and the first coordinates force the constant to be x_v/x. Equating the second coordinates gives us $y_c = M + NT_0$, where

$$M = \frac{(1 - x_v)xd \cos u - x_v y}{x - x_v} \quad \text{and} \quad N = \frac{x x_v \omega_v \cos i - x_v \omega_e}{x - x_v},$$

and equating the third coordinates gives us $z_c = P - QT_0$, where

$$P = \frac{(1 - x_v)xd \sin u - x_v z}{x - x_v} \quad \text{and} \quad Q = \frac{x x_v \omega_v \sin i}{x - x_v}.$$

Then $y_c^2 + z_c^2 = \hat{R}^2$ becomes $(M + NT_0)^2 + (P - QT_0)^2 = \hat{R}^2$, which can easily be solved for T_0. (The $+$ sign in the quadratic formula gives the correct root.)

Our observer at longitude λ and latitude β should see the internal contact at egress $T_0 - T$ hours later than a hypothetical observer at the center of the Earth, assuming the difference is positive, and $|T_0 - T|$ hours earlier if the difference is negative. To illustrate, we restate the values given previously: for an observer at the Cape of Good Hope, $T_0 - T = 0.10222$ hours, or 6′ 8″ later, whereas for an observer at Greenwich, $T_0 - T = -0.01972$ hours, or 1′ 11″ earlier.

We now address the problem of determining values for the many parameters that have been introduced into our work. There are three sources for values of these parameters: First, Short specifically lists a few of the values that he uses; second, and most importantly, many of the values can be computed from Short's values using the underlying hypothesis that the solar parallax is 8.5″ on the day of the transit; third, there are a few values related to the Earth's daily and annual motions that Short undoubtedly knew, but did not specify. For these last-mentioned values, we resort to modern sources that readily supply the necessary information. In all cases, we shall use Short's values for the latitudes and longitudes of the observers as given in Short's first paper, for accurate determination of the longitude was a significant problem in 1761, and the use of modern values would seriously affect the results. In using Short's longitude values, one must be careful to note that not all are measured with respect to Greenwich, a standard that evolved sometime after 1761.

Let us reiterate that our coordinate system is chosen so that the centers of the Earth and Sun are exactly one unit apart at midtransit, as seen from the center of the Earth. All distances in the following work will be measured in terms of this unit. Under the 8.5″ hypothesis, the radius of the Earth is therefore $r_e = \sin 8.5''$. Short gives the difference in the parallaxes of Venus and the Sun as 21.35″ on the day of the transit, so we may take 29.85″ as the parallax of Venus. Therefore the Venus-to-Earth distance is given by $r_e/\sin 29.85'' = \sin 8.5''/\sin 29.85''$, so that $x_v = 1 - \sin 8.5''/\sin 29.85''$.

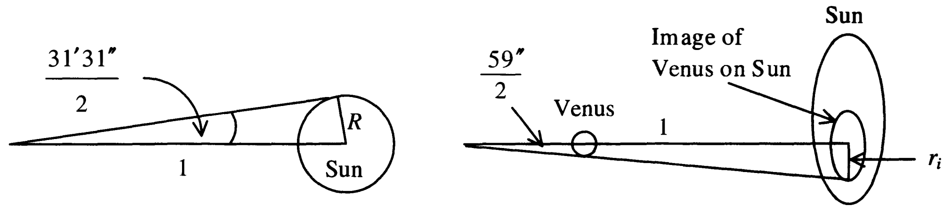


Figure 8 The radii of the Sun and the image of Venus from their angular diameters

Next, Short's value of 31′ 31″ for the angular diameter of the Sun on the day of the observation gives us a value for the solar radius $R = \sin(31'31''/2)$; likewise his value of 59″ for the angular diameter of Venus gives us the value $r_i = \tan(59''/2)$ (see FIGURES 6 and 8). Thus $\hat{R} = R - r_i$ is known. Short also gives the minimum angular separation of the centers of Venus and the Sun as 9′ 32″, as seen from the center of the Earth. This figure, based on the actual transit observation, gives us $d = \tan(9'32'')$. And last, Short gives us two values regarding the motion of Venus, the first of which is $\omega_v = 3'\,59.8''$ of arc per hour. The second is i, the inclination of Venus's orbit with respect to the plane of the Earth's orbit. The value published in Short's paper **[10]** is $i = 8^\circ\,30'\,10''$, *a value that is surely the result of a typesetting error.* For this value produces nonsensical results, whereas the value $i = 3^\circ\,30'\,10''$ not only agrees well with the modern value **[9]** but produces results that match Short's quite well. It is inconceivable that the best data available in 1761 had a 5° error in the inclination.

For the parameters that Short omits from his paper, modern references by Montenbruck and Pfleger **[8]** and Roy **[9]** provide us with the appropriate 1761 values. For the Earth's daily and annual motions, we have used $\epsilon = 23.47^\circ$, $\phi = 256^\circ$, $\theta_0 = -25^\circ$, and $\omega_e = 2'\,25.0''$ per hour. To obtain θ from θ_0, we have assumed the constant t value of 3 hours, which approximates the semi-transit time and which therefore allows us to determine the observer's xyz coordinates at the moment of internal contact at egress. Computational experience suggests that the calculations are quite sensitive to changes in ω_e, but much less so for ϵ, ϕ, and θ.

The lack of certainty as to the exact values James Short used for i, ϵ, ϕ, θ_0, and ω_e, the sensitivity of the computations to ω_e, and our rather different method of computing the $T_0 - T$ values make it impossible to match Short's values exactly. But the results are consistently close, differing from Short's by roughly 1%. Thus the $T_0 - T$ value for the Cape of Good Hope, computed by the above method, is 6′ 12″, compared to Short's value of 6′ 8″. The table below lists data for four other locations, the last being a location in present day Finland. The second column shows our computed $T_0 - T$ value followed by Short's value in parentheses. The third column is the difference between the Cape and the given location, again with Short's value in parentheses. The fourth column shows the difference in time of internal contact at egress between the given location and the Cape, as actually reported by the observers. The last column shows the resulting solar parallax on the day of the transit computed as in (1), with Short's value in parentheses.

Location	$T_0 - T$	Diff. from Cape	Observed	Parallax
Greenwich	−1′ 12″ (−1′ 11″)	7′ 24″ (7′ 19″)	7′ 15″	8.33″ (8.42″)
Rome	−0′ 14″ (−0′ 13″)	6′ 26″ (6′ 21″)	6′ 26″	8.50″ (8.61″)
Stockholm	−2′ 20″ (−2′ 18″)	8′ 32″ (8′ 26″)	8′ 25″	8.38″ (8.48″)
Cajaneburg	−3′ 1″ (−2′ 59″)	9′ 13″ (9′ 7″)	8′ 56″	8.24″ (8.33″)

For further comparison, we note that Short's computed values for the solar parallax on the day of the transit fall between 8.07″ and 8.86″, based on data from the fifteen sites studied in his first paper [**10**].

Short's second method

Short's second method for computing the solar parallax is really just a small variation on the first method, and so it will be very easy for us to describe. This method compares the observed *duration* of the transit, that is, the time between internal contact at ingress and internal contact at egress, to the theoretical duration of the transit computed from the 8.5″ hypothesis. In the notation of the previous section, the theoretical duration for an observer at the center of the Earth is simply $2T$, which can readily be computed from (7). By this method, the duration is $5^h\,57'\,59''$, whereas Short gives the value $5^h\,58'\,1''$. For an observer on the surface of the Earth, one can compute a theoretical duration from the 8.5″ hypothesis as follows. First, compute T_0 (the time between midtransit and internal contact at egress) just as before. Next, by evaluating $\theta = \theta_0 + 15t$ at $t = -3$ instead of $t = 3$, and by using the $-$ sign in the quadratic formula used to determine T_0, we obtain the time *before* midtransit at which the internal contact at ingress should occur. Subtracting this (negative) value from the original T_0 value gives us the theoretical duration of the transit for our observer. We can then use a proportion much like (1) to reconcile the actual observed duration with this theoretical duration.

For example, Short's theoretical duration for Tobolsk is $5^h\,48'\,58''$, which differs from his center-of-the-Earth duration by $9'\,3''$ or 543 seconds. But the observed duration at Tobolsk was $5^h\,48'\,50''$, which differs from his center-of-the-Earth duration by $9'\,11''$ or 551 seconds. Then

$$\frac{8.5}{\alpha} = \frac{\text{center-of-Earth} - \text{theoretical}}{\text{center-of-Earth} - \text{observed}} = \frac{543}{551},$$

yielding a solar parallax of 8.63″ on the day of the transit. To illustrate further, our theoretical duration for Cajaneburg is $5^h\,49'\,54''$, Short's is $5^h\,49'\,56''$, and the observed duration was $5^h\,49'\,54''$. For Stockholm, our theoretical value is $5^h\,50'\,27''$, Short's is $5^h\,50'\,27''$, and the two reported observations are $5^h\,50'\,45''$ and $5^h\,50'\,42''$.

Short uses this second method to compute the solar parallax on the day of the transit using data from sixteen different observers. He thus obtains sixteen values ranging from 8.03″ to 8.98″, with a mean of 8.48″. In a manner that modern statisticians can only envy for its simplicity, he concludes that "if we reject the observations of number 7th, 8th, 9th, 10th, 12th, 13th, and 14th, which differ the most from the rest, the mean of the nine remaining ones gives the Sun's parallax = 8.55", agreeing, to a surprising exactness, with that found by the observations of the internal contact at the egress." If half of the data doesn't support your conclusion, just use the other half!

The transit of 1769 and conclusions

The reader may have sensed by now that the transit of 1761 did not produce the definitive result that Halley had predicted in his 1716 paper. Despite the cleverness of the method and the extraordinary efforts that had gone into making the observations, the conclusions that were drawn from the data still varied widely, with much of the uncertainty due to the lack of accurate longitude data [**15**]. But the experience gained in 1761 only served to whet the appetites and improve the skills of those who would follow in 1769, when an equally vast international effort was undertaken to observe the second Venus transit of the decade. Captain James Cook was hired to transport observers to the South Pacific, and our friend Chappe (of Tobolsk in Siberia) observed the transit from Baja California, where he died soon thereafter. The luckless Le Gentil, who missed the 1761 transit as a result of the war, waited eight years in the Indian Ocean area for the 1769 transit, only to be defeated by cloudy weather. And how does Euler fit into the story? He certainly did not witness the 1769 transit, for by then he was totally blind, but this did not stop him from writing about it [**2**]. Other than these few items of trivia, we shall not go into the 1769 transit in any detail, for the mathematics did not change significantly from the work already described.

The range of solar parallax values derived from the 1769 transit, and thus the length of the astronomical unit, drew ever closer to the values accepted today. We close by providing the details of a comparison that was mentioned in the introduction: a modern radar-based value for the astronomical unit is 92,955,000 miles [**9**]. And based on his analysis of the 1769 transit of Venus, Thomas Hornsby [**5**] wrote in 1771 that "The parallax on the 3d of June being $8.65''$, the mean parallax will be found to be $= 8.78''$; and if the semidiameter of the Earth be supposed $= 3985$ English miles, the mean distance of the Earth from the Sun will be 93,726,900 English miles."

Eight-tenths of a percent difference. Absolutely remarkable.

Notes on the sources A very large number of papers on the transits of Venus in 1761 and 1769 appeared in the *Philosophical Transactions of the Royal Society*. Thanks to a project known as the Internet Library of Early Journals (ILEJ), many of these [**5, 10, 11**] are available at `www.bodley.ox.ac.uk/ilej/`. (Hint: sometimes the "next page" arrows at this site don't work, but adjusting the page number in the URL does.) Also, Halley's paper [**4**] can be found online, starting at `www.dsellers.demon.co.uk/`.

One can find a number of web sites devoted to the upcoming Venus transit of June 8, 2004. In particular, a web site maintained by the U.S. Naval Observatory (`http://aa.usno.navy.mil/data/docs/Venus2004.pdf`) gives precise information about the times of the transit predicted for various locations all over the world. According to this web site, the entire transit will be visible throughout most of Europe, Asia, and Africa. The very end of the transit will be visible in the eastern U.S. just after sunrise.

REFERENCES

1. A. Acker and C. Jaschek, *Astronomical Methods and Calculations*, John Wiley and Sons, Chichester, 1986.
2. L. Euler, A deduction of the quantity of the sun's parallax from the comparison of the several observations of the late transit of Venus, made in Europe, with those made in George Island in the South Seas, *Philosophical Transactions of the Royal Society* **62** (1772), pp. 69–76.
3. C. Gillispie, editor, *Dictionary of Scientific Biography*, Scribner, New York, 1980.
4. E. Halley, A new method of determining the parallax of the sun, or his distance from the earth, *Philosophical Transactions of the Royal Society, Abridged* **VI** (1809), pp. 243–249.
5. T. Hornsby, The quantity of the sun's parallax, as deduced from the observations of the transit of Venus, on June 3, 1769, *Philosophical Transactions of the Royal Society* **61** (1771), 575–579.
6. W. Ley, *Watchers of the Skies*, Viking Press, New York, 1969.
7. E. Maor, *June 8, 2004: Venus In Transit*, Princeton University Press, Princeton, New Jersey, 2000.
8. O. Montenbruck and T. Pfleger, *Astronomy on the Personal Computer*, Springer-Verlag, New York, 1991.
9. A. E. Roy, *Orbital Motion*, 3rd ed., Adam Hilger, Ltd., Bristol, UK, 1988.

J. Short, The observations of the internal contact of Venus with the sun's limb, in the late transit, made in different places of Europe, compared with the time of the same contact observed at the Cape of Good Hope, and the parallax of the sun from thence determined, *Philosophical Transactions of the Royal Society* **52** (1761), 611–628.

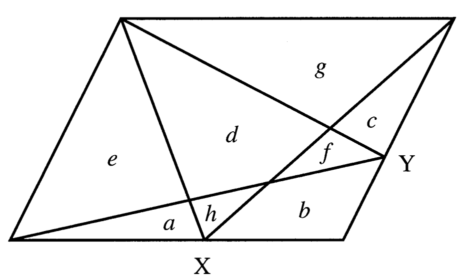
J. Short, Second paper concerning the parallax of the sun determined from the observations of the late transit of Venus, in which this subject is treated more at length, and the quantity of the parallax more fully ascertained, *Philosophical Transactions of the Royal Society* **53** (1763), 300–345.

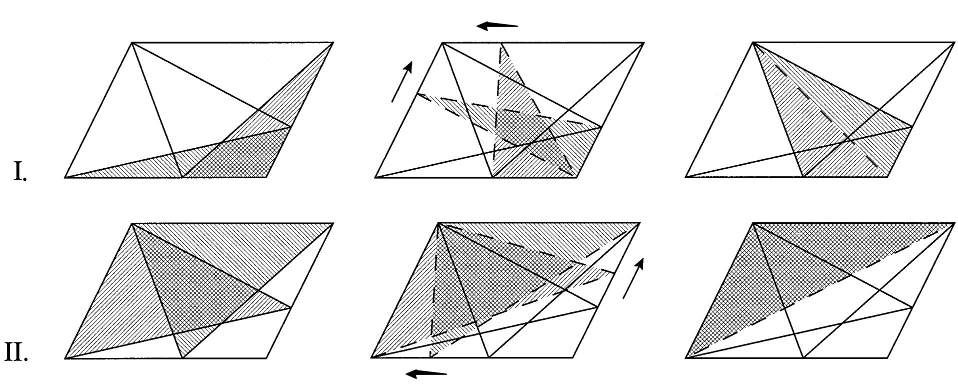
12. L. G. Taff, *Celestial Mechanics: A Computational Guide for the Practitioner*, Wiley, New York, 1985.
13. A. Van Helden, *Measuring the Universe*, University of Chicago Press, Chicago, 1985.
14. A. B. Whatton, *The Transit of Venus Across the Sun... by the Rev. Jeremiah Horrox... to which is prefixed A Memoir of his Life and Labours*, W. Macintosh, London, 1859.
15. H. Woolf, *The Transits of Venus: A Study of Eighteenth Century Science*, Princeton University Press, Princeton, New Jersey, 1959.

Proof Without Words: Equal Areas in a Partition of a Parallelogram

I. $a+b+c=d$

II. $e+f=g+h$





—PHILIPPE R. RICHARD
UNIVERSITÉ DE MONTRÉAL
QUÉBEC, CANADA

Fits and Covers

JOHN E. WETZEL
University of Illinois at Urbana-Champaign
Urbana, IL 61801-2975
j-wetzel@uiuc.edu

Combinatorial Euclidean geometry is replete with appealing, uninvestigated questions of unpredictable depth and difficulty. Interesting, innocent-looking, beguilingly simple questions having to do with the ways shapes fit together, they are fascinating problems to tinker with, and, well, you never know....

The situation that interests us here has to do with fitting one geometric figure into another. A *figure* Φ and a *target set* Γ are given.

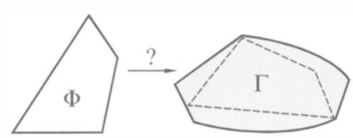


Figure 1 Does Φ fit in Γ?

We say that the figure Φ *fits* in the target Γ, or, equivalently, the target Γ *covers* the figure Φ, when there is a rigid motion μ so that $\mu(\Phi) \subseteq \Gamma$, that is, when Γ has a subset *congruent* to Φ (FIGURE 1). We assume throughout that the figure Φ is compact and the target Γ is closed, convex, and (usually) compact.

It might happen that a specific figure Φ and target Γ are given, and one wants to know if the figure fits in the target, or perhaps how large a figure similar to Φ fits in the target Γ. We call such problems *fitting problems*, or *fits*.

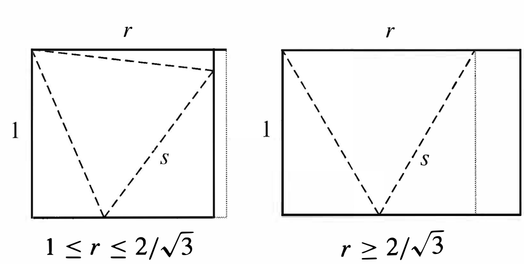


Figure 2 Fitting an equilateral triangle in a rectangle

For example, one might want to know how large an equilateral triangle can fit in a $1 \times r$ rectangle ($r \geq 1$). The answer depends on r, of course. If $1 \leq r \leq 2/\sqrt{3}$ the largest such equilateral triangle has side $2\sqrt{r^2 - \sqrt{3}r + 1}$, and if $r \geq 2/\sqrt{3}$ it has side $2/\sqrt{3}$ (see FIGURE 2). In particular, as is well known, the largest equilateral triangle that fits in a unit square has side $2\sqrt{2 - \sqrt{3}} = \sqrt{6} - \sqrt{2}$. See, for example, Madachy [**36**, pp. 116, 129–131].

Alternatively, a family of figures might be given, and one seeks a target set, perhaps of prescribed shape, that covers every figure in the family and is small in some specified sense. We refer to questions of this kind as *covering problems,* or *covers.*

For example, one might ask for the smallest square region large enough to contain a congruent copy of every triangle with perimeter two. (The smallest such square has side $1/\sqrt{2}$. See Wetzel [**52**].)

Our intention here is to examine a few typical problems, solved and unsolved, of these two kinds.

When the target Γ is convex, the questions and methods are geometric in nature; but when the target Γ is not assumed convex, the questions of interest are fractal or measure theoretic in nature, and they typically lie much deeper. In the last section of this survey we cite some of the surprising results that are known in the nonconvex case.

Fits

Suppose a pair (Φ, Γ) of compact, convex shapes is given, each shape having nonempty interior. There is a largest figure $\Phi_{\max}$ similar to Φ that fits in Γ, and Φ itself fits in Γ if and only if it is no larger than $\Phi_{\max}$. Dually, there is a smallest target $\Gamma_{\min}$ similar to Γ into which Φ fits, and Φ itself fits in Γ if and only if Γ is no smaller than $\Gamma_{\min}$. Finding a necessary and sufficient condition for Φ to fit in Γ is equivalent to finding the largest figure $\Phi_{\max}$ similar to Φ that fits in Γ and also to finding the smallest target $\Gamma_{\min}$ similar to Γ that covers Φ.

Another pair of problems arises from interchanging the roles of figure and target. We could ask for the largest figure similar to Γ that fits in the given target Φ, or, equivalently, the smallest target similar to Φ that contains the figure Γ. We call these (Γ, Φ) questions *converse* to the (Φ, Γ) questions. So, associated with any pair of compact, convex shapes (with nonempty interiors) there are four fit problems, which are equivalent in pairs. There seems to be little relationship between a fit problem and its converse.

Triangles and disks

Here and throughout we employ the standard notations for the parts of triangle ABC. We write a, b, and c for both the segments BC, CA, and AB and their lengths, and α, β, and γ for the (measures of the) opposite angles.

Euclid surely knew that the incircle is the largest circle that fits in a given triangle even though this fact is not mentioned in his *Elements* (see Heath [**24**, IV, Prop. 4]). It follows that a disk of diameter d fits in the triangle T with sides a, b, c precisely when $d \leq 2r$, where r is the inradius of T. An expression for r in terms of the sides follows from Heron's formula for the area F combined with the observation that $r(a+b+c) = 2F$, and we conclude that a disk of diameter d fits in a triangle with sides a, b, c if and only if

$$d \leq \sqrt{\frac{(-a+b+c)(a-b+c)(a+b-c)}{a+b+c}}.$$

Dually, the smallest triangle similar to T that contains the disk of diameter d is the one with inradius $d/2$.

In the 1870s Cayley [**8**] posed the *converse* problem in the *Educational Times,* asking for a necessary and sufficient condition on a, b, c, d for a triangle with sides a, b, c to fit in a disk with diameter d.

EXERCISE 1. *Find a necessary and sufficient condition on a, b, c, d for the triangle T with sides a, b, c $(a \geq b \geq c)$ to fit into the disk D of diameter d. What is the* *dual result? (Suggestion: Use the law of sines, $1/(2R) = \sin\alpha/a$, with Heron's formula to express the circumradius R in terms of the sides of T.)* [Answer on page 398.]

Are there analogous results for spheres and tetrahedra?

Squares and triangles

In 1836, Ernst Brune, a mid-level Prussian civil servant in the Prussian General Widows Pension Office, found the largest square that fits in an acute or right triangle [**6**]. Despite the title, "Grösstes Quadrat im Dreiecke," Brune actually found the largest square that is *inscribed* in the given triangle in the classical sense that all four corners of the square must all lie on the *sides* of the given triangle.

For nonobtuse triangles, Brune showed by synthetic means that the largest inscribed square rests on the shortest side of the triangle. If the triangle has sides a, b, c with $a \geq b \geq c$, then the largest square rests on the shortest side c, and its side $s_{\max}$ is easily seen from similar triangles to be $s_{\max} = ch_c/(c + h_c) = (c^{-1} + h_c^{-1})^{-1}$, half the harmonic mean of the shortest side c and the longest altitude h_c. It follows that a square of side s fits into the triangle if and only if $s \leq s_{\max} = (c^{-1} + h_c^{-1})^{-1}$. As far as I am aware, this is the earliest question about fits in the literature, although there may well be earlier examples.

For the obtuse case, Brune noted that there is only one square inscribed in the triangle, so it is, *a fortiori*, the largest.

Sullivan [**51**] has recently noted the very useful fact that if a convex polygon fits in a triangle, then it must always also fit in such a way that one of its sides lies along a side of the containing triangle (indeed, Sullivan showed that the polygon can always be moved continuously within the triangle so that one of its sides is brought to rest on one side of the containing triangle). It follows that in the obtuse case, one side of the largest square rests on one side of the triangle, but *that side isn't always the longest side of that triangle.* The precise result is complicated because an unanticipated cubic condition appears.

THEOREM 1. (WETZEL [**58**]) *An obtuse triangle T is given, with sides a, b, and c, where $a > b \geq c$. Then the side $s_{\max}$ of the largest square that fits in T is*

$$s_{\max} = \begin{cases} \dfrac{h_a}{\sin\gamma + \cos\gamma} & \text{if } h_a \geq a\sqrt{2}\sin(\gamma + 45^\circ) \\ \dfrac{ah_a}{a + h_a} & \text{if } h_a < a\sqrt{2}\sin(\gamma + 45^\circ). \end{cases}$$

A square of side s fits in T precisely when $s \leq s_{\max}$.

The argument is straightforward using the easily established formulas

$$s_a = \frac{ah_a}{a + h_a}, \qquad s_b = \frac{h_a}{\sin\gamma + \cos\gamma}, \qquad s_c = \frac{h_a}{\sin\beta + \cos\beta}$$

for the sides of the largest squares in T with one side on a, b, c, respectively, of T.

The boundary between the two cases in the Theorem is given by equality $h_a = a\sqrt{2}\sin(\gamma + 45^\circ)$, which gives the locus of points A for which the largest squares on the longest side BC and the middle side AC have equal sides. In a cartesian coordinate system in which $C = (0, 0)$, $B = (-a, 0)$, and $A = (x, y)$ with $y > 0$ (FIGURE 3) this constraint is the cubic curve whose equation is

$$y^3 + x^2y + 2ax^2 + 2ay^2 + 2a^2x = 0,$$

which is asymptotic to the line $y = -2a$ and whose intercepts are the points $(-a, 0)$, $(0, 0)$, and $(0, -2a)$. Since T is obtuse, the notational normalization $a > b \geq c$ forces A to lie in the (upper) semicircle with diameter BC and to the left of the perpendicular bisector of BC. If A lies above the cubic (in the region lightly shaded in FIGURE 3), then the largest square in $T = ABC$ rests on the middle side AC, but if A lies below the cubic (in the region more heavily shaded), then the largest square in T rests on the longest side BC. This result can be reformulated in terms of just the sides a, b, c using the familiar formulas of trigonometry, but the results are ugly.

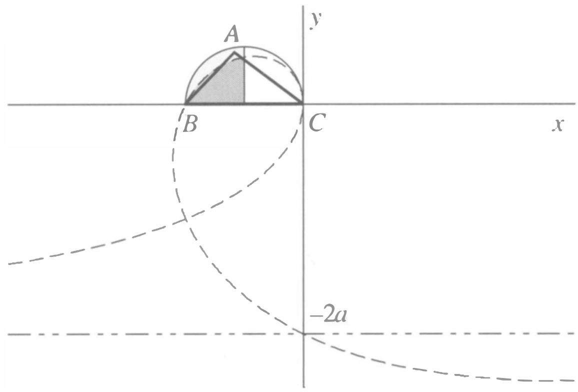


Figure 3 The obtuse case

EXERCISE 2. *A square S of side s and a triangle T_0 are given. Describe the smallest triangle T similar to T_0 that can cover S.*

EXERCISE 3. (CALABI) *Show that there exists exactly one nonequilateral triangle for which the three maximal squares, one resting on each side, have equal areas.*

In the introduction, we mentioned the familiar fact that the largest equilateral triangle that fits in the unit square has side $\sqrt{6} - \sqrt{2}$. The problem of finding the largest triangle of any given shape that fits in a given square has recently been studied (Jerrard and Wetzel [**27**]). To my knowledge, nothing is known about the analogous problems for cubes and tetrahedra in $\mathbb{R}^3$ or, more generally, for r-simplices and d-cubes in $\mathbb{R}^d$.

Triangle in triangle

Natural fit problems need not have nice answers. A fundamental fit question was posed in 1964 by H. Steinhaus [**48**], who asked for a necessary and sufficient condition on the six sides a, b, c, a', b', c' for the triangle T' with sides a', b', c' to fit in the triangle T with sides a, b, c. Separate necessary conditions and sufficient conditions are easy to give, but a condition that is both necessary and sufficient is more elusive. In 1993 the Dutch mathematician Karel Post, motivated, he wrote, "by a problem that arose in industry," gave a list of 18 inequalities whose disjunction is both necessary and sufficient.

THEOREM 2. (POST [**44**]) *Write F, F' for the areas of T, T', respectively. Then T' fits in T if and only if one of the eighteen inequalities obtained by permuting a, b, c and cyclically permuting a', b', c' in the following inequality is correct:*

$$\begin{aligned} &\max\left\{F\left(b'^2 + c'^2 - a'^2\right), F'\left(b^2 + c^2 - a^2\right)\right\} \\ &\quad + \max\left\{F\left(a'^2 + c'^2 - b'^2\right), F'\left(a^2 + c^2 - b^2\right)\right\} \leq 2Fcc'. \end{aligned} \tag{1}$$

Post's proof is elementary but clever, and it involves only a bit of trigonometry. It relies on the useful preliminary result that if one triangle fits in a second in any way whatsoever, then it also fits in such a way that one of its sides lies on a side of the containing triangle. (It is this result of Post's that Sullivan [**51**] generalized.)

The inequalities (1) can be rewritten via the Law of Cosines so as to involve the angles of the triangles, and the resulting inequalities, also noted by Post, are somewhat simpler. But at best these conditions are opaque and intractable.

Virtually nothing is known about the analogous problems in space: find a necessary and sufficient condition for a triangle (that is, a thin triangular plate) to fit in a tetrahedron, or for one tetrahedron to fit into another. The corresponding problems for a p-simplex in a q-simplex in $\mathbb{R}^d$ boggle the mind.

Equilateral triangle in triangle

Specializing Post's inequalities (1) by setting $a' = b' = c' = s$ would give a necessary and sufficient condition on the four sides s, a, b, c for an equilateral triangle of side s to fit into a triangle with sides a, b, c, but the resulting condition, even when the notational normalization $a \geq b \geq c$ is taken into account, is far from simple. Jerrard and Wetzel [**25**] have recently established somewhat more geometric conditions by determining the side $s_{\max}$ of the largest equilateral triangle that fits in the given triangle T. It is interesting that again a cubic constraint appears. We sketch this result.

Arrange the notation for T so that $a \geq b \geq c$, and introduce a rectangular coordinate system in such a way that $B = (0, 0)$, $C = (a, 0)$, and A has positive ordinate. Then A lies to the left of the perpendicular bisector MD of BC and inside the circular arc BD with center C and radius a (FIGURE 4). It turns out that when $\beta \leq 60°$ the largest equilateral triangle in T fits with one side along the longest side BC, but when $\beta \geq 60°$, the location of the largest equilateral triangle in T depends on the sign of

$$c \sin\left(60° - \frac{1}{2}\beta\right) - a \sin\left(60° + \frac{1}{2}\beta\right). \tag{2}$$

In coordinates, the locus of the points $A(x, y)$ for which the expression (2) vanishes is an arc of the cubic

$$\sqrt{3}x^3 - x^2y + \sqrt{3}xy^2 - y^3 - 2\sqrt{3}ax^2 - 2\sqrt{3}ay^2 + \sqrt{3}a^2x + a^2y = 0,$$

which is asymptotic to the line ℓ with equation $\sqrt{3}x - y = 2\sqrt{3}a$ (FIGURE 4). The curve crosses its asymptote at the point $(a, -\sqrt{3}a)$. The arc PQ of the cubic that forms part of the boundary between the two significant regions is generated for $60° \leq \beta \leq 80°$.

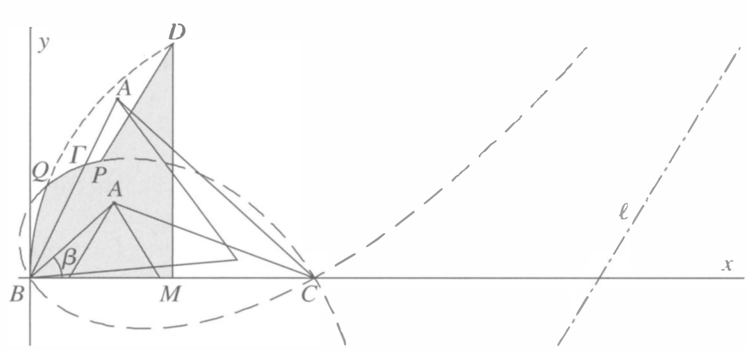


Figure 4 The critical regions

Then one can show that the largest equilateral triangle in T rests on the shortest side c when A lies in the *unshaded* region $QPDQ$ and on the longest side a when A lies in the *shaded* region $BMEDPQB$. Details can be found in Jerrard and Wetzel [**25**].

EXERCISE 4. *Show that the isosceles triangle with apex angle 20° is the only nonequilateral triangle for which the three largest equilateral triangles, one on each side, are equal.*

EXERCISE 5. *Let x be a side of a triangle T, and let s_{in}^x be the side of the largest equilateral triangle that fits in T with one side along the side x and s_{out}^x the side of the smallest equilateral triangle that contains T with one side along x (possibly extended). Show that $s_{\text{in}}^x s_{\text{out}}^x = 4F/\sqrt{3}$, irrespective of the shape of T, where F is the area of T. (See Jerrard and Wetzel [25].)*

The converse problem, to determine precisely when the triangle T fits in the equilateral triangle Δ with side s, has apparently not been examined, and other interesting fit problems involving two triangles remain unstudied. For example, precisely when does a right triangle with legs a, b fit in a right triangle with legs c, d?

Rectangle in rectangle

In 1914, Flanagan [**18**] asked for the second side d of a $c \times d$ rectangle inscribed in an $a \times b$ rectangle (a question with many precursors in the problem literature, most asking about the dimensions of a rug laid diagonally on a rectangular floor), and a few years later Dunkel [**13**] provided an elaborate discussion of the resulting quartic equation. In 1956, Ford [**19**] asked when a $p \times q$ rectangle fits in an $a \times b$ rectangle. A necessary and sufficient condition on the sides p, q, a, b was soon supplied by Carver [**7**]. Here is Carver's nice result in an elegant formulation having its roots in Dunkel's discussion. Carver's argument is straightforward; a quite different formulation appears in Wetzel [**55**].

THEOREM 3. (CARVER [7]) *Suppose an $a \times b$ rectangle T is given, with the notation arranged so that $a \geq b$. Then a $p \times q$ rectangle R with $p \geq q$ fits into T if and only if*

a. $p \leq a$ and $q \leq b$, or

b. $p > a$, $q \leq b$, and

$$\left(\frac{a+b}{p+q}\right)^2 + \left(\frac{a-b}{p-q}\right)^2 \geq 2.$$

Little seems to be known about the corresponding problems in space. A necessary and sufficient condition on p, q, a, b, c for a $p \times q$ rectangle (that is, a thin rectangular plate) to fit into an $a \times b \times c$ box (a rectangular parallelepiped) remains unknown, and determining more generally when a $p \times q \times r$ box fits in an $a \times b \times c$ box seems very difficult. The analogous questions in higher dimensions seem completely beyond reach.

One might ask in particular for the side of the largest m-cube that fits in an n-cube. This question, at least in the case of a cube in a tesseract, was raised by Gardner [**21**, p. 172], who summarized what little is known about this general question.

Jerrard and I have recently found the largest rectangle of given shape that fits in a unit cube [**26**]. The special case of fitting a square in the unit cube is equivalent to the well-known Prince Rupert's problem of passing a unit cube through a suitable hole in another unit cube. Rupert showed that a hole can be made in a cube of side one that is large enough to permit a second cube of side one to pass through, and a Dutch scientist, Pieter Nieuwland, proved that a cube of side less than $\frac{3}{4}\sqrt{2} \approx 1.06066$ can pass through a suitable hole in a unit cube. The history of this result was beautifully summarized in 1950 by Schreck [**47**].

Such *passage* questions are not exactly fit questions, but they are closely related. Jerrard and I [**26**] have recently generalized Nieuwland's result and determined precisely when an $L \times \lambda L \times \mu L$ box (where $0 < \mu \leq \lambda \leq 1$) can pass through a suitable hole in a unit cube. Jerrard has also shown that a regular tetrahedron with edge slightly larger than one can pass through a suitable hole in a regular tetrahedron of edge one. More generally, how large a Platonic solid can pass though a suitable hole in a second Platonic solid with edge one? To my knowledge nothing much is known about such questions.

Further fits

Even in the plane the subject is filled with uninvestigated questions that seem to require little more than insight and elementary mathematics to study, and in higher dimensions—even in $\mathbb{R}^3$—interesting questions abound, of unknown depth and difficulty. Precisely when does a regular n-gon of side s fit in the triangle with sides a, b, c? in a regular m-gon? in a box with edges a, b, c? in a Platonic solid of edge e? I have investigated the question of precisely when a $u \times v$ rectangle fits into a triangle with sides a, b, c [**56**], but the converse question of fitting a triangle into a rectangle apparently remains unexamined. When does a given triangle, regular n-gon, rectangle, ... fit in a given ellipse? sphere? ellipsoid? Little is known about when one polyhedron fits into another, although Croft, Falconer, and Guy [**11**, pp. 52–53] report a few results for Platonic solids. There seems little chance of a general theory of fits.

A fool can ask more questions than seven wise men can answer. —Old Proverb

Covers

In problems of *cover* type, a family of figures is given, and one seeks a compact (convex) target set, perhaps of prescribed shape, that covers every figure in the family and is small in some specified sense, area, perimeter, etc.

A prototypical example dating from 1914 is the well-known Lebesgue *universal cover problem*, which asks for the convex set of least area that is large enough to cover each set of diameter one (see Lebesgue [**34**], Croft, Falconer, and Guy [**11**, pp. 125–27]). The problem is known to have a solution, and the least area is known to lie between 0.8263 and 0.8442, but little more is known.

Here is an easy introductory example.

EXERCISE 6. *A* kite *is a quadrilateral region in the plane whose diagonals intersect and are perpendicular. Find the smallest disk that can cover (a) every kite whose diagonals have given lengths a and b, (b) every kite whose diagonals have given sum s, and (c) every kite whose perimeter is at most p. (Hint: the median of a triangle is shorter than the average of the two adjacent sides.)* [Answer on page 398.]

Worm problems

In most of our examples, the figures to be covered are rectifiable arcs of various kinds. Suppose a family C of arcs is given (for instance, arcs of unit length, convex unit arcs, closed unit arcs, unit polygonal arcs having at most two, three, n segments), and suppose a convex region F is given in the plane. A *worm problem* for F and $\mathcal{C}$ is to find the smallest convex region similar to F that contains a congruent copy of every arc in $\mathcal{C}$. The earliest problem of this type of which I am aware was posed by L. Moser **[39]** more than 35 years ago (see also W. Moser [**40**]):

> *Find the (convex) set of least area that contains a congruent copy of each arc in the plane of length one.*

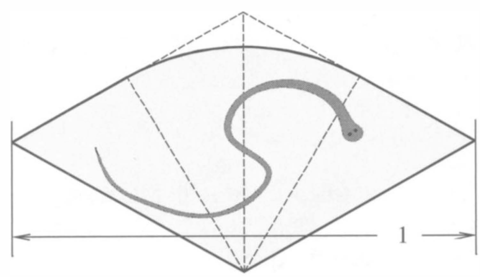


Figure 5 The smallest known worm blanket

The unit arcs are sometimes called *worms* (*inchworms*?), and the problem has been phrased in different ways in the literature: The architect's version (find the smallest comfortable living quarters for a unit worm), the humanitarian version (find the shape of the most efficient worm blanket), the sadistic version (find the shape of the best mallet head...), and so on. Partial results are known, including the existence of such a minimum cover, but its shape and area remain unknown. The best bounds presently known for its area μ are

$$0.21946 < \mu < 0.27524. \tag{3}$$

Lost in a forest An old problem of Bellman [**2**] is closely related to these worm problems. The problem can be phrased in the following picturesque way (see Finch and Wetzel [**17**]):

> *A hiker is lost in a forest whose shape and dimensions are precisely known to him. What path should he follow to escape from the forest in the shortest possible time?*

The problem is to find the shortest path that is guaranteed to reach the boundary of the forest, given complete knowledge of the size and shape of the forest but no information about where one is within the forest or what initial direction to set out. There seem to be no general methods to attack problems of this sort, and the best path has been found in only a very few special cases. The length of a shortest escape path appears to be a basic geometric property of the region, like its perimeter, area, diameter, and minimum width. See Berzsenyi [**3**] and Finch and Wetzel [**17**] for a survey of this problem area.

Bellman's question is related to the worm problem in the following way. The shortest escape path γ from a forest F is the shortest path that does not fit in the interior of F. It need not be unique in shape but, of course, its length is uniquely determined. If a shortest escape path has length ℓ, then F must be a cover for the family of all arcs of length ℓ. Indeed, the length ℓ of a shortest escape path from a forest F is the largest number s for which F is a cover for all arcs of length s.

The broadworm Bellman asked in particular about a forest in the shape of an infinite strip of unit width. The shortest escape path, pictured bold in FIGURE 6, was described in 1961 by Zalgaller [**62**], rediscovered in 1968 by Schaer [**45**] (who named it the "broadworm" because it is the curve whose minimum width is as large as possible for curves of given length), discovered again 1986 by Klötzler [**31, 32**] (who called it the "universal escape curve"), and rediscovered again in 1989 by Adhikari and Pitman (who called it the "caliper") [**1**]. We sketch Schaer's description of the unit broadworm (from Schaer [**45**]).

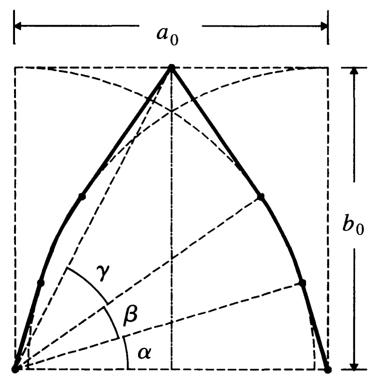


Figure 6 The unit broadworm

The broadworm of length 1 is formed by four line segments and two circular arcs as shown, where (switching briefly to radians),

$$\alpha = \arcsin\left[\frac{1}{6} + \frac{4}{3}\sin\left(\frac{1}{3}\arcsin\frac{17}{64}\right)\right] \approx 0.290046,$$

$$\gamma = \arctan\left(\frac{1}{2}\sec\alpha\right) \approx 0.480931,$$

$$\beta = \frac{\pi}{2} - \alpha - 2\gamma \approx 0.318888, \quad \text{and}$$

$$a_0 = b_0 \sec\alpha \approx 0.458058, \qquad b_0 = \frac{1}{2}\left(\beta + \tan\alpha + \tan\gamma\right)^{-1} \approx 0.438925,$$

expressions that arise from the trigonometric solution of the cubic that appears in an extremum problem. The angle at the top of the broadworm is about 69.8°, and the minimum width of the curve is the vertical height b_0.

By duality, the shortest curve with given minimal width is the broadworm, suitably scaled. The shortest escape path for the infinite strip of width one is the broadworm scaled to have minimal width 1, whose length is $1/b_0 \approx 2.278292$. This path is also the best escape path for oblong rectangular forests, namely, $1 \times r$ rectangular forests for which $r \geq \sqrt{b_0^{-2} - 1} \approx 2.047099$ (see Finch and Wetzel [**17**]).

EXERCISE 7. *Show that the best escape curve for a* $1 \times r$ *rectangular forest with* $r < \sqrt{b_0^{-2} - 1}$ *is a line segment of length* $\sqrt{1 + r^2}$.

At this point we return to Moser's worm problem and consider the area bounds (3). Let $\mathcal{C}$ be the family of all curves of length one. A convex cover for $\mathcal{C}$ must, in particular, cover the line segment of length one and also the unit broadworm, so its diameter must be at least 1 and its width in the direction perpendicular to the diameter must be at least b_0. Consequently its area must be at least $(b_0 \times 1)/2 \approx 0.21946$ (Schaer and Wetzel **[46]**, Chakerian and Klamkin **[9]**).

The upper bound is the area of the smallest cover for C presently known, the trimmed quadrilateral pictured in FIGURE 5 found by Norwood, Poole, and Laidacker **[42]** in 1992. Two equilateral triangles with altitude $1/2$ are joined at the base, and a corner is pruned off the resulting rhombus by a circular arc. This cover is probably not minimal.

Worm problems have considerable appeal for amateurs. In 1996, Stewart **[50**, pp. 98–99] devoted a *Scientific American* Mathematical Recreations column entitled "Mother worm's blanket" to insightful work by one David Reynolds, then a software engineer at the Credence Systems Corporation. Reynolds was reported to have found a cover with area about 0.239. The claim was withdrawn a few months later when a reader found a zigzag curve of length 1 that does not fit in Reynold's proposed cover.

The problem has been studied in higher dimensions, too, where Stewart **[49]** asks for the "smallest sleeping bag for a baby snake." Schaer and Wetzel **[46]** showed that the cube with principal diagonal 1 in $\mathbb{R}^d$ is such a sleeping bag. In $\mathbb{R}^3$ a cube of diagonal 1 has volume about 0.19245. Lindström **[35]** found a convex cover in $\mathbb{R}^3$ whose volume is about 0.15953; but the smallest (convex) sleeping bag in $\mathbb{R}^3$ presently known, published in 2001 by Håstad, Linusson, and Wästlund **[23]**, has volume about 0.075803. The probable shape of the broadworm in $\mathbb{R}^3$, that is, the curve in $\mathbb{R}^3$ of length one whose minimum width is as large as possible, was described in 1994 by Zalgaller **[63]**, but nothing is known in $\mathbb{R}^d$ for $d \geq 4$.

Worm covers of prescribed shape

Many partial results for covers of given shape can be found in the literature, but best possible results are scarce. We continue to write C for the family of all arcs of unit length.

EXERCISE 8. *Show that every curve in C lies in a disk of diameter* 1. [Answer on page 398.]

The smallest *semidisk* cover for C (that is, the region bounded by a semicircle and its diameter) has diameter 1. Meir's elegant proof of this fact, reported in Wetzel **[57]**, also works in higher dimensions **[53]** and shows that the *semiball* (bounded by a hemisphere of radius $1/2$ and a hyperplane through its center) is a cover for the family of all curves of length 1. Reynolds (see Stewart **[50]**) gives a very nice argument for this result in the plane.

The smallest *square* cover for C has side $1/\sqrt{2}$ and area 0.5; the smallest *rectangular* cover is $b_0 \times \sqrt{(1 - b_0^2)}$ and has area about 0.3944. With area about 0.3927, Meir's semidisk is a little smaller than this smallest rectangle.

Norwood, Poole, and Laidacker **[42]** conjectured that a 30°-60°-90° right triangle large enough to contain a square of side $1/3$ with a corner at the right angle is a cover for C. This was recently disproved by Dr. Eric Ferguson, a retired engineer living in Eindhoven, The Netherlands, who noted in an email message on March 10, 2000 that the "staple," formed by three segments each of length $1/3$ at angles of 91°, does not fit. Norwood, Poole, and Laidacker also mention two conjectures that I made in the early 1970s, neither of which has been resolved:

CONJECTURE 1. (WETZEL) *The circular sector with angle* 30° *and radius* 1 *is a cover for C.*

The area of this sector is $\pi/12 \approx 0.26180$.

CONJECTURE 2. (WETZEL) *The smallest* 30°-60°-90° *right triangle that contains a square of side* 1/3 *resting on its hypotenuse is a cover for* C.

The hypotenuse of this right triangle is $(3+4\sqrt{3})/9 \approx 1.10313$, its legs are $(3+4\sqrt{3})/18 \approx 0.55157$ and $(4+\sqrt{3})/6 \approx 0.955342$, and its area is $(24+19\sqrt{3})216 \approx 0.26347$.

EXERCISE 9. *Prove that a semicircle of length* 1 *fits into the region in each of these conjectures.*

The most *frustrating* conjecture concerning covers for C of prescribed shape is the dual of a conjecture made by Besicovitch 35 years ago in response to a question posed by Graham [**22**, p. 98]: *Find the shortest arc in the plane that does not fit in an open equilateral triangle of side 1.* (In terms of the "lost in the forest" metaphor, the problem is to find the shortest escape path from an equilateral-triangular forest of side 1.) Besicovitch [**4**] described a three-segment polygonal arc of length $\sqrt{27/28} < 1$ that does not fit in an open equilateral triangle of side 1, and he conjectured that every shorter arc can be covered by the closed equilateral triangle of side 1.

Figure 7 The Besicovitch Z

This so-called Besicovitch Z (FIGURE 7) is the centrally symmetric three-segment unit polygonal arc $ABCD$ with $AB = BC = CD = \sqrt{3/28}$, $AB \parallel CD$, and $\angle CAB = \arcsin 1/\sqrt{28} \approx 10.9°$. (Besicovitch determined the length numerically; the radical expression was found by Steven Knox in 1994.) The dual of Graham's question is a covering problem, and for it the Besicovitch conjecture amounts to the following:

CONJECTURE 3. (BESICOVITCH [**4**]) *The smallest equilateral triangle that covers* C *has side* $\sqrt{28/27} \approx 1.01835$.

It is known that any arc of length 1 that cannot be covered by the equilateral triangle with side $\sqrt{28/27}$ must be a *zigzag* in the sense that it crosses the (open) line segment that joins its endpoints. A bit more can be said, but a proof of the general Besicovitch conjecture, which is almost certainly true, remains frustratingly beyond grasp.

The same question can be asked for triangles of any shape. Here, for example, is the analogous conjecture for isosceles right triangles:

CONJECTURE 4. *The smallest isosceles right triangle that covers* C *has hypotenuse* $\sqrt{10/9} \approx 1.05409$.

The analogue of Besicovitch's problem in higher dimensions is unexplored: Find the smallest regular simplex that can cover every arc of length one in $\mathbb{R}^d$, or, equivalently, find the shortest arc that cannot be covered by an (open) regular simplex of side 1. One could ask the same question for the Platonic solids or, indeed, for any polytope in any dimension.

Covers for other families of curves

Similar questions have been asked for other families of curves (and indeed for other collections of figures), but few have been answered.

Convex arcs of unit length Let $\hat{C}$ be the collection of all convex unit arcs (that is, simple arcs of length one that lie on the boundary of their convex hull). Every such arc lies in an isosceles right triangle with hypotenuse 1, whose area is 1/4.

EXERCISE 10. *Show that every convex unit arc lies in some isosceles right triangle of hypotenuse* 1. *(Suggestion: pause and reflect. See Moser [**39**, pp. 218–19], and Johnson, Poole, and Wetzel [**28**].)*

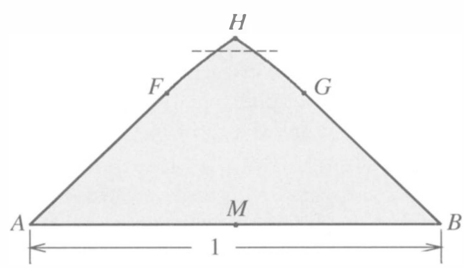


Figure 8 The smallest known cover for convex unit arcs

But tall arcs can be turned, so some of the space near the right angle vertex isn't needed. One can show that the isosceles right triangle ABC with hypotenuse AB of length 1 can be clipped parallel to AB at height $MH = \sqrt{2}/3 \approx 0.47140$ to give a cover for $\hat{C}$ whose area is $(\frac{1}{2} - \frac{1}{3}\sqrt{2})^2 \approx 0.00082$ smaller than 1/4. The smallest cover for $\hat{C}$ presently known, pictured in FIGURE 8, was recently described by Johnson, Poole, and Wetzel [**28**]. The points F and G are above the trisection points of the hypotenuse, and the boundary arcs FH and GH are symmetric in the median line MH, parabolic with axes vertical, tangent at F and G, and rise to the point H. This cover has area about 0.24656. Perhaps one can further clip down to the height of the broadworm, $b_0 \approx 0.4389$, indicated by the dashed line.*

No better lower bound for the area of covers for $\hat{C}$ than the one for Moser's problem is known, so the least area of a cover for $\hat{C}$ lies between 0.21946 and 0.24656.

Closed curves of unit length The collection C_0 of all *closed* curves of length 1 is also of particular interest. The best estimate in the literature for the least area μ_0 of a cover for the collection C_0 is $0.09633 \le \mu_0 \le 0.12274$ (Chakerian and Klamkin [**9**], Schaer and Wetzel [**46**]). Z. Füredi and I have recently improved these bounds slightly to $0.09667 \le \mu_0 \le 0.11754$, but the gap remains wide.

Some results are known for covers of prescribed shapes for C_0. Somewhat surprisingly, the smallest *triangular* cover for C_0 of any given similarity class is known. It follows from an averaging inequality due to Eggleston [**14**, 157–58] that a triangle T is a cover for C_0 if and only if its incircle has circumference 1 (see Wetzel [**54, 60**]). The smallest of these covering triangles is equilateral with side $\sqrt{3}/\pi$, and its area is about 0.13162. The smallest rectangular cover (Schaer and Wetzel [**46**]) for C_0 is the $1/\pi \times \sqrt{\pi^2 - 4}/2\pi$ rectangle, whose area is about 0.12274.

Triangles of unit perimeter The family $\mathcal{T}_0$ of triangles of perimeter 1 is of interest for itself, but studying it might also give some insights into the family C_0. It is easy

*Added in proofs: Wacharin Wichiramala has found a cover for the family $\hat{C}$ that is a few ten-thousandths smaller.

to show that the smallest disk that covers $\mathcal{T}_0$ has diameter 1/2. Wetzel [**59**] found the smallest equilateral triangle that is a cover for $\mathcal{T}_0$.

EXERCISE 11. *Find the rectangular cover for $\mathcal{T}_0$ that has least area. (See Wetzel [52].* [Answer on page 398.]

The smallest convex cover for this family has recently been found, so there may be some hope that the problem for the family $\mathcal{C}_0$ of all closed curves of length 1 might eventually yield. It is interesting that the smallest cover turns out to be a triangle! This smallest cover has area about 0.072375.

THEOREM 4. (FÜREDI AND WETZEL [**20**]) *The smallest convex set in the plane that contains a congruent copy of every triangle of perimeter 1 is the triangle $T = ABC$ with $AB = 1/3$, $\angle B = 60°$, and $BC = s_0$, where s_0 be the maximum of the trigonometric function*

$$f(\theta) = \frac{\sqrt{3}\sin\left(\frac{1}{3}\pi + \theta\right)}{3\left(1 + \sin\frac{1}{2}\theta\right)}$$

on the interval $[0, \pi/3]$.

A little numerical work shows that $s_0 \approx 0.50143$. As with most results of this kind, the proof, which relies only on trigonometry and a bit of the calculus, is entirely elementary.

Triangles of unit diameter The collection $\mathcal{T}_1$ of all triangles of diameter 1 is also of interest. Twenty years ago, M. D. Kovalev [**33**] showed, somewhat surprisingly, that the unique smallest convex cover for $\mathcal{T}_1$ is a certain *triangle.* In [**59**] Wetzel remarked that the smallest equilateral triangular cover for the family $\mathcal{T}_1$ has side $(2\cos 10°)/\sqrt{3}$, and Füredi and Wetzel [**20**] used this result to give a different proof of Kovalev's result. More generally, Yuan and Ding [**61**] have recently found the smallest triangular cover of prescribed similarity class that is a cover for $\mathcal{T}_1$.

The general worm problem

This situation can be formulated in considerable generality. We continue with the plane, but analogous questions can, of course, be asked in arbitrary dimensions.

> *A collection $\mathcal{C}$ of plane figures F and a transitive group $\mathcal{G}$ of motions of the plane are given. Find a minimal convex target set T (minimal in the sense of having least area, perimeter, or whatever) so that for each $F \in \mathcal{C}$ there is a motion $\mu \in \mathcal{G}$ with $\mu(F) \subseteq T$. We call T a $\mathcal{G}$-cover for $\mathcal{C}$.*

If the collection $\mathcal{C}$ is called the *can* and the figures are called *worms*, posing a problem of this kind is "opening a can of worms." Existence of $\mathcal{G}$-covers can be guaranteed under certain natural hypotheses by fundamental compactness results like the Blaschke Selection Theorem. See, for example, Kelly and Weiss [**30**]. Unfortunately, little more is known in this generality.

Sometimes the subgroup of all *direct* (that is, orientation preserving) motions is appropriate, when the problem does not permit an arc to be replaced by its mirror image. For other problems (including the original Moser worm problem) opposite (that is, orientation reversing) motions are permitted. Various problems for which the motion group is the group of translations have been studied in the literature. See, for example, Croft, Falconer, and Guy [**11**], Bezdek and Connelly [**5**], Wetzel [**57**].

The nonconvex case If the target set is not assumed convex, the context of the problem changes dramatically, from geometric to measure-theoretic or fractal, and "area" or "measure" are supplemented or even replaced by questions about the Hausdorff measure and Hausdorff dimension. The smallest convex set that contains a unit line segment in every direction (that is, that contains a translate of every line segment of length one) is the equilateral triangle of altitude 1 (Pál [**43**]), but Besicovitch has described a set of Lebesgue measure zero having the same property. The smallest convex set that contains a congruent copy of every rectangle whose sides are at most one is the unit square, but the *Cantor Tartan*, $T = (K \times I) \cup (I \times K)$, where K is the classical Cantor "middle third" set in the interval $I = [0, 1]$, also has this property and has measure 0. More generally, Jones and Schaer [**29**] have shown that for each $\varepsilon > 0$ there is a set of measure less than ε that contains a congruent copy of every polygonal arc; and Davies [**12**] has shown the existence of such a set having measure zero! On the other hand, Marstrand [**37**] (see also [**11**, pp. 174–75]) has shown that every measurable set that is a cover of the collection $\mathcal{C}$ of all curves of length one must have positive measure. Perhaps for each $\varepsilon > 0$ there is such a cover with measure less than ε, but if true, this likely lies quite deep. Falconer's book [**15**, §7.4] offers a glimpse of this difficult material.

Note

This article is a revised and expanded version of a survey lecture given at the University of Illinois April 30, 2001 and at Eastern Illinois University August 17, 2001.

REFERENCES

→ A. Adhikari and J. Pitman, The shortest planar arc of width 1, *Amer. Math. Monthly* **96** (1989), 309–327.
2. R. Bellman, Minimization problem, *Bull. Amer Math. Soc.* **62** (1956), 270.
3. G. Berzsenyi, Lost in a forest, *Quantum* **6** (1995), 41.
→ A. S. Besicovitch, On arcs that cannot be covered by an open equilateral triangle of side 1, *Math. Gaz.* **49** (1965), 286–288.
→ K. Bezdek and R. Connelly, Covering curves by translates of a convex set, *Amer. Math. Monthly* **96** (1989), 789–806.
6. E. W. Brune, Grösstes Quadrat im Dreiecke, *Crelle's Journal* (= *J. reine angew. Math.*) **15** (1836), 365–66.
→ W. B. Carver, Solution to Problem E1225, *Amer. Math. Monthly* **64** (1957), 114–6.
8. A. Cayley, Problem 3564, *Mathematical Questions with their Solutions from the "Educational Times,"* Vol. XVII, C. F. Hodgson & Sons, London, 1872, p. 72.
→ G. D. Chakerian and M. S. Klamkin, Minimal covers for closed curves, this MAGAZINE, **46** (1973), 55–61.
→ R. J. Chapman, Solution to Problem 10251, *Amer. Math. Monthly* **102** (1995), 465–67.
11. H. T. Croft, K. J. Falconer, and R. K. Guy, *Unsolved Problems in Geometry*, Springer-Verlag, New York, 1991.
12. R. O. Davies, Some remarks on the Kakeya problem, *Proc. Cambridge Philos. Soc.* **69** (1971), 417–421.
→ O. Dunkel, Solution I to Problem 416, *Amer. Math. Monthly* **27** (1920), 327–30.
14. H. G. Eggleston, *Problems in Euclidean Space: Applications of Convexity*, Pergamon, New York, 1957.
15. K. J. Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, Cambridge, 1985.
16. S. R. Finch, *Mathematical Constants*, Cambridge University Press, Cambridge, 2003.
17. S. R. Finch and John E. Wetzel, Lost in a forest, to appear in *Amer. Math. Monthly.*
→ C. E. Flanagan, Problem 416, *Amer. Math. Monthly* **22** (1914), 156.
→ L. R. Ford, Problem E1225, *Amer. Math. Monthly* **63** (1956), 421.
20. Z. Füredi and J. E. Wetzel, The smallest convex cover for triangles of perimeter two, *Geom. Dedicata* **81** (2000), 285–293.
21. M. Gardner, *The Colossal Book of Mathematics*, W. W. Norton, New York, 2001.
22. R. Graham, Problem 41, *Proceedings, 1963 Number Theory Conference,* Univ. of Colorado, Boulder, Colorado, 1963.
23. J. Håstad, S. Linusson, and J. Wästlund, A smaller sleeping bag for a baby snake, *Discrete Comput. Geom.* **26** (2001), 173–181.
24. T. L. Heath, *The Thirteen Books of Euclid's Elements,* 3 vols., 2nd ed., Dover, New York, 1956.

→ R. P. Jerrard and J. E. Wetzel, Equilateral triangles and triangles, *Amer. Math. Monthly* **109** (2002), 909–915.

→ ———, Prince Rupert's rectangles, to appear in *Amer. Math. Monthly* **111**:1 (2004), 22–31.

27. ———, Triangles in squares, *Elem. Math.* **58** (2003), 65–72.

28. J. A. Johnson, G. D. Poole, and J. E. Wetzel, A small cover for convex unit arcs, to appear in *Discrete Comput. Geom.*.

29. J. P. Jones and J. Schaer, The worm problem, Dept. of Math., Stat., and Comp. Sci. Research Paper No. 100, University of Calgary, May, 1970.

30. P. J. Kelly and M. L. Weiss, *Geometry and Convexity: A Study in Mathematical Methods,* John Wiley & Sons, New York, 1979.

31. R. Klötzler, Universale Rettungskurven, I, *Z. Anal. Anwendungen* **5** (1986), 27–38.

32. R. Klötzler and S. Pickenhain, Universale Rettungskurven, II, *Z. Anal. Anwendungen* **6** (1987), 363–369.

33. M. D. Kovalev, A minimal convex covering for triangles (in Russian), *Ukrain. Geom. Sb.* **26** (1983), 63–68.

34. H. Lebesgue, Sur le problème des isopéremètres et sur les domaines de largeur constante, *Bull. Soc. Math. France* **42** (1914), 72–76.

35. B. Lindström, A sleeping bag for a baby snake, *Math. Gaz.* **81** (1997), 451–452.

36. J. S. Madachy, *Madachy's Mathematical Recreations,* Dover, New York, 1979.

37. M. Marstrand, Packing smooth curves in R^q, *Mathematika* **26** (1979), 1–12.

→ J. G. Mauldon, Problem 10251, *Amer. Math. Monthly* **99** (1992), 782.

39. L. Moser, *Poorly formulated unsolved problems in combinatorial geometry,* mimeographed, (undated, but about 1966).

40. W. O. J. Moser, Problems, problems, problems, *Discrete Appl. Math.* **31** (1991), 201–225.

41. W. O. J. Moser and J. Pach, *100 research problems in discrete geometry,* McGill University, Montreal, 1986,

42. R. Norwood, G. Poole, and M. Laidacker, The worm problem of Leo Moser, *Discrete Comput. Geom.* **7** (1992), 153–162.

43. J. Pál, Ein Minimumproblem für Ovale, *Math. Ann.* **83** (1921), 311–19.

44. K. A. Post, Triangle in a triangle: On a problem of Steinhaus, *Geom. Dedicata* **45** (1993), 115–20.

45. J. Schaer, The broadest curve of length 1, Dept. of Math., Stat., and Comp. Sci. Research Paper No. 52, University of Calgary, May, 1968.

46. J. Schaer and J. E. Wetzel, Boxes for curves of constant length, *Israel J. Math.* **12** (1972), 257–265.

47. D. J. E. Schreck, Prince Rupert's Problem and its extension by Pieter Nieuwland, *Scripta Math.* **16** (1950), 73–80, 261–267.

48. H. Steinhaus, *One Hundred Problems in Elementary Mathematics,* Basic Books, New York, 1964.

49. I. Stewart, *Game, Set and Math: Enigmas and Conundrums*, Blackwell, London, 1989.

50. ———, Mathematical Recreations: Mother worm's blanket, *Scientific American*, Jan., 1996, 98–99. Mathematical Recreations: Feedback, *Scientific American*, June, 1996, 103.

51. J. M. Sullivan, Polygon in triangle: generalizing a theorem of Post. (In preparation.)

→ J. E. Wetzel, Boxes for isoperimetric triangles, this MAGAZINE, **73** (2000), 315–319.

53. ———, Covering balls for curves of constant length, *Enseignement Math.* **17** (1971), 275–277.

54. ———, On Moser's problem of accommodating closed curves in triangles, *Elem. Math.* **25** (1970), 78–81.

→ ———, Rectangles in rectangles, this MAGAZINE, **73** (2000), 204–211.

56. ———, Rectangles in triangles, to appear in *J. Geom.*.

57. ———, Sectorial covers for curves of constant length, *Can. Math. Bull.* **16** (1973), 367–376.

→ ———, Squares in triangles, *Math. Gazette* **86** (2002), 28–34.

→ ———, The smallest equilateral cover for triangles of perimeter two, this MAGAZINE, **70** (1997), 125–130.

60. ———, Triangular covers for closed curves of constant length, *Elem. Math.* **25** (1970), 78–81.

61. L. Yuan and R. Ding, The smallest triangular cover for triangles of diameter one. (Preprint.)

62. V. A. Zalgaller, How to get out of the woods? On a problem of Bellman (Russian), *Matematicheskoe Prosveshchenie* **6** (1961), 191–195.

63. ———, The problem of the shortest space curve of unit width (Russian), *Mat. Fiz. Anal. Geom.* **1** (1994), 454–461.

The Operator $(x\frac{d}{dx})^n$ and Its Applications to Series

PETER M. KNOPF
Pace University
Pleasantville, NY 10570
pknopf@pace.edu

My research into the operator $(x\frac{d}{dx})^n$ has been a mathematical adventure. It led me to the rare book collection at Columbia University in Manhattan, where I read books written in the 18th century that were literally falling apart in my hands. Independently, I discovered how to obtain the full Maclaurin series of the elementary trigonometric functions other than the sine and cosine. I used the operator to sum familiar series from calculus that can't be summed using the standard techniques covered in a calculus course. Although the operator $(x\frac{d}{dx})^n$ is unfamiliar to many, it has a surprising number of applications. As far back as 1740, Leonhard Euler used the operator as a tool in his work [**1**, pp. 1080–1081]. This article is devoted to applications of this operator that may interest students of calculus.

As a natural starting point, consider the infinite series $\sum_{k=0}^{\infty} k^n x^k$ that frequently arises in calculus, at least for particular values of x and n. The focus in calculus is often to determine the domain of convergence. What about the exact sum? If $n = 0$, we have the familiar geometric series

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \tag{1}$$

for $-1 < x < 1$. To handle $n = 1$, we apply the operator $x\frac{d}{dx}$ to each side of equation (1) to obtain

$$\sum_{k=0}^{\infty} kx^k = x(1-x)^{-2}$$

for $-1 < x < 1$. To find the sum for any positive integer value of n takes some work. If we apply the operator $x\frac{d}{dx}$ n times, which we denote by the symbol $(x\frac{d}{dx})^n$, to equation (1) we have

$$\sum_{k=0}^{\infty} k^n x^k = \left(x\frac{d}{dx}\right)^n \left(\frac{1}{1-x}\right). \tag{2}$$

Once we determine the result of the action of the operator $(x\frac{d}{dx})^n$ on the function $(1-x)^{-1}$, we will have the desired sum.

Stirling numbers of the second kind and summing infinite series

All of the applications in this paper hinge on the effect of the operator $(x\frac{d}{dx})^n$ on a function. For an arbitrary infinitely differentiable function f, let's check by induction that

$$\left(x\frac{d}{dx}\right)^n f(x) = \sum_{k=1}^{n} \left\{ {n \atop k} \right\} x^k \frac{d^k f}{dx^k}, \tag{3}$$

where $\left\{ {n \atop k} \right\}$ are constants, which for the time being are unknown. In this paper, I will refer to (3) as the *main formula*. The coefficients $\left\{ {n \atop k} \right\}$ are known as Stirling numbers of the second kind. They are of particular interest to specialists in combinatorics, because $\left\{ {n \atop k} \right\}$ is equal to the number of ways a set of n objects can be partitioned into k nonempty subsets [**2**, Section 8.2].

For $n = 1$, the formula is obviously true with $\left\{ {1 \atop 1} \right\} = 1$. Assume that (3) is true for a particular value of n and apply the product rule to find

$$\left(x\frac{d}{dx}\right)^{n+1} f(x) = \sum_{k=1}^{n}\left(k\left\{ {n \atop k} \right\}x^k\frac{d^kf}{dx^k} + \left\{ {n \atop k} \right\}x^{k+1}\frac{d^{k+1}f}{dx^{k+1}}\right)$$

$$= \left\{ {n \atop 1} \right\}x\frac{df}{dx} + \sum_{k=2}^{n}\left(k\left\{ {n \atop k} \right\} + \left\{ {n \atop k-1} \right\}\right)x^k\frac{d^kf}{dx^k} + \left\{ {n \atop n} \right\}x^{n+1}\frac{d^{n+1}f}{dx^{n+1}}.$$

Thus, the formula is true for $n+1$, as long as we set $\left\{ {n+1 \atop 1} \right\} = \left\{ {n \atop 1} \right\}$, $\left\{ {n+1 \atop k} \right\} = k\left\{ {n \atop k} \right\} + \left\{ {n \atop k-1} \right\}$ for $k = 2, \dots, n$, and $\left\{ {n+1 \atop n+1} \right\} = \left\{ {n \atop n} \right\}$. Since $\left\{ {1 \atop 1} \right\} = 1$, we immediately obtain $\left\{ {n \atop 1} \right\} = 1$ and $\left\{ {n \atop n} \right\} = 1$ for $n \geq 1$.

The Stirling numbers are easy to compute, as we will show that they satisfy

$$\left\{ {n \atop k} \right\} = \sum_{j=1}^{k}\frac{(-1)^{k-j}j^n}{j!\,(k-j)!}, \tag{4}$$

for $1 \leq k \leq n$ and $n \geq 1$. To establish (4), we will follow a proof of Heinrich Ferdinand Scherk from 1824 [**6**]. Although this proof originally appeared in 1824, it is reproduced in Scherk's book *Mathematische Abhandlungen* from 1825 [**7**, pp. 11–14]. The book gives a more expanded and accessible treatment of his work.

Scherk was considering a problem slightly different from ours. Given a function $f(x)$ and letting $x = e^y$, he obtained an equation for the nth derivative $\frac{d^nf}{dy^n}$ in terms of the first n derivatives $\{\frac{d^kf}{dx^k}\}_{k=1}^n$. It is easy to see that this equation is the same as our main formula, if one substitutes $\frac{d^nf}{dy^n}$ for $(x\frac{d}{dx})^n f(x)$. Following in Scherk's footsteps, we apply the operator $(x\frac{d}{dx})^n$ to each side of the identity $e^x = \sum_{j=0}^{\infty} x^j/j!$ and use the main formula to obtain

$$e^x\sum_{k=1}^{n}\left\{ {n \atop k} \right\}x^k = \sum_{j=1}^{\infty}\frac{j^n}{j!}x^j. \tag{5}$$

Multiplying both sides of (5) by e^{-x} and substituting the Maclaurin series for e^{-x} gives

$$\sum_{k=1}^{n}\left\{ {n \atop k} \right\}x^k = \sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\frac{(-1)^i j^n}{i!\,j!}x^{i+j}.$$

Changing the summation variable from i to k, where $k = i + j$ results in

$$\sum_{k=1}^{n}\left\{ {n \atop k} \right\}x^k = \sum_{k=1}^{\infty}\sum_{j=1}^{k}\frac{(-1)^{k-j}j^n}{j!\,(k-j)!}x^k. \tag{6}$$

Equating corresponding coefficients of x^k in (6) for $k = 1, \dots, n$ gives (4), as desired.

We return to the question of summing the series $\sum_{k=0}^{\infty} k^n x^k$. Since $\frac{d^k}{dx^k}\left((1-x)^{-1}\right) = k!\,(1-x)^{-k-1}$, applying (2) and the main formula gives the sum of the series:

$$\sum_{k=0}^{\infty} k^n x^k = \sum_{k=1}^{n} k! \left\{ {n \atop k} \right\} x^k (1-x)^{-k-1} \quad \text{for} \quad -1 < x < 1.$$

Another class of series studied in calculus is $\sum_{k=0}^{\infty} k^n x^k / k!$, usually for particular values of x and n. We can sum this series easily by applying the operator $(x\frac{d}{dx})^n$ to the identity $\sum_{k=0}^{\infty} x^k / k! = e^x$. Invoking the main formula gives the desired sum:

$$\sum_{k=0}^{\infty} \frac{k^n x^k}{k!} = e^x \sum_{k=1}^{n} \left\{ {n \atop k} \right\} x^k \quad \text{for} \quad -\infty < x < \infty.$$

Maclaurin series

As mentioned in the introduction, the Maclaurin series for sine and cosine are easy to obtain, but the other elementary trigonometric functions present more of a challenge. In 1671, James Gregory wrote in a correspondence to John Collins [**9**, pp. 61–64] that he had obtained the Maclaurin series for $\tan x$ and $\sec x$. However, no record of his derivations survives. Some eighty years later, Euler obtained the Maclaurin series for $\tan x$, $x \cot x$, and $x \csc x$ in terms of Bernoulli numbers [**3**, pp. 540–543]. He also obtained the Maclaurin series for $\sec x$, where each of the coefficients is expressed as an infinite sum [**3**, pp. 543–545]. Our approach will be in the spirit of Scherk's work from 1825 [**7**, p. 25].

To get my hands on these old manuscripts, I took a train into Manhattan and caught the subway from Grand Central Station to 116th Street, finding the rare book collection on the top floor of Butler Library at Columbia University.* Materials can only be viewed in a glass-walled room, under the watchful eye of a vigilant librarian. Once retrieved from the archive, books must be perched on special foam bases to protect their bindings. I had little difficulty deciphering the mathematics, since the notation, even in the 18th century, was remarkably similar to modern day usage, but the Latin was sometimes a challenge. One old book had deteriorated so badly that several small pieces of the binding fell off, despite all my care.

The approach in this paper arises from that trip back in time via original manuscripts, mixing methods that Scherk used with some that I developed independently to produce relatively streamlined results and proofs (with some help from a referee). The methods to obtain the Maclaurin series for $\tan x$, $x \cot x$, and $x \csc x$ resemble one another. Obtaining the Maclaurin series for $\sec x$ offers an extra twist.

Maclaurin series for $\tan x$ In order to obtain the Maclaurin series for the tangent function, Scherk starts with the identity $\tan x = -i - (e^{ix}+i)^{-1} + (e^{ix}-i)^{-1}$ and uses his analogue to our main formula (3) to compute the higher derivatives of $\tan x$. His form for the nth coefficient of the Maclaurin series of $\tan x$ involves the quantities $\sin((k+1)\pi/4)$ for $k = 1, \ldots, n$. We prefer the identity

$$\tan x = \frac{2i}{e^{2ix}+1} - i, \tag{7}$$

*EDITOR'S NOTE: Readers inspired to view original sources for themselves—rare and often beautiful volumes in the history of mathematics—may find it worth traveling to visit one of the great collections. Among them are: the Burndy Library of the Dibner Institute at MIT, the Lownes Collection at Brown University, the Linda Hall Library in Kansas City, the John Crerar Library at the University of Chicago, and the Barchas Collection at Stanford.

since it yields a simpler form for the Maclaurin series. This identity is easily obtained from Euler's formulas $\sin x = (e^{ix} - e^{-ix})/2i$ and $\cos x = (e^{ix} + e^{-ix})/2$. In order to derive the nth coefficients of the tangent and other trigonometric functions, we will find the following formula useful:

$$\frac{d^n}{dz^n}\left(\frac{1}{e^z+1}\right)\Big|_{z=0} = \sum_{k=1}^{n}\frac{(-1)^k k!}{2^{k+1}}\left\{{n \atop k}\right\}, \tag{8}$$

where z may be complex-valued. To prove (8), let $u = e^z$ and define $f_n(u) = (u\frac{d}{du})^n(u+1)^{-1}$. Since $\frac{du}{dz} = u$, the chain rule gives $\frac{d^n}{dz^n}(e^z+1)^{-1} = f_n(e^z)$. By the main formula (3), we have

$$\frac{d^n}{dz^n}\left(\frac{1}{e^z+1}\right) = \sum_{k=1}^{n}\left\{{n \atop k}\right\}(e^z)^k(-1)^k k!\,(e^z+1)^{-k-1},$$

which gives (8) after evaluating at $z = 0$.

We now easily obtain the coefficients for the Maclaurin series of the tangent function. Letting $z = 2ix$ and referring to (7) and (8), we have

$$\frac{d^n}{dx^n}(\tan x)\Big|_{x=0} = (2i)^{n+1}\frac{d^n}{dz^n}\left(\frac{1}{e^z+1}\right)\Big|_{z=0} = i^{n+1}\sum_{k=1}^{n}(-1)^k 2^{n-k} k!\left\{{n \atop k}\right\}. \tag{9}$$

Since $\tan x$ is real-valued, (9) is zero for even n. Thus the Maclaurin series for $\tan x$ is

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

One can also find this form for the Maclaurin series in Schwatt's book, *An Introduction to the Operations with Series* [**8**]. His book includes most of our results, along with a wealth of information on series. However, his methods are less tidy.

Maclaurin series for $x\cot x$ The function $\cot x$ does not have a Maclaurin series since it is not defined at $x = 0$. On the other hand, the function $x\cot x$ does, since

$$x\cot x = \frac{\cos x}{(\sin x)/x} = \frac{1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots}{1 - \frac{1}{3!}x^2 + \frac{1}{5!}x^4 - \cdots},$$

and we extend $x\cot x$ to equal 1 for $x = 0$. By a method similar to Scherk's [**7**, p. 29], we will obtain the Maclaurin series for $x\cot x$, using a result of Laplace [**5**, pp. 108–109]:

$$\lim_{z\to 0}\frac{d^n}{dz^n}\left(\frac{z}{e^z-1}\right) = \frac{-n}{2^n-1}\frac{d^{n-1}}{dz^{n-1}}\left(\frac{1}{e^z+1}\right)\Big|_{z=0}, \tag{10}$$

where z may be complex-valued. To obtain (10) we differentiate n times the identity $2z/(e^{2z}-1) = z/(e^z-1) - z/(e^z+1)$ and let $z \to 0$, so that

$$\lim_{z\to 0}\frac{d^n}{dz^n}\left(\frac{2z}{e^{2z}-1}\right) = \lim_{z\to 0}\frac{d^n}{dz^n}\left(\frac{z}{e^z-1}\right) - \frac{d^n}{dz^n}\left(\frac{z}{e^z+1}\right)\Big|_{z=0}. \tag{11}$$

Letting $u = 2z$, by the chain rule it is immediate that

$$\lim_{z\to 0}\frac{d^n}{dz^n}\left(\frac{2z}{e^{2z}-1}\right) = \lim_{u\to 0}2^n\frac{d^n}{du^n}\left(\frac{u}{e^u-1}\right). \tag{12}$$

Substituting (12) into (11) and rearranging, we have

$$\lim_{z\to 0}\frac{d^n}{dz^n}\left(\frac{z}{e^z-1}\right)=\frac{-1}{2^n-1}\frac{d^n}{dz^n}\left(\frac{z}{e^z+1}\right)\Big|_{z=0}. \tag{13}$$

It is easy to check that any infinitely differentiable function f satisfies

$$\frac{d^n}{dz^n}\big(zf(z)\big)=z\frac{d^n}{z^n d}f(z)+n\frac{d^{n-1}}{dz^{n-1}}f(z). \tag{14}$$

Applying (14) to the function $f(z)=(e^z+1)^{-1}$ and substituting this result into (13) gives Laplace's result.

To obtain the Maclaurin series for $x\cot x$, we express this function in the form $x\cot x = 2ix/(e^{2ix}-1)+ix$. If we let $z=2ix$, it is obvious that

$$\lim_{x\to 0}\frac{d^n}{dx^n}(x\cot x)=(2i)^n\lim_{z\to 0}\frac{d^n}{dz^n}\left(\frac{z}{e^z-1}\right) \tag{15}$$

for $n\geq 2$. By (8) and (10) we see that

$$\lim_{x\to 0}\frac{d^n}{dx^n}(x\cot x)=\frac{-(2i)^n n}{2^n-1}\sum_{k=1}^{n-1}\frac{(-1)^k k!}{2^{k+1}}\left\{{n-1 \atop k}\right\}. \tag{16}$$

Since $x\cot x$ is real, the right side of (16) equals zero when n is odd. We conclude that the Maclaurin series for $x\cot x$ is

$$1+\sum_{n=1}^{\infty}\frac{(-1)^{n+1}2^{2n}nx^{2n}}{(2^{2n}-1)(2n)!}\sum_{k=1}^{2n-1}\frac{(-1)^k k!}{2^k}\left\{{2n-1 \atop k}\right\}.$$

Maclaurin series for $x\csc x$ In order to obtain the Maclaurin series for $x\csc x$, we express this function as $x\csc x = ix/(e^{ix}+1)+ix/(e^{ix}-1)$. The reader can adapt our techniques to verify that the Maclaurin series for $x\csc x$ is

$$1+\sum_{n=1}^{\infty}\frac{(-1)^n(2^{2n}-2)nx^{2n}}{(2^{2n}-1)(2n)!}\sum_{k=1}^{2n-1}\frac{(-1)^k k!}{2^k}\left\{{2n-1 \atop k}\right\}.$$

Maclaurin series for $\sec x$ The techniques used above would give the Maclaurin series for the secant function, but the nth coefficient of the series would involve the awkward quantity $\cos((k+1)\pi/4)$ for $k=1,\dots,n$. This is the result in Scherk's book [7, p. 18]. Using a somewhat different approach, we will find a cleaner form for its Maclaurin series. We write $\sec x = 2e^{ix}/(e^{2ix}+1)$. Define $g_n(u)=(u\frac{d}{du})^n\big(u/(u^2+1)\big)$ and note that

$$\frac{d^n}{dx^n}(\sec x)=2i^n g_n(e^{ix}) \tag{17}$$

by the chain rule. We will show that

$$g_n(u)=\sum_{k=0}^{n}a_{nk}u^{2k+1}(u^2+1)^{-k-1}, \quad \text{where} \tag{18}$$

$$a_{nk}=\sum_{j=0}^{k}(-1)^j\binom{k}{j}(2j+1)^n \tag{19}$$

for $k = 0, 1, \ldots, n$ and $\binom{k}{j}$ are binomial coefficients. It is trivial to show by induction that (18) is true, where $a_{n0} = 1$ and the a_{nk} are constants, similar to the way we verified the main formula. At least I believe you will find this step trivial. But as some of you may not know, mathematicians are notorious for abusing the term *trivial*. Often *trivial* is not so trivial. The renowned physicist Richard Feynman **[4]**, after listening to the interchange between mathematicians during tea time at Princeton, concluded that *trivial* simply means proved.

To prove that the a_{nk} satisfy equation (19), we first express $u(u^2+1)^{-1}$ in its Maclaurin series form $\sum_{j=0}^{\infty}(-1)^j u^{2j+1}$ and then express $g_n(u)$ as a power series:

$$g_n(u) = \left(u\frac{d}{du}\right)^n \left(\frac{u}{u^2+1}\right) = \sum_{j=0}^{\infty}(-1)^j(2j+1)^n u^{2j+1}. \tag{20}$$

Using the form of $g_n(u)$ given by (18) and writing $(u^2+1)^{-k-1}$ as a Maclaurin series, we can also represent $g_n(u)$ as

$$g_n(u) = \sum_{k=0}^{n} a_{nk} u^{2k+1} \sum_{i=0}^{\infty}(-1)^i \binom{k+i}{k} u^{2i}.$$

Making the change of summation from i to j with $j = k + i$, we obtain

$$g_n(u) = \sum_{j=0}^{n}\sum_{k=0}^{j}(-1)^{j-k}\binom{j}{k} a_{nk} u^{2j+1} + \sum_{j=n+1}^{\infty}\sum_{k=0}^{n}(-1)^{j-k}\binom{j}{k} a_{nk} u^{2j+1}. \tag{21}$$

Setting corresponding coefficients of u^{2j+1} in (20) and (21) equal for $j = 0, 1, \ldots, n$, we have

$$(2j+1)^n = \sum_{k=0}^{j}(-1)^k \binom{j}{k} a_{nk}. \tag{22}$$

We next use the following mathematical tidbit that I would like to thank the referee for bringing to my attention. For any sequences $\{v_j\}_{j=0}^n$ and $\{w_k\}_{k=0}^n$, it is true that

$$w_j = \sum_{k=0}^{j}\binom{j}{k}(-1)^k v_k \text{ for all } j = 0, \ldots, n \Longleftrightarrow$$

$$v_k = \sum_{j=0}^{k}\binom{k}{j}(-1)^j w_j \text{ for all } k = 0, \ldots, n. \tag{23}$$

By the symmetry of the relationship between v_j and w_k, it is enough to prove (23) in just one direction (say $\Rightarrow$). So we assume that the left side of the biconditional (23) is true, and prove that the right side of (23) is true. Observe that

$$\begin{aligned}\sum_{j=0}^{k}\binom{k}{j}(-1)^j w_j &= \sum_{j=0}^{k}\binom{k}{j}(-1)^j \sum_{i=0}^{j}\binom{j}{i}(-1)^i v_i \\ &= \sum_{i=0}^{k}\binom{k}{i} v_i \sum_{j=i}^{k}\binom{k-i}{j-i}(-1)^{j-i} \\ &= v_k + \sum_{i=0}^{k-1}\binom{k}{i} v_i \sum_{j=i}^{k}\binom{k-i}{j-i}(-1)^{j-i}.\end{aligned}$$

By the binomial theorem

$$0 = (1-1)^{k-i} = \sum_{j=i}^{k} \binom{k-i}{j-i} (-1)^{j-i}$$

for $i = 0, \ldots, k-1$ and so (23) follows.

Choosing $w_j = (2j+1)^n$ and $v_k = a_{nk}$ in (23) gives (19), as we wished to show. Evaluating equation (17) at $x = 0$ and referring to (18), we have

$$\frac{d^n}{dx^n}(\sec x)\mid_{x=0} = i^n \sum_{k=0}^{n} a_{nk} 2^{-k}.$$

Hence the Maclaurin series for $\sec x$ is

$$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} \sum_{k=0}^{2n} a_{2n,k} 2^{-k}.$$

On a lighter note, 18th and early 19th century mathematicians were interested in the computational aspects of their formulas. So, they would often include in their papers lists of values of relevant quantities. After having obtained a Maclaurin series for $\sec x$, Euler listed the coefficients for the series up to the term x^{18} [**3**, p. 542]. Scherk discovered that Euler had miscalculated the coefficient for x^{18} [**7**, footnote on p. 7]. Scherk remarks that this value of the coefficient was reproduced in numerous subsequent publications without any of the authors checking Euler's calculations, perhaps because the respect for Euler was so great.

Bernoulli numbers

The Bernoulli numbers B_n, introduced by James Bernoulli in 1713, are widely used in mathematics. Just to mention two examples, an extension of Stirling's formula allows us to compute $n!$ to arbitrarily good precision:

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \exp\left[\sum_{k=1}^{\infty} \frac{B_{2k}}{(2k-1)2kn^{2k-1}}\right]$$

and Euler discovered that the zeta function $\zeta(z) = \sum_{n=1}^{\infty} n^{-z}$ evaluated at the positive even integers $2k$ satisfies

$$\zeta(2k) = \frac{(-1)^{k-1}(2\pi)^{2k} B_{2k}}{2(2k)!}.$$

The Bernoulli numbers are defined by

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n x^n}{n!}. \tag{24}$$

Since the right side of (24) is simply the Maclaurin series for $x/(e^x - 1)$, we must have

$$B_n = \lim_{x \to 0} \frac{d^n}{dx^n} \left(\frac{x}{e^x - 1}\right).$$

Applying (8) and (10), we immediately obtain a formula for computing the Bernoulli numbers:

$$B_n = \frac{n}{2^n - 1}\sum_{k=1}^{n-1} \frac{(-1)^{k+1}k!}{2^{k+1}} \left\{ {n-1 \atop k} \right\}.$$

Acknowledgments. I would like to thank one of the referees for numerous helpful suggestions, especially for introducing me to the rich history of the topics in this paper. I would also like to thank David Rowe and Sheila Hu for their assistance in my historical research.

REFERENCES

1. R. Ayoub, Euler and the zeta function, *Am. Math. Monthly* **81** (1974), 1067–1085.
2. Richard A. Brualdi, *Introductory Combinatorics*, 2nd ed., Elsevier Science Publishing Company, 1992.
3. Leonhard Euler, *Institutiones Calculi Differentialis*, Vol. II, Academie Imperialis Scientiarum Petropolitanae, 1755.
4. Richard P. Feynman, *Surely You're Joking, Mr. Feynman*, W. W. Norton & Company, Inc., New York, 1985.
5. S. F. Lacroix, *Traité du calcul différentiel et du calcul intégral*, 2nd ed., Vol. III, Paris (1810–1819).
6. H. F. Scherk, *De evolvenda functione* $y\partial.y\partial.y\partial\ldots y\partial X/\partial x^n$ *disq. nonnull. analyt.*, Regiomonti, 1824.
7. H. F. Scherk, *Mathematische Abhandlungen*, G. Reimer, Berlin, 1825.
8. I. J. Schwatt, *An Introduction to the Operations with Series*, 2nd ed., Chelsea Publishing Company, New York, 1924.
9. H. W. Turnbull, *The Correspondence of Isaac Newton*, Camb. Univ. Press, Vol. I, 1959.

75 Years Ago in the MAGAZINE
(called, at that time, the *Mathematics News Letter*)

A.C. Maddox, The minimum training prerequisite to the successful pursuit of college mathematical study, Vol. **3**, No. 4, (Dec., 1928), 14–19:

> A student is not studying college mathematics successfully if he habitually commits errors in simple computation, particularly if he usually fails to detect the errors.
>
> He cannot reasonably be regarded as successful in his college mathematical study if he rarely remembers, or if he never deigns, to check the results of his exercise and problem solutions except by reference to answer books.
>
> He is certainly not really successful in college mathematical endeavors as long as he is tenaciously antagonistic toward the subject of mathematics.
>
> He is really not succeeding in the college mathematical field if he never experiences any thrills from mathematical adventures.

Apart from the relentless gender pronouns, which the modern ear finds grating, these seventy-five year old thoughts do not seem particularly out of date.

Symmetric Polynomials in the Work of Newton and Lagrange

GREG ST. GEORGE
University of Montana
Missoula, MT 59812
stgeorge@selway.umt.edu

We all know everything about quadratic equations, don't we? Well maybe, but here is a question which, at least for me, did not have an obvious answer:

Given two quadratic equations,

$$ax^2 + bx + c = 0 \quad \text{and} \quad Ax^2 + Bx + C = 0$$

(where $a \neq 0$ and $A \neq 0$) when do they have a common root? That is, what condition or conditions on the coefficients a, b, c, A, B, and C will assure us that there is a value r such that

$$ar^2 + br + c = Ar^2 + Br + C = 0? \tag{1}$$

The analogous question, for cubics, was asked by Isaac Newton. D. T. Whiteside, in his masterful edition of Newton's papers, dates the work to early 1665, when Newton would have been 22 years old.

Newton introduced his inquiry by saying

"For taking of unknowne quantitys out of intricate Equations it may bee convenient to have severall formes." [**8**, p. 517]

A modern paraphrase of Newton's remark might be that it is valuable to have "several different ways of looking at things." Although we will work with quadratic equations instead of the cubics Newton asked about, the insight gained will be much the same, and the attendant computations will be less tedious. When we are finished you should be able to do the cubic for yourself, if you so desire, and compare your answer to Newton's. In the course of our investigation, we will discover something that is worth much more than a simple solution to our problem. We will find evidence of this in another source, the work of Lagrange on the cubic and quartic equations.

I have taught classes on the history of mathematics three times, and none of the students in any of those classes had encountered this material before. Indeed, I first encountered these ideas in a historical context, reading about the work of Newton. However, seventy or so years ago, when a course called the "Theory of Equations" was still common, this would probably not have been the case. Although the course is all but extinct, I hope that you will find these ideas from it to be natural, accessible, and exciting.

The resultant of two quadratics

Since we may divide our quadratic equations by their leading coefficients, we may as well assume we are starting with the equations

$$f(x) = x^2 + dx + e = 0 \quad \text{and} \quad g(x) = x^2 + Dx + E = 0.$$

Polynomials like these, whose leading coefficient is one, are called *monic* polynomials. Now the equation

$$f(x) = x^2 + dx + e = 0$$

has two roots, which we shall call r_1 and r_2. They may be the same or different, real or complex. We are interested in the case when one or both of these roots is also a root of the other quadratic equation, $g(x) = 0$.

To say that r_1 or r_2 are roots of g means that

$$g(r_1) = r_1^2 + Dr_1 + E = 0 \quad \text{or} \quad g(r_2) = r_2^2 + Dr_2 + E = 0.$$

Can we express these two equations as a single equation? If $g(r_1) = 0$ or $g(r_2) = 0$, we must always have $g(r_1) \cdot g(r_2) = 0$. That is, g will have a root in common with f if and only if

$$(r_1^2 + Dr_1 + E)(r_2^2 + Dr_2 + E) = 0.$$

Multiply this out. We get nine terms, which we combine by the coefficients of g:

$$r_1^2r_2^2 + D(r_1^2r_2 + r_1r_2^2) + E(r_1^2 + r_2^2) + D^2r_1r_2 + DE(r_1 + r_2) + E^2 = 0. \quad (2)$$

We now have a condition for the two quadratic equations, f and g, to share a root. However, it is written in terms of the roots of f and the coefficients of g. We want only coefficients, so we must replace the roots of f with an expression in terms of its coefficients. To say that r_1 and r_2 are roots of f is the same as saying that f factors as

$$f(x) = x^2 + dx + e = (x - r_1)(x - r_2).$$

Multiplying the right-hand side and grouping terms gives

$$x^2 + dx + e = x^2 - (r_1 + r_2)x + r_1r_2 \quad (3)$$

so that by equating coefficients we have

$$r_1 + r_2 = -d \quad (4)$$

and

$$r_1r_2 = e. \quad (5)$$

This shows that we can write the expressions $r_1 + r_2$ and r_1r_2 using the coefficients of $f(x)$. The expressions using r_1 and r_2 in (2) are more complicated. The critical question is: Can we use (4) and (5) to replace all the expressions involving roots r_1 and r_2 in (2) with the coefficients d and e? The answer is yes. You might try it now, and we will do most of the calculations ourselves shortly. However, because this is the heart of the issue, we first would like to discuss the special sort of expressions that appear in (2). Since we want you to think of them as polynomials, we will write them down again, replacing r_1 with p and r_2 with q. These expressions are then

$$p^2q^2,\ p^2q + pq^2,\ p^2 + q^2,\ pq,\ \text{and}\ p + q.$$

What is special about these polynomials? The first thing that you might notice is that in the polynomials with multiple terms, the degree of each term is the same. We call such polynomials *homogeneous*. However, we are more interested in the fact that all these polynomials are *symmetric*. A symmetric polynomial is unchanged by exchanging or interchanging the variables that occur in it. In more precise but less accessible language they are "invariant under permutations of the variables." Here we have only two variables, so there is only one nontrivial permutation possible, exchanging $\boldsymbol{p}$ and $\boldsymbol{q}$. Observe that all of the polynomials in our list are unchanged by this permutation. For instance, for the second polynomial in our list above we get

$$\begin{array}{cccccc} p^2 & q & + & p & q^2 \\ \downarrow & \downarrow & & \downarrow & \downarrow \\ q^2 & p & + & q & p^2. \end{array}$$

So the polynomial $p^2q + pq^2$ acted on by the permutation that exchanges p and q becomes $q^2p + qp^2$, which, of course, is exactly the same. (The polynomial $p^2 + p + q + q^2$ is symmetric but not homogeneous, while the polynomial $p^2 + pq$ is homogeneous without being symmetric, so these two concepts are distinct.)

We would like to write the symmetric expressions occurring in (2), or generally any symmetric polynomial, in terms of certain special symmetric polynomials represented, in the case of two variables, by (4) and (5). These special symmetric polynomials are called the *elementary symmetric polynomials.* In any monic polynomial equation (for example, $x^2 + ax + b = 0$ or $x^3 + ax^2 + bx + c = 0$), we can always express the coefficients (a, b, c, etc.), up to sign, as an elementary symmetric polynomial evaluated at the roots of the monic polynomial.

For example, consider the third degree monic polynomial

$$0 = x^3 - \sigma_1 x_2 + \sigma_2 x - \sigma_3.$$

(Notice the alternating signs and unusual choice of subscripts; the reason for this will be evident shortly.) Such a polynomial should have three roots (which may be the same, real or complex), which we refer to as r_1, r_2 and r_3. Then the polynomial should factor as

$$x^3 - \sigma_1 x_2 + \sigma_2 x - \sigma_3 = (x - r_1)(x - r_2)(x - r_3).$$

If we multiply the right-hand side and compare like terms we find that

$$\begin{aligned} \sigma_1 &= r_1 + r_2 + r_3 \\ \sigma_2 &= r_1 r_2 + r_1 r_3 + r_2 r_3 \quad \text{and} \\ \sigma_3 &= r_1 r_2 r_3. \end{aligned}$$

The polynomial expressions on the right-hand side of these equations are the elementary symmetric polynomials for three variables r_1, r_2, and r_3. These polynomials reflect the way the roots are coded into the coefficients. In the case of four variables they would be

$$\begin{aligned} \sigma_1 &= r_1 + r_2 + r_3 + r_4 \\ \sigma_2 &= r_1 r_2 + r_1 r_3 + r_1 r_4 + r_2 r_3 + r_2 r_4 + r_3 r_4 \\ \sigma_3 &= r_1 r_2 r_3 + r_1 r_2 r_4 + r_1 r_3 r_4 + r_2 r_3 r_4 \quad \text{and} \\ \sigma_4 &= r_1 r_2 r_3 r_4. \end{aligned} \tag{6}$$

We hope the pattern is now visible: In n variables the kth elementary symmetric polynomial, σ_k, is the sum of all possible monomials of k *distinct* variables. Thus, σ_k is always homogeneous with each term of degree k.

Let $h(x_1, x_2, \ldots, x_t)$ be a polynomial in the variables $x_1, x_2, \ldots, x_t$. If we replace each variable x_i with an elementary symmetric polynomial σ_i, written in terms of the x_is, then the resulting polynomial

$$h(\sigma_1, \sigma_2, \ldots, \sigma_t)$$

is symmetric in the quantities x_i, messy though it may be. Remarkably, *any* symmetric polynomial in n variables can always be expressed as a polynomial evaluated at the elementary symmetric polynomials. This is the content of the Fundamental Theorem of Symmetric Polynomials. A nice discussion, including a statement and an inductive proof of the theorem appears in Edward's book *Galois Theory* [**4**, p. 9].

Let's see how this works for our problem. As we saw, the coefficients $d = -\sigma_1 = -(r_1 + r_2)$ and $e = \sigma_2 = r_1 r_2$ of $f(x)$, are (with one sign change) just the elementary symmetric polynomials evaluated at the roots. The polynomials occurring in (2) are symmetric, and so the Fundamental Theorem of Symmetric Polynomials assures us that we can write them in terms of the elementary symmetric polynomials, or their proxies, d and e. For example, the expression $r_1^2 r_2 + r_1 r_2^2$ can be written

$$\begin{aligned} r_1^2 r_2 + r_1 r_2^2 &= r_1 r_2 (r_1 + r_2) \\ &= -de. \end{aligned}$$

The only other expression in (2) that presents any difficulty is

$$\begin{aligned} r_1^2 + r_2^2 &= (r_1 + r_2)^2 - 2 r_1 r_2 \\ &= d^2 - 2e. \end{aligned}$$

The complete list of expressions appearing in (2), written in terms of the coefficients of f, is:

$$\begin{aligned} r_1^2 r_2^2 &= e^2 \\ r_1^2 r_2 + r_1 r_2^2 &= -de \\ r_1^2 + r_2^2 &= d^2 - 2e \\ r_1 + r_2 &= -d \\ r_1 r_2 &= e. \end{aligned}$$

With these substitutions, (2) becomes

$$e^2 + D(-de) + E(d^2 - 2e) + D^2(e) + DE(-d) + E^2 = 0,$$

or simplified,

$$e^2 - deD + d^2E - 2eE + eD^2 - dDE + E^2 = 0. \tag{7}$$

The expression on the left is called the *resultant* of the two quadratic polynomials. In fact if you type

```
Resultant[ x^2+d*x+e, x^2+D*x+E1, x]
```

into *Mathematica*, you will obtain the result we have found here (with $E1$ replacing E, which is reserved for 2.71828... in *Mathematica*). Notice that in our version of the resultant, exchanging the capital with the lower case ds and es leaves it invariant. This must be the case since the nature of the problem dictates that the two quadratics should be interchangeable.

Let's test the criterion in (7). The quadratic equations

$$2x^2 + 5x - 3 = (2x - 1)(x + 3) = 0$$

and

$$4x^2 - x - \frac{1}{2} = \left(x - \frac{1}{2}\right)(4x + 1) = 0$$

share the root $x = 1/2$. Therefore the coefficients of these quadratics should make the resultant zero. To test, we divide by the lead coefficients and then substitute $d = 5/2$, $e = -3/2$, $D = -1/4$, and $E = -1/8$ into our formula for the resultant, (7). We obtain

$$\left(-\frac{3}{2}\right)^2 - \left(\frac{5}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{1}{4}\right) + \left(-\frac{5}{2}\right)^2\left(-\frac{1}{8}\right)$$
$$-2\left(-\frac{3}{2}\right)\left(-\frac{1}{8}\right) + \left(-\frac{3}{2}\right)\left(-\frac{1}{4}\right)^2 - \left(\frac{5}{2}\right)\left(-\frac{1}{4}\right)\left(-\frac{1}{8}\right) + \left(-\frac{1}{8}\right)^2$$
$$= \frac{9}{4} - \frac{15}{16} - \frac{25}{32} - \frac{3}{8} - \frac{3}{32} - \frac{5}{64} + \frac{1}{64} = 0.$$

So we have seen that to determine if two quadratic polynomials have a root in common, we need only to evaluate the resultant, a function of their coefficients, without actually having to find the roots of the polynomials. There are general formulas for resultants of any pair of polynomials, and symmetric polynomials lie at the heart of these too.

Another application of symmetric polynomials

Our second application, which we will give in a more succinct manner, occurs in Lagrange's paper on the solution of equations [7, sections 30 and 31], which was a precursor to the work of Galois. We would like to show that if you can solve a general cubic polynomial equation, then you can also solve the quartic. This can be done directly [6] but we will take the more general approach that Lagrange introduced. This work was an important step in introducing the idea of studying polynomial equations by examining permutations of their roots.

We assume that the four roots of the quartic equation

$$x^4 - ax^3 + bx^2 - cx + d = 0 \tag{8}$$

are r_1, r_2, r_3, and r_4, so that the coefficients of the equation are just the elementary symmetric polynomials in those roots, as in (6). Consider the polynomial

$$r_1r_2 + r_3r_4.$$

There are $4! = 24$ possible permutations of the roots $r_1, \ldots, r_4$. We are interested in the effect of these permutations on this polynomial. For instance the permutation

$$\begin{pmatrix} r_1 & r_2 & r_3 & r_4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ r_2 & r_3 & r_1 & r_4 \end{pmatrix}$$

converts $r_1r_2 + r_3r_4$ into

$$r_2r_3 + r_1r_4$$

while the permutation

$$\begin{pmatrix} r_1 & r_2 & r_3 & r_4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ r_3 & r_4 & r_2 & r_1 \end{pmatrix}$$

converts $r_1r_2 + r_3r_4$ into the polynomial

$$r_3r_4 + r_2r_1.$$

Note that the first permutation changed our polynomial into something different, while the second actually didn't change it at all. (For a more detailed analysis of this example and its importance for group theory, see the article by Roth [**11**] in this MAGAZINE.)

We now ask how many different forms this polynomial can take under all the permutations of the four roots? We claim that there are only three:

$$r_1r_2 + r_3r_4, \quad r_1r_3 + r_2r_4, \quad r_1r_4 + r_2r_3.$$

We use these three to form a new, *cubic* polynomial in the variable y:

$$(y - (r_1r_2 + r_3r_4))(y - (r_1r_3 + r_2r_4))(y - (r_1r_4 + r_2r_3)). \tag{9}$$

The important thing about this new polynomial (besides the fact that it is cubic) is that any permutation of the four roots $r_1, \ldots r_4$ will leave it unchanged. That is, the coefficients of this new cubic polynomial will be symmetric polynomials in $r_1, \ldots r_4$. The Fundamental Theorem of Symmetric Polynomials, therefore, assures us that its coefficients can be written in terms of the coefficients of the original quartic. Multiplying out (9) and grouping terms gives us

$$y^3 - \left(\sum r_ir_j\right) y^2 + \left(\sum r_i^2r_jr_k\right) y - \sum r_i^3r_jr_kr_l - \sum r_i^2r_j^2r_k^2$$

where the sums are over all possible different terms of the given form with the indices taken from $\{1, 2, 3, 4\}$ without replacement. As the fundamental theorem assures us, this can be written as

$$y^3 - by^2 + (ac - 4d)y - a^2d + 4bd - c^2 = 0. \tag{10}$$

(This is easier to check than to work out—for a derivation of surpassing clarity, read Lagrange's version [**7**].) Lagrange showed that if we know any one of the roots of this cubic, then we can find all the roots of the quartic. For a somewhat different approach, read Kiernan [**5**].

Let's suppose that we know $y = r_1r_2 + r_3r_4$. Recall that from the original quartic, (8), we also know the coefficient $d = r_1r_2r_3r_4$. Now in a (monic) quadratic equation the coefficient of the middle (linear) term is the negative of the sum of the roots and the constant coefficient is the product of the roots (look back at (3)). So we can view r_1r_2 and r_3r_4 as the roots of the quadratic equation

$$t^2 - yt + d = 0. \tag{11}$$

Assume that we have solved this and that we know the two solutions

$$t' = r_1r_2 \quad \text{and} \quad t'' = r_3r_4.$$

We also know, from the coefficient of the x term in the quartic, that

$$\begin{aligned} c &= r_1r_2r_3 + r_1r_2r_4 + r_1r_3r_4 + r_2r_3r_4 \\ &= r_1r_2(r_3+r_4) + r_3r_4(r_1+r_2) \\ &= t'(r_3+r_4) + t''(r_1+r_2) \end{aligned} \tag{12}$$

Using the coefficient of the x^3 term of the quartic, we find

$$a = r_1 + r_2 + r_3 + r_4 \quad \Rightarrow \quad r_3 + r_4 = a - (r_1 + r_2)\,. \tag{13}$$

If $t' \neq t''$, then we can rewrite (12) as

$$\begin{aligned} c &= t'(a - (r_1+r_2)) + t''(r_1+r_2) \\ c - t'a &= (r_1+r_2)\left(t'' - t'\right) \\ r_1 + r_2 &= \frac{c - t'a}{t'' - t'}. \end{aligned} \tag{14}$$

(If $t' = t''$, the roots can be extracted from the equations for a and b, in much the same way as we proceed to do for the general case. Details are left to the reader.) This shows that we can compute the sum $r_1 + r_2$; since we also know the product $r_1r_2 = t'$, we can find r_1 and r_2 by solving the quadratic equation

$$X^2 - \frac{c - t'a}{t'' - t'}\,X + t' = 0. \tag{15}$$

Now that we know $r_1 + r_2$, (13) will give us $r_3 + r_4$. Since we also know the product $t'' = r_3r_4$, we can again solve a quadratic equation to find the last two roots.

As an example, consider the quartic equation

$$x^4 - 2x^3 - 13x^2 + 14x + 24 = 0$$

so that, $a = 2$, $b = -13$, $c = -14$, and $d = 24$. The associated cubic, (10), is

$$y^3 + 13y^2 - 124y - 1540 = 0.$$

We are assuming that we can find at least one root of this cubic. In general, this would involve some work, but, since I created this example, let's just take $y = 11$. Our quadratic, (11), becomes

$$t^2 - 11t + 24 = 0,$$

which has solutions $t'' = 8$ and $t' = 3$. Therefore, the sum of the roots $r_1 + r_2$ (from (14)) is

$$\frac{c - t'a}{t'' - t'} = \frac{-14 - (3)(2)}{8 - 3} = -4,$$

so that the quadratic equation for r_1 and r_2, (15), is

$$X^2 + 4X + 3 = 0.$$

Solving this, we have $r_1 = -1$ and $r_2 = -3$. Substituting these into (13), we find that

$$r_3 + r_4 = -2 - (-1 + -3) = 6.$$

So the quadratic equation for r_3 and r_4 is

$$X^2 - 6X + 8 = 0,$$

which we solve to obtain $r_3 = 2$ and $r_4 = 4$.

We do not claim that this is a practical way of solving quartic equations. The value is theoretical. By using permutations of the roots and properties of symmetric polynomials, Lagrange elucidated the internal machinery of the quartic, showing how its solution is related to that of the cubic. This new way of attacking polynomial equations helped prepare the way for the triumphs of Galois and is a significant landmark on the road to modern algebra.

Acknowledgments. Substantial improvements in both form and content are due to several anonymous referees. The author offers his sincere thanks.

REFERENCES

1. Maxime Bôcher, *Introduction to Higher Algebra,* Macmillan, New York, 1907. (also Dover Publications, New York, 1964.)
2. David Cox, John Little, and Donal O'Shea, *Ideals, Varieties, and Algorithms 2nd Ed.* Springer Verlag, New York, 1997.
3. Floyd Fisk Decker, *The Symmetric Function Tables of the Fifteenthic*, Carnegie Institution of Washington (no. 120), 1910.
4. Harold M. Edwards, *Galois Theory*, Spring-Verlag, Cambridge, New York, 1984.
5. B. M. Kiernan, Galois theory from Lagrange to Artin, *Archive for History of the Exact Sciences* **8**, (1972) 41–153.
6. J. L. Lagrange, *Lectures on Elementary Mathematics*, (translation by Thomas J. McCormack), The Open Court Publishing Co., Chicago 1901.
7. J. L. Lagrange, Réflexions sur la Résolution Algébrique des Équations, in *Oeuvres vol.* **3**, 205–421. (A working translation of the relevant sections, 30 and 31, can be found at the MAGAZINE website).
8. Isaac Newton, *The Mathematical Papers of Isaac Newton*, vol. 1, edited by D. T. Whiteside, Cambridge University Press, Cambridge, 1967.
9. Isaac Newton, *The Mathematical Works of Isaac Newton*, edited by D. T. Whiteside, vol. 2, Johnson Reprint Corp. 1967.
10. E. D. Roe Jr., On symmetric functions, *Amer. Math. Monthly* (in six parts, Jan.-May and June/July 1899).
11. Richard L. Roth, A history of Lagrange's Theorem on groups, this MAGAZINE, **74** (2001), 99–108.

---

75 Years Ago in the MAGAZINE
(called, at that time, the *Mathematics News Letter*)

Editorial comment by S. T. Sanders, Vol. **3**, No. 4, (Dec., 1928), p. 2:

Aggressive Mathematics

The rich disciplines of properly directed mathematical study are real enough to the mathematician. But the genuine lover and student of mathematics is prone to preoccupation with the object of his affection. Too seldom is he fired with a missionary zeal to tell the story of these disciplines to those who know them not. Much less disposed is he to challenge the attitudes of the numberless modern youths whose hurried social programs are inconsistent with the rigid exactions of mathematical study. Thus is there need of a certain militancy of attitude on the part of the professional mathematical worker to all influences which are operating to discount mathematics.

NOTES

Deconstructing Bases: Fair, Fitting, and Fast Bases

THOMAS Q. SIBLEY
St. John's University
Collegeville, MN 56321-3000
tsibley@csbsju.edu

Elementary school students wrestling with decimals quickly realize that not all fractions are created equal. While the relatively awkward fraction $17/32$ turns into the modestly nice finite decimal 0.53125, other seemingly simple ones, like $2/3 = 0.66666666\ldots$ or $1/7 = 0.14285714\ldots$, confront us with the sophisticated notion of infinite repeating decimals. Can we find a *finitely fair* base, one in which all fractions have finite representations?

It is natural to start our hunt for a finitely fair base by changing from the familiar base 10 to base b, where b is any integer greater than 1. Unfortunately, the following review of representations base b reveals that for any b some fractions must have infinite repeating representations, while other fractions have finite representations. Recall that $0.a_1a_2a_3\ldots_b = a_1/b + a_2/b^2 + a_3/b^3 + \cdots = \sum_{n=1}^{\infty} a_n/b^n$, where the subscripted b indicates the base and a_n is an integer satisfying $0 \le a_n \le b-1$. A fraction p/q in reduced form has a k-place representation in base b exactly when q divides b^k but doesn't divide b^{k-1}. For example, $32 = 2^5$ divides 10^5 but not 10^4, so base 10 uses five decimal places for $17/32$. If q has a prime factor not in b, the reduced fraction p/q has an infinite repeating representation base b. Since no fixed b has every prime factor, every base has some fractions with infinite repeating representations.

Clearly, a finitely fair base requires something new, a mathematical "deconstruction" of the idea of a base. The postmodern term *deconstruction* describes something mathematicians have done for two centuries: probe a familiar concept more deeply to expose new interpretations and understandings. An initial deconstruction of base in the next section leads to a finitely fair base, called base $\{n!\}$. A further deconstruction in the middle section leads more generally to variable bases, which we use to find bases that "fit" a given real number with a specified representation. The final section critiques bases in yet another way, leading to competing measures for finitely fair bases. The Prime Number Theorem enables us to approximate one of these measures in terms of the other one and so resolve the competition between them. For ease we consider only representations of real numbers between 0 and 1, although the reader is invited to extend these ideas to the integer parts.

A finitely fair base Finite representation for all fractions requires a reinterpretation of the concept of a base. The familiar infinite series $e^x = \sum_{n=0}^{\infty} x^n/n!$ suggests replacing the powers b^n in the denominators of the base with the factorials $n!$. We will denote this new base with the subscript $_{\{n!\}}$. For example, $1/3!$ becomes $0.001_{\{n!\}}$ and $3/5!$ becomes $0.00003_{\{n!\}}$. Then $e = 2 + \sum_{n=2}^{\infty} 1/n!$ has the memorable representation $2 + 0.01111\ldots_{\{n!\}}$ and $5/6$ equals $0.012_{\{n!\}}$.

DEFINITION. By $0.a_1a_2a_3\ldots_{\{n!\}}$ we mean $\sum_{n=1}^{\infty} a_n/n!$, where a_n is an integer satisfying $0 \le a_n \le n-1$.

The first place in base $\{n!\}$ is always a useless 0, but retaining it makes the nth place correspond to $1/n!$. Since q divides $q!$, the representation of p/q never needs more than q places, showing base $\{n!\}$ is finitely fair. Indeed, only when q is a prime or 4 does the reduced fraction p/q need q places.

The condition $0 \le a_n \le n-1$, which corresponds to the condition $0 \le a_i \le b-1$ in base b, ensures that every fraction has a unique finite representation. Suppose, to illustrate the uniqueness of this representation, we try to find a second representation for $1/3! = 0.001_{\{n!\}}$. The most we can put in the fourth place is a 3, but $0.0003_{\{n!\}} = 3/4!$, which is $1/3! - 1/4!$ since $1/3! = 4/4!$. Again, the most we can add in the fifth place is a 4, and $0.00034_{\{n!\}} = 1/3! - 1/5!$. In general $(i-1)/i! = i/i! - 1/i! = 1/(i-1)! - 1/i!$, and so $0.000345\ldots(k-1)_{\{n!\}} = 1/3! - 1/k!$.

Irrationals related to e can have nice representations base $\{n!\}$. Recall the series $\cosh(x) = (e^x + e^{-x})/2 = \sum_{n=0}^{\infty} x^{2n}/(2n)! = 1 + x^2/2! + x^4/4! + \ldots$. Then $\cosh(1) = 1 + 0.010101\ldots_{\{n!\}}$, and similarly $\sinh(1) = 1 + 0.00101010\ldots_{\{n!\}}$. Since $e^{-1} = \cosh(1) - \sinh(1)$ we can use elementary borrowing in base $\{n!\}$ to find the representation of e^{-1}:

$$\begin{aligned} &1.010101010\ldots_{\{n!\}} \\ -\ &1.001010101\ldots_{\{n!\}} \\ =\ &0.002040608\ldots_{\{n!\}} \end{aligned}$$

In base $\{n!\}$ the representation of π starts off $3 + 0.00031565\ldots_{\{n!\}}$. It seems unlikely that the integers in this representation will form a pattern anyone could describe, although that might well be a deep mathematical question. Some might hope to improve on base $\{n!\}$ so that every irrational would have a repeating representation or other easily recognized pattern. However, a cardinality argument shows that not all of the uncountably many real numbers can have recognizable patterns, even in such a lovely base as $\{n!\}$. Whatever the term *recognizable* might mean, such a pattern must at least be describable in finitely many terms. For example, we can describe the pattern $0.00204060\ldots_{\{n!\}}$ for e^{-1} by "the digit in place $2n+1$ for $n \ge 1$ is $2n$ and all other digits are 0." In any language or base, there are only countably many symbols and so only countably many things describable in a finite number of terms [**2**, p. 223]. Thus no matter what base we invent, there will always be uncountably many real numbers with nonrepeating, and even indescribable representations. However, in the next section we see how to find a base fitting any real number with almost any desired representation.

Finding fitting bases For any given real number between 0 and 1 we seek a base with a specified representation of that number. First we need to deconstruct the idea of a base beyond base $\{n!\}$ or base b. Note that the successive denominators in either case $(1, 2, 6, 24, \ldots$ or $b, b^2, b^3, \ldots)$ are multiples of previous denominators. To see the advantage of this property, consider an attempted base $\{\$\}$ built on making change with American coins. We could write $0.a_1a_2a_3a_{4\{\$\}} = a_1/4 + a_2/10 + a_3/20 + a_4/100$, where a_1, a_2, a_3, and a_4 are the number of quarters, dimes, nickels, and pennies, respectively. To make change, we never need more than four pennies, so we restrict $0 \le a_4 \le 4$. Similarly, $0 \le a_3 \le 1$ since two nickels equals a dime. We need up to two dimes to make change, so $0 \le a_2 \le 2$. Even with this restriction, we don't have a unique way to make change since $0.0210_{\{\$\}} = 0.1000_{\{\$\}}$. This ambiguity occurs because we can't divide a quarter evenly into dimes.

DEFINITION. A sequence $\{b_n\}$ of positive integers is a *variable base* provided each b_n divides b_{n+1} and $\lim_{n\to\infty} b_n = \infty$. Define *place ratios* r_n recursively by $r_1 = b_1$ and $r_{n+1} = b_{n+1}/b_n$. We define the *base* $\{b_n\}$ *representation* $0.a_1a_2a_3\ldots_{\{b_n\}}$ to equal $\sum_{n=1}^{\infty} a_n/b_n$, where a_n is an integer satisfying $0 \le a_n \le r_n - 1$.

Base b and base $\{n!\}$ clearly qualify as variable bases: in base b, we have $b_n = b^n$ and $r_n = b$, whereas $b_n = n!$ and $r_n = n$ in base $\{n!\}$. Recall the earlier example showing that $1/3!$ has a unique finite representation in base $\{n!\}$. The keys underlying that argument generalize to the conditions $b_n = r_n b_{n-1}$ and $0 \le a_n \le r_n - 1$, which correspond to the requirements for a variable base representation. Thus for all variable bases $\{b_n\}$ finite representations in base $\{b_n\}$ are unique. Because $\lim_{n\to\infty} b_n = \infty$ every real number between 0 and 1 is approximated by finite representations, and so it has a finite or infinite representation base $\{b_n\}$. Further, because $a_n = 0$ when $r_n = 1$, the following inequalities show that the infinite series for any representation $0.a_1a_2a_3\ldots_{\{b_n\}}$ converges to a real number:

$$\sum_{n=1}^{\infty} \frac{a_n}{b_n} \le \sum_{r_n>1} \frac{r_n - 1}{b_n} < \sum_{r_n>1} \frac{1}{b_{n-1}} < \sum_{j=0}^{\infty} \frac{1}{2^j} = 2.$$

The reader can verify that, in fact, this sum always lies between 0 and 1.

Let's take as our test case π, whose representation in base $\{n!\}$ appears as indescribable as its representation in base 10. Can we find some $\{b_n\}$ so that $\pi = 3 + 0.1111\ldots_{\{b_n\}}$? For a_1 to be 1, $1/b_1 < 0.14159\ldots_{10} < 2/b_1$. Then $8 \le b_1 \le 14$. Once we pick b_1, say $b_1 = 8$, we have $1/b_2 < 0.14159\ldots_{10} - 1/b_1 = 0.01659\ldots_{10} < 2/b_2$. Thus $60 < b_2 < 120$. Further, b_2 must be a multiple of $b_1 = 8$. If we always pick the smallest denominator at each step, we get the base $\{b_n\} = \{8, 8^2, 8^2\cdot 17, 8^2\cdot 17\cdot 19, \ldots\}$, although this sequence seems no more memorable than the digits of π in base 10. Still, Theorem 1 below provides an explicit construction of a base fitting any number between 0 and 1 with a specified infinite representation.

THEOREM 1. *Let $0 < \gamma < 1$ and let $\{a_n\}$ be any sequence of positive integers. Then there is a base $\{b_n\}$ such that $\gamma = 0.a_1a_2a_3\ldots_{\{b_n\}} = \sum_{n=1}^{\infty} a_n/b_n$.*

Proof. We use recursion to construct a base $\{b_n\}$. Because

$$0 < \gamma \quad \text{and} \quad \lim_{k\to\infty} \frac{a_1}{k} = 0,$$

there are values k such that $a_1/k < \gamma$. Let b_1 be the smallest k such that $a_1/k < \gamma < (a_1+1)/k$. Set $\gamma_1 = \gamma - a_1/b_1$, so that $0 < \gamma_1 < 1/b_1 = a_2/(a_2b_1)$. Choose b_2 to be the least multiple of b_1 such that $a_2/b_2 < \gamma_1 < (a_2+1)/b_2$. In general, suppose we have b_i for $1 \le i \le n$ such that b_i divides b_{i+1} for $i < n$ and

$$\sum_{i=1}^{n} \frac{a_i}{b_i} < \gamma < \left(\sum_{i=1}^{n} \frac{a_i}{b_i}\right) + \frac{1}{b_n}.$$

For $\gamma_n = \gamma - \sum_{i=1}^{n} a_i/b_i$ we have $0 < \gamma_n < 1/b_n = a_{n+1}/(b_n a_{n+1})$. Let b_{n+1} be the least multiple of b_n such that $a_{n+1}/b_{n+1} < \gamma_n < (a_{n+1}+1)/b_{n+1}$. Clearly b_n properly divides b_{n+1} so $\lim_{n\to\infty} b_n = \infty$ and $\{b_n\}$ is a base. Further,

$$\sum_{i=1}^{n+1} \frac{a_i}{b_i} < \gamma < \left(\sum_{i=1}^{n+1} \frac{a_i}{b_i}\right) + \frac{1}{b_{n+1}} \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{b_{n+1}} = 0.$$

Hence $\sum_{i=1}^{\infty} a_i/b_i = \gamma$. ∎

While the preceding proof chooses the smallest possible b_{n+1} each time, there may be other choices of bases fitting the same representation to γ. For example, by suitably continuing the sequence $\{c_n\} = \{9, 36, 396, 5940, \ldots\}$ we would also have $\pi = 3 + 0.111\ldots_{\{c_n\}}$. Determining how many different bases fit a given real number with a specific representation appears quite difficult. We illustrate the range of possibilities for the number of bases giving a particular representation using perhaps the easiest examples. The reader is invited to fill in the details of the induction proofs.

- There is a unique base $\{b_n\}$ (namely $b_n = 3^n$) such that $1/2 = 0.111\ldots_{\{b_n\}}$.
- There are uncountably many bases $\{b_n\}$ such that $1/3 = 0.111\ldots_{\{b_n\}}$.

Sketch of the proof. For each $k \in \mathbb{N}$ consider the options ($r_{2k-1} = 4$ and $r_{2k} = 4$) and ($r_{2k-1} = 5$ and $r_{2k} = 2$). Using the notation in the preceding proof, either way $\gamma_{2k} = 1/3b_{2k}$, so the next step will have the same options. Since $2^{\mathbb{N}}$ is uncountable, there are uncountably many ways of choosing the values r_n in pairs. (There are other possible bases as well.)

- There is a countable infinity of bases $\{b_n\}$ such that $1/4 = 0.111\ldots_{\{b_n\}}$.

Sketch of the proof. If the first k choices of r_n are all 5 then $\gamma_k = 1/4b_k$ and we have three choices for r_{k+1}, namely 5, 6, and 7. Choosing $r_{k+1} = 5$ leaves these options open for r_{k+2}. However, if $r_{k+1} = 6$, then $r_{k+i} = 3$ for all $i \geq 2$. Similarly, the choice of $r_{k+1} = 7$ forces $r_{k+2} = 2$ and for all $i \geq 3$, $r_{k+i} = 3$. Thus there are countably many times one can choose to deviate from $r_n = 5$. However, once that choice is made, there are no further options.

We abandon the difficulties of fitting bases to desired representations of given numbers. Instead, we return our attention to finitely fair bases, considering competing comparisons of them with base $\{n!\}$.

Fast finitely fair bases To consider alternatives to base $\{n!\}$ we need to determine when a base is finitely fair. Minimally every fraction m/p^k, where p is a prime, needs to be finitely represented. That is, for all primes p and all natural numbers k, there is a natural number n such that p^k divides b_n. Actually this condition also suffices: factor the denominator of m/q into powers of primes $q = p_1^{k_1} \ldots p_j^{k_j}$ and pick b_n to be the maximum of the b_i corresponding to the $p_i^{k_i}$. While base $\{n!\}$ is finitely fair, it seems slow in one way and fast in another. First of all it is slow in that it can need as many as q places to represent p/q. However, its place ratios, $r_n = n$, grow fast and without bound. Each criterion alone seems mathematically uninteresting: we could speed up the way we represent fractions by choosing huge values for the r_n, or we could, on average, slow the growth of the place ratios by taking almost all of the r_n to be 1, sprinkling in the primes just often enough to get infinitely many of each eventually. Combining these two senses of speed leads to a more interesting result. We measure the overall growth of the place ratios r_n with their geometric mean rather than their arithmetic mean because of their multiplicative nature.

DEFINITION. The nth *average place ratio* of $\{b_n\}$ is $\sqrt[n]{b_n} = \sqrt[n]{\prod_{i=1}^{n} r_i}$.

For base b the nth average place ratio is constant: $\sqrt[n]{b^n} = b$. From Stirling's approximation for $n!$ the nth average place ratio for base $\{n!\}$ is $\sqrt[n]{n!} \approx \sqrt[n]{\sqrt{2\pi n}(n/e)^n} > n/e$, which goes to infinity as n does [**4**, 127–128].

DEFINITION. A finitely fair base $\{b_n\}$ is *q-fast* if the representation in base $\{b_n\}$ of a reduced fraction p/q never needs more than q places.

Although base $\{n!\}$ is q-fast, factorials lead to fairly fast growth of the average place ratios. Fortunately, a q-fast base doesn't need factorials. We need $r_2 = 2$, and $r_3 = 3$ to handle halves and thirds by the second and third place, respectively. But we only need $r_4 = 2$ to accommodate fourths in the fourth place since $r_4 = 2$ gives $b_4 = 2 \cdot 3 \cdot 2 = 12$, a multiple of 4. Similarly, $r_6 = 1$ suffices since $b_4 = 12$ is already a multiple of 6. The *slowest q-fast base* increases the denominators b_n only as much as needed. This base has the sequence of smallest place ratios, which are $r_{p^k} = p$ for every power of a prime p^k and $r_n = 1$ otherwise. The following table gives initial values of r_n, b_n and $\sqrt[n]{b_n}$ for this slowest q-fast base.

TABLE 1: Values of r_n, b_n, and $\sqrt[n]{b_n}$ for the slowest q-fast base.

n	1	2	3	4	5	6	7	8	9	10	11
r_n	1	2	3	2	5	1	7	2	3	1	11
b_n	1	2	6	12	60	60	420	840	2520	2520	27720
$\sqrt[n]{b_n}$	1	1.41	1.86	1.9	2.27	1.98	2.37	2.32	2.39	2.19	2.53

While the preceding table suggests the average place ratios $\sqrt[n]{b_n}$ increase slowly but unevenly, surprisingly the sequence $\{\sqrt[n]{b_n}\}$ has a bound related to e.

THEOREM 2. *The slowest q-fast base $\{b_n\}$ has $\sqrt[n]{b_n} < e^{1.105} \approx 3.02$ for all n.*

Proof. In a q-fast base if $n \geq p^k$, for a power of a prime, then p^k must divide b_n. The slowest q-fast base has no extra powers of any p, meaning when we factor b_n into primes, for any prime p there are exactly k factors of p in b_n, where $p^k \leq n < p^{k+1}$. The needed power of a prime p is $\lfloor \log_p n \rfloor$, where $\lfloor x \rfloor$ is the floor function of x, the greatest integer less than or equal to x. Hence $b_n = \prod_{p=2}^{n} p^{\lfloor \log_p n \rfloor}$, where p varies over all primes between 2 and n. Now $\prod_{p=2}^{n} p^{\lfloor \log_p n \rfloor} < \prod_{p=2}^{n} p^{\log_p n}$ and $p^{\log_p n} = n$, so $b_n < n^{p(n)}$, where $p(n)$ is the number of primes less than or equal to n. The Prime Number Theorem gives the approximation $p(n) \approx n/\ln(n)$ with an upper bound of $p(n) < 1.105n/\ln(n)$ [3, p. 830]. Thus $\sqrt[n]{b_n} < (n^{1.105n/\ln(n)})^{1/n} = n^{1.105/\ln(n)} = n^{1.105 \log_n e} = e^{1.105}$. ∎

Remarks: The value $\sqrt[n]{b_n}$ exceeds e for the first time when $n = 19$, giving approximately 2.76. The maximum value of $\sqrt[n]{b_n}$ appears to be just under 2.8 when $n = 31$.

We can readily generalize the idea of the slowest q-fast base. A base $\{b_n\}$ is *βq-fast* if its representation of p/q never needs more than βq places (or 1 place if $\beta q < 1$). For example, the slowest $q/2$-fast base needs $r_1 = b_1 = 6$ to take care of halves and thirds by the first place, $r_2 = 10$ to take care of fourths and fifths by the second place, and so on. The preceding theorem generalizes as follows.

COROLLARY 1. *The slowest βq-fast base $\{b_n\}$ has $\sqrt[n]{b_n} < e^{1.105/\beta}$.*

Proof. We need only replace $b_n < n^{p(n)}$ in the previous proof with $b_n < n^{p(n)/\beta}$. ∎

Thus a suitable fast finitely fair base can slow the growth of the place ratios and simultaneously represent fractions as quickly as desired. Mathematicians derive much pleasure from "deconstructing" a commonly accepted concept to find a more fundamental and general one. Often these generalizations give profound insights and applications. Others, such as fair, fitting, and fast bases, are simply fun.

Acknowledgments. I am indebted to a long forgotten person who spoke at a Boston University graduate school seminar between 1975 and 1978 and gave the original idea of base $\{n!\}$. I thank the referees for drawing my attention to Bergman's article on an irrational base [**1**], another, although unrelated, " deconstruction" of base.

REFERENCES

1. G. Bergman, A number system with an irrational base, this MAGAZINE **31**:2 (1957), 98–110.
2. P. Fletcher and C. W. Patty, *Foundations of Higher Mathematics*, 3rd ed., Brooks Cole, Pacific Grove, CA, 1996.
3. M. Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford U. Press, New York, 1972.
4. M. Woodroofe, *Probability with Applications*, McGraw-Hill, New York, 1975.

Another Look at the Euler-Gergonne-Soddy Triangle

RAYMOND A. BEAUREGARD
E. R. SURYANARAYAN
University of Rhode Island
Kingston, RI 02881
beau@math.uri.edu

We assume that the reader is familiar with such things as the incenter, circumcenter, and orthocenter of a triangle and is ready for more wonders in the same vein. Take any triangle, $\triangle ABC$, and consider the following three lines, as shown in FIGURE 1:

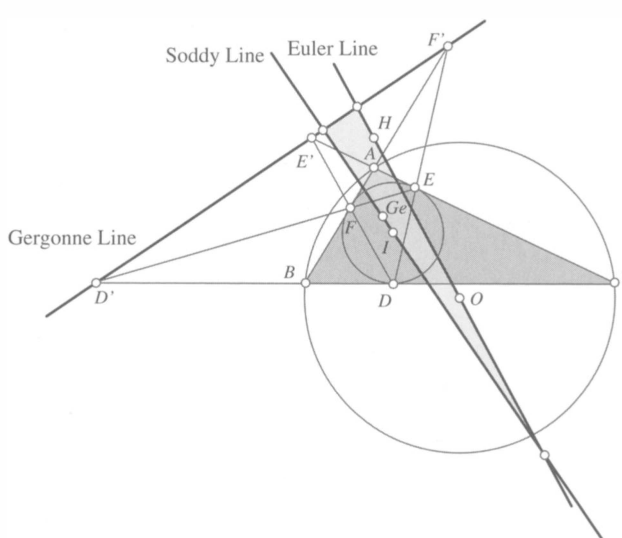


Figure 1 The EGS triangle

- The *Euler line*, which passes through the circumcenter O and the orthocenter H (among other notable points).

- The *Gergonne line* (a term recently coined by Oldknow [5]), which passes through the points D', E', F' in the diagram. These points are obtained from the circle inscribed in $\triangle ABC$, as follows: label its points of contact with the sides as D, E, F, and extend the lines DE, EF, and DF; the points of intersection of these lines with the extended lines AB, BC, and AC are F', D', and E'. That they are always collinear is surprising.
- The *Soddy line*, which passes through the incenter I and the Gergonne point Ge, which is the intersection of the three lines AD, BE, and CF.

Coxeter's book [3] is a good source for those wanting more detail on these and other aspects of a triangle. If the triangle is isosceles then the Euler and Soddy lines coincide. If it is not isosceles, the points of intersection of these three lines form a triangle, the Euler-Gergonne-Soddy (EGS) triangle. In 1996, A. Oldknow [5] described the Gergonne line and showed that it is always perpendicular to the Soddy line, making the EGS triangle a right triangle. This discovery arose from his analysis of the "four-coin problem," which deals with the relationship of three circular discs of assorted sizes, placed in a plane so that they are mutually tangent (each one touching the other two), and a fourth disc, which can then be placed tangentially in the middle region bounded by the first three.

Oldknow used *trilinear coordinates* to avoid messy computations that arise with Cartesian coordinates. The fact that the EGS triangle is always a right triangle suggests the question:

> When is the EGS triangle a Pythagorean triangle, that is, a right triangle with *integer-length* sides?

To attack this problem we need Cartesian coordinates since trilinear coordinates are less than ideal for finding distances. By placing the original triangle strategically in the plane, we can establish Oldknow's result with Cartesian coordinates. We are then able to attack more easily certain numerical problems like the one just mentioned.

In determining conditions under which the EGS triangle of $\triangle ABC$ is Pythagorean, we assume that $\triangle ABC$ itself has integer sides. Although this general problem remains open, we settle the case where $\triangle ABC$ is a right triangle. In particular, we prove that for an infinite number of nonsimilar Pythagorean triangles, the EGS triangle is also Pythagorean, and we show how this problem reduces to a classic question in number theory, which was solved by Aubry in 1910.

Using Cartesian coordinates We place $\triangle ABC$ in a Euclidean plane so that the longest side, which we assume is BC, lies on the x-axis and the altitude from A, whose length is h, lies along the y-axis. It will be convenient to divide sides and altitude by $h/2$ so that the resulting triangle has altitude 2. Then we may label the lengths of the sides as in FIGURE 2 where $0 < p, q \leq 1$, but $pq \neq 1$. To show why this is so, we split the triangle along its altitude and consider right and left halves separately. Thus, it suffices to argue the case of a right triangle with altitude 2 and hypotenuse $a \geq 2$. We let $a = p + 1/p$, where $p = (a - \sqrt{a^2 - 4})/2$, so that $p \leq 1$ (solve the equation $ap = p^2 + 1$ for p). Then computation shows that $-p + 1/p = \sqrt{a^2 - 4}$, the base of the right triangle.

Confirming Oldknow's result We assume that $\triangle ABC$ is not isosceles since the Euler and Soddy lines coincide in that case. Thus, we let p and q be distinct nonnegative real numbers with $p \leq 1$ and $q \leq 1$ and suppose that $\triangle ABC$ has vertices

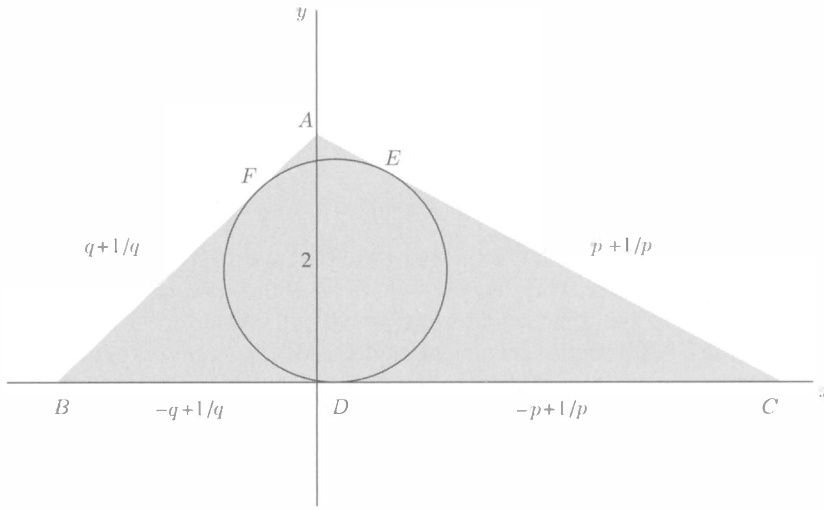


Figure 2 A triangle with altitude 2

$$A = (0, 2), \quad B = \left(-\frac{1}{q} + q, 0\right), \quad C = \left(\frac{1}{p} - p, 0\right).$$

Its area is $\Delta = 1/p + 1/q - p - q$ and its semiperimeter is $s = 1/p + 1/q$. Thus [**3**], the radius of the incircle is Δ/s, which is equal to $1 - pq$, so this is the y-coordinate of I. The x-coordinate of I, which matches the x-coordinate of D is easily found: $BF = BD = -q + 1/q + x$ gives $AF = 2q - x$; similarly, $AE = 2p + x$; thus $AF = AE$ and so $x = q - p$.

This is one example of the straightforward means we use to compute the coordinates of the various points in FIGURE 1. As another example, the point E can be computed as the closest point on the line AC to I. We summarize the results:

$$I = (q - p, 1 - pq) \qquad H = \left(0, \frac{(1-p^2)(1-q^2)}{2pq}\right)$$

$$O = \left(\frac{(q-p)(1+pq)}{2pq}, 1 - \frac{(1-p^2)(1-q^2)}{4pq}\right)$$

$$Ge = \left(\frac{(p+q)^2(q-p)}{1+pq+p^2+q^2}, \frac{2(1-pq)}{1+pq+p^2+q^2}\right)$$

$$D = (q - p, 0) \qquad D' = \left(\frac{p^2+q^2-2}{q-p}, 0\right)$$

$$E = \left((p+q)\frac{1-p^2}{1+p^2}, \frac{2(1-pq)}{1+p^2}\right) \qquad E' = \left(-\frac{(p+q)(1-p^2)}{1-2pq-p^2}, \frac{2(1-pq)}{1-2pq-p^2}\right)$$

$$F = \left(-(p+q)\frac{1-q^2}{1+q^2}, \frac{2(1-pq)}{1+q^2}\right) \qquad F' = \left(\frac{(p+q)(1-q^2)}{1-2pq-q^2}, \frac{2(1-pq)}{1-2pq-q^2}\right)$$

Notice that the denominator $1 - 2pq - p^2$ is nonzero, otherwise $p = -q + \sqrt{q^2+1}$ but then $-p + 1/p = 2q$, so $\triangle ABC$ would have base $q + 1/q$, an isosceles case that we have ruled out. Similarly, $1 - 2pq - q^2 \neq 0$, and of course $q - p \neq 0$.These coordinates yield the following equations for the lines of interest:

$$\text{Euler line:} \quad \left(3(1-p^2)(1-q^2)-4pq\right)x+2(q-p)(1+pq)y = \frac{(q-p)(1+pq)(1-p^2)(1-q^2)}{pq},$$

$$\text{Gergonne line:} \quad (q-p)x+(p^2+q^2+pq-1)y=p^2+q^2-2,$$

$$\text{Soddy line:} \quad (1-p^2-q^2-pq)x+(q-p)y=(q-p)\left(2-(p^2+q^2)\right).$$

We obtained these equations by expressing the two-point form of a line as an equation with determinants, verifying our computations with *Mathematica.* Note that the slope of the Gergonne line is the negative reciprocal of the slope of the Soddy line. This confirms that the EGS triangle has a right angle. A *Mathematica* program accessible at `www.math.uri.edu/~beau` displays the EGS triangle for any given nonisosceles triangle.

Pythagorean triangles The easiest way to consider the right-triangle case is to take $q=1$, so that B lies at the origin. By taking p to be rational, we will obtain a Pythagorean triangle after rescaling. FIGURE 3 shows a right triangle together with its EGS lines. The equations of these lines simplify to the following.

$$\text{Euler line:} \quad y=\frac{2p}{1-p^2}x,$$

$$\text{Gergonne line:} \quad y=-\frac{(1-p)}{p(1+p)}x+\frac{p-1}{p},$$

$$\text{Soddy line:} \quad y=\frac{p(1+p)}{1-p}x+1-2p-p^2.$$

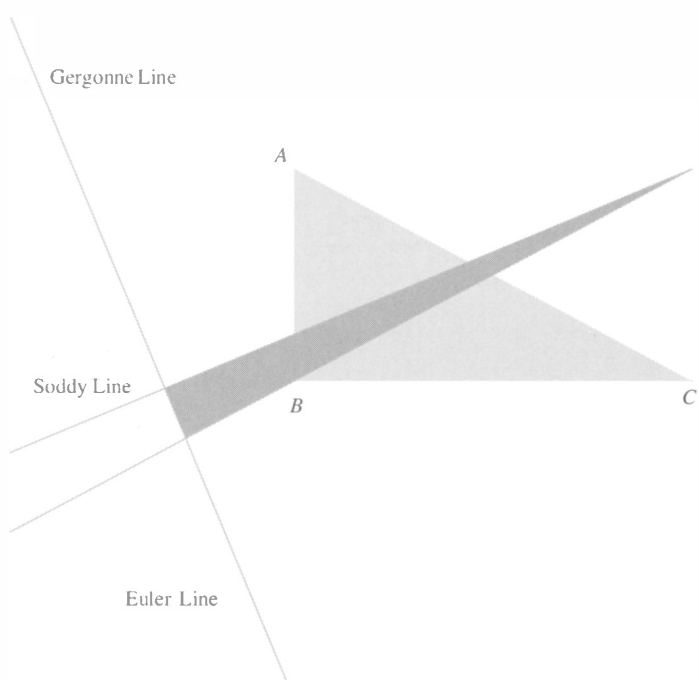


Figure 3 The EGS triangle of a right triangle

Solving these equations two at a time, we find the vertices of the EGS triangle and compute the lengths of the sides. The result is

$$\begin{aligned} &\text{Leg 1:} && \frac{(1+p)(1-p+2p^2)}{p\sqrt{1-2p+2p^2+2p^3+p^4}} \\ &\text{Leg 2:} && \frac{(1-p)|1-2p-p^2|(1-p+2p^2)}{(1-2p+3p^2)\sqrt{1-2p+2p^2+2p^3+p^4}} \\ &\text{Hypotenuse:} && \frac{(1+p^2)(1-p+2p^2)}{p(1-2p+3p^2)}. \end{aligned}$$

The quantity $1-2p-p^2$ appearing between absolute value signs (Leg 2) is negative when $-1+\sqrt{2}<p<1$. In this case the vertical leg of $\triangle ABC$ is longer than the horizontal leg. Notice that $p=-1+\sqrt{2}$ gives the isosceles case which cannot arise since p is rational. The EGS triangle will have rational sides if and only if the radicand in Legs 1 and 2 is a perfect square. Writing $p=m/n$ where m and n are relatively prime integers, the sides become

$$\begin{aligned} &\text{Leg 1:} && \frac{(m+n)(2m^2-mn+n^2)}{m\sqrt{Q}} \\ &\text{Leg 2:} && \frac{(n-m)|m^2+2mn-n^2|(2m^2-mn+n^2)}{n(3m^2-2mn+n^2)\sqrt{Q}} \\ &\text{Hypotenuse:} && \frac{(m^2+n^2)(2m^2-mn+n^2)}{mn(3m^2-2mn+n^2)}, \quad \text{where} \\ & && Q=(n^2+m^2)^2-2nm(n^2-m^2). \end{aligned} \tag{1}$$

If we multiply by the scaling factor mn, $\triangle ABC$ has its sides of $2, n/m-m/n$, $n/m+m/n$ expanded to $r=2nm$, $s=n^2-m^2$, $t=n^2+m^2$ and (r,s,t) is a PT (Pythagorean triple) with parameters n and m. Thus $Q=r^2-rs+s^2$ and the corresponding EGS triangle has rational sides if and only if Q is a square, say $Q=w^2$.

We seek integers r, s satisfying $r^2+s^2=t^2$, $r^2-rs+s^2=w^2$. The smallest pair of positive integers that works is $(r,s)=(8,15)$ with $t=17$ and $w=13$. In order to apply (1) we need the PT parameters of (8, 15, 17). These can be found (following a standard method [**2**]) by computing $n/m=(15+17)/8=4/1$ (thus $n=4$ and $m=1$). The legs of the EGS triangle are obtained using the formulas in (1) multiplied by $mn=4$. They are

$$\text{Leg 1: } \frac{280}{13}, \qquad \text{Leg 2: } \frac{294}{143}, \qquad \text{Hypotenuse: } \frac{238}{11}.$$

If the (8,15,17) triangle is magnified by the scaling factor (13)(143)(11) then its EGS triangle is a Pythagorean triangle. FIGURE 3 actually represents the triangles in this case drawn to scale.

A problem in number theory This square requirement for Q together with the requirement that r^2+s^2 is a square is a special case of the following Diophantine system

$$r^2+s^2=t^2, \qquad r^2+rs+s^2=w^2, \tag{2}$$

where r, s are integers with $r>0$. L. Aubry ([**1, 4**]) showed that there are an infinite number of primitive (components relatively prime) integer pairs (r,s) satisfying system (2). We will refer to a solution (r, s) as positive or negative according as s is positive or negative. Solutions are generated as follows.

Let (r, s) be a primitive solution of (2), and let $r = 2nm$ and $s = n^2 - m^2$ (n and m can be found by taking $(t + s)/r = n/m$ in lowest terms). Next let $(t \pm w)/s = u/v$, reduced to lowest terms. Then the values

$$N = ru^2 + sv^2, \qquad M = su^2 + rv^2$$

give rise to two new primitive solutions (R, S), where $R = 2NM$ and $S = N^2 - M^2$. In this way we obtain an infinite set of (primitive) solutions. For example, the smallest solution to (2) is $(r, s) = (8, -15)$, where $t = 17$ and $w = 13$. This gives rise to the solutions

$$(1768, 2415) \quad \text{with} \quad t = 2993 \quad \text{and} \quad w = 3637,$$
$$(10130640, -8109409) \quad \text{with} \quad t = 12976609 \quad \text{and} \quad w = 9286489.$$

The negative solution represents a new right triangle whose EGS triangle has integer sides; but the positive solution (1768, 2415) does not. However, the positive solutions are important. It is not difficult to show that, starting with a positive solution, Aubry's formulas always produce both a positive and a negative solution thus assuring an infinite number of negative primitive solutions and so an infinite number of incongruent Pythagorean triangles whose EGS triangle is of the same type.

REFERENCES

1. L. Aubry, Untitled, *L'Intermediaire des Math.*, **17** (1910), 283–285.
2. R. A. Beauregard and E. R. Suryanarayan, Pythagorean triples: the hyperbolic view, *College Math. J* **27** (1996), 170–181.
3. H. S .M. Coxeter, *Introduction to Geometry*, Wiley, New York, 1969.
4. L. E. Dickson, *History of the Theory of Numbers*, Vol. II, Chelsea, New York, 1934.
5. A. Oldknow, The Euler-Gergonne-Soddy triangle of a triangle, *Amer. Math. Monthly* **103** (1996), 319–329.

Historical Mathematical Blunders: The Case of Barbaro's Cannonballs

GEORGE M. HOLLENBACK
5401 Rampart St., Apt. 413
Houston TX 77081-1324
gmh616@yahoo.com

When the Ottoman Turks took Constantinople in 1453, their success was due in large part to the breaches made in the city's walls by siege guns that shot enormous granite cannonballs. Nicolo Barbaro, a ship's surgeon attached to the Venetian contingent that helped defend the city, described some of these cannonballs in a diary he kept during the siege; the source of his information is not known. According to Barbaro, the largest Turkish gun shot a ball measuring 13 *quarte* in circumference and weighing **1200** ***livre***, and a smaller gun shot a ball measuring 9 *quarte* in circumference and weighing **800** *livre* [**1**, p. 30; **2**, p. 21]. The diary was published by Cornet in 1856, and an English translation by Jones appeared in 1969; it has since become one of the indispensable sources for siege scholarship.

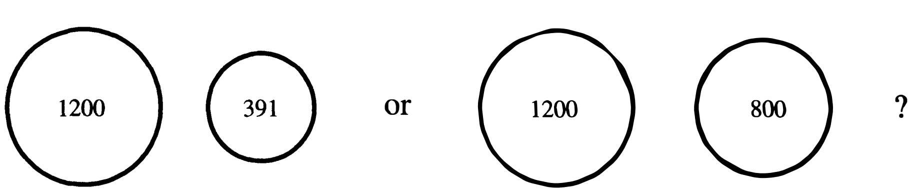
A major problem with Barbaro's description of the cannonballs, however, has gone unnoticed in published commentary on the ordnance used during the siege. If the dimensions and weights of the two cannonballs are reasonably accurate, then their weight to volume densities should closely correspond. This is decidedly not the case, however; in terms of *livre* per cubic *quarta*—rounded to the nearest *livra*—the larger ball would have a density of 32, and the smaller ball would have a density of 65. Which, if either, is correct?

In order to find out, one must know the value of the *quarta*, the value of the *livra*, and the specific gravity of granite. The *quarta*—one quarter of the Venetian *braccio* of 63.9 cm—had a length of nearly 16 cm [**5**, pp. 47, 211]. The *livra* or Venetian "pound" had a weight of 0.301 kg [**5**, pp. 129, 134, 138]. Granite has a specific gravity of 2.66, meaning that it is 2.66 times as heavy as an equal volume of water or, equivalently, that one liter weighs 2.66 kg. Given this information, it is possible to calculate the correct density of 36—much closer to the density of the larger ball than the smaller ball.

There is also a possibility, however, that Barbaro might have been using his Venetian terminology to describe near-equivalent Byzantine weights and measures used in and around Constantinople. The Byzantine "double handbreadth" or *dichas* of 15.62 cm, for example, was a little shorter than the Venetian *quarta* [**3**, pp. 18–19]. The Byzantine "pound" or *litra* of 0.319 kg was a little heavier than the Venetian *livra* [**4**]. The weight to volume density calculated from these Byzantine units and rounded to the nearest *litra* comes out to 32, corresponding to the density derived from Barbaro's figures for the larger ball. It would appear, then, that Barbaro was calling the Byzantine units by Venetian names. The larger ball would have had a circumference of 203 cm and a weight of 376 kg.

What about the smaller ball, though, whose apparent density was impossibly twice that of the larger ball? More than likely, someone knew either the circumference or the weight of the smaller ball, but not both.

Since the known datum for the smaller ball (whether weight or circumference) was about 2/3 that of the corresponding datum for the larger ball, it was erroneously assumed that the unknown datum for the smaller ball would have also been about 2/3 that of the corresponding datum for the larger ball. In reality, if the circumference of a smaller ball is taken as a fraction, k, of the circumference of a larger ball, its volume would be k^3 that of the larger ball; thus if the circumference of the smaller ball is 2/3 that of the larger, its volume would be 8/27 that of the larger. On the other hand, if its volume is 2/3 that of the larger, then its circumference would be about 88% that of the larger. Thus a ball with a known circumference of 9 units might have been attributed a weight of 800 units when its actual weight was only 391 units, or a ball with a known weight of 800 units might have been attributed a circumference of only 9 units when its actual circumference was about 11.4 units. Either mistake would have led to an inflated density for the smaller ball.



REFERENCES

1. Nicolo Barbaro, *Diary of the Siege of Constantinople*, 1453, trans. J. R. Jones, Exposition Press, New York, 1969.

2. Nicolo Barbaro, *Giornale dell' Assedio di Constantinopoli*, 1453, ed. Enrico Cornet, Libreria Tendler & Comp., Vienna, 1856.
3. Erich Schilbach, *Byzantinische Metrologie, Handbuch der Altertumswissenschaft XII.4*, C. H. Beck'sche, Munich, 1970.
4. Erich Schilbach, Litra, *Oxford Dictionary of Byzantium*, Vol. 2, Oxford U. Press, London, 1991.
5. Ronald Edward Zupko, *Italian Weights and Measures from the Middle Ages to the Nineteenth Century*, American Philosophical Society, Philadelphia, 1981.

An *LDU* Factorization in Elementary Number Theory

WARREN P. JOHNSON
Bates College
Lewiston, ME 04240
wjohnson@bates.edu

I want to point out a simple result suitable for introductory courses in linear algebra (where I use it as a source of examples) or number theory. Recall Euler's phi (or totient) function [**1**]: For a positive integer n, $\varphi(n)$ denotes the number of positive integers $\leq n$ and relatively prime to n, where $\varphi(1) := 1$. The only fact we need about $\varphi(n)$ is

$$n = \sum_{d|n} \varphi(d). \tag{1}$$

To keep this note self-contained we sketch a well-known proof of (1). Let $S_d(n)$ denote the subset of $\{1, 2, \ldots, n\}$ whose members each have greatest common divisor d with n, and let $|S|$ denote the cardinality of the set S. Since $S_d(n)$ is empty if d does not divide n, we have

$$n = \sum_{d|n} \left|S_d(n)\right|. \tag{2}$$

The potential members of $S_d(n)$ are the numbers kd, where $k \in \{1, 2, \ldots, n/d\}$. However, kd will be in some $S_e(n)$, where e is a multiple of d, unless k and n/d are relatively prime. Therefore

$$\left|S_d(n)\right| = \varphi\left(\frac{n}{d}\right),$$

and so (2) becomes

$$n = \sum_{d|n} \varphi\left(\frac{n}{d}\right).$$

But this is equivalent to (1), since as d runs through the divisors of n so does n/d.

Against this background, we define several matrices. Let $L(n)$ be the $n \times n$ matrix whose ijth entry is 1 if j divides i, and is 0 otherwise. For example,

$$L(8) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Note that $L(n)$ is always lower triangular with 1s on the main diagonal. Let $\Phi(n)$ be the diagonal $n \times n$ matrix whose ith diagonal entry is $\varphi(i)$. For example,

$$\Phi(8) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

Finally, let $G(n)$ be the $n \times n$ matrix whose ijth entry is the greatest common divisor of i and j. For example,

$$G(8) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 \\ 1 & 1 & 3 & 1 & 1 & 3 & 1 & 1 \\ 1 & 2 & 1 & 4 & 1 & 2 & 1 & 4 \\ 1 & 1 & 1 & 1 & 5 & 1 & 1 & 1 \\ 1 & 2 & 3 & 2 & 1 & 6 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 7 & 1 \\ 1 & 2 & 1 & 4 & 1 & 2 & 1 & 8 \end{pmatrix}$$

It is appropriate to call our result **Le Paige's Theorem**:

$$G(n) = L(n)\Phi(n)L(n)^T. \tag{3}$$

Proof. Since $\Phi(n)$ is diagonal, the ikth entry on the right side of (3) is

$$\sum_{j=1}^{n} \left(L(n)\right)_{ij}\left(\Phi(n)\right)_{jj}\left(L(n)^T\right)_{jk} = \sum_{j=1}^{n} \left(L(n)\right)_{ij}\varphi(j)\left(L(n)\right)_{kj}$$
$$= \sum_{j|(i \text{ and } k)} \varphi(j).$$

According to (1), the last sum equals the greatest common divisor of i and k, which is also the ikth entry on the left side of (3). ∎

As a corollary we get a determinant evaluation of Smith [**5**]:

$$\det G(n) = \det \Phi(n) = \prod_{j=1}^{n} \varphi(j)$$

since $\det L(n) = 1$. In essence (3) is due to Le Paige [**2**], who was looking (as I was) for a natural proof of Smith's theorem; actually he evaluated a generalization of Smith's determinant due to Mansion [3], although he did not state his method as a matrix factorization. See Muir's history [4] for more details.

Acknowledgments. I thank John Rhodes and Paul Campbell for their comments and encouragement.

REFERENCES

1. Leonhard Euler, Theorema arithmetica nova methodo demonstrata, *Nov. comm. acad. sci. Petrop.* **8** (1763), 74–104. Reprinted in *Opera Omnia*, Series Prima, Vol. 2, B. G. Teubner, Leipzig, 1915, 531–555.
2. C. Le Paige, Sur un théorème de M. Mansion, *Nouv. Corresp. Math.* **4** (1878), 176–178.
3. P. Mansion, On an arithmetical theorem of Professor Smith's, *Messenger of Math.* **7** (1877), 81–82.
4. Thomas Muir, *The Theory of Determinants in the Historical Order of Development*, Vol. 3, Macmillan and Co., Limited, London, 1920. Reprinted by Dover Publications, New York, 1960.
5. H. J. S. Smith, On the value of a certain arithmetical determinant, *Proc. London Math. Soc.* **7** (1876), 208–212. Reprinted in J. W. L. Glaisher, editor, *The collected mathematical papers of Henry John Stephen Smith*, Vol. 2, The Clarendon press, Oxford, 1894, 161–165.

A Laplace Transform Technique for Evaluating Infinite Series

JAMES P. LESKO
Grace College
Winona Lake, Indiana 46590
leskojp@grace.edu

WENDY D. SMITH
Bowling Green State University
Bowling Green, OH 43403
wendys@bgnet.bgsu.edu

In an article in this MAGAZINE, Efthimiou [2] shows how the Laplace transform can be used as a tool for evaluating infinite series. We review his method and illustrate a more general application of the technique.

Efthimiou's technique Efthimiou [2] finds closed-form expressions for series of the form

$$\sum_{n=1}^{\infty} \frac{1}{(n+a)(n+b)}, \qquad \text{where} \quad a, b \notin \{-1, -2, -3, \dots\}. \tag{1}$$

He applies the same methods to sum series of the form $\sum_{n=1}^{\infty} Q(n)/P(n)$, where P and Q are polynomials with $\deg(P) - \deg(Q) = 2$ and P factors completely into linear factors with no roots in $\{1, 2, 3, \dots\}$.

Efthimiou's technique applies when $\sum_{n=1}^{\infty} u_n$ is an infinite series whose summand u_n can be realized as a Laplace transform integral $u_n = \int_0^{\infty} e^{-nx} f(x)\, dx$. In such a case, what appeared to be a sum of numbers is now written as a sum of integrals. This may not seem like progress, but interchanging the order of summation and integration (with proper justification of course!) yields a sum that we can evaluate easily, namely, a geometric series. Here are the steps:

$$\sum_{n=1}^{\infty} u_n = \sum_{n=1}^{\infty} \int_0^{\infty} e^{-nx} f(x)\, dx$$
$$\stackrel{?}{=} \int_0^{\infty} f(x) \sum_{n=1}^{\infty} e^{-nx}\, dx = \int_0^{\infty} f(x) \left(\frac{e^{-x}}{1-e^{-x}}\right) dx.$$

If this integral is easy to evaluate, we will have our sum.

To illustrate this technique, consider the series (1) where $a \neq b$; without loss of generality, we can assume that $b > a > -1$. The nth term can be written (use partial fractions) as the Laplace transform integral

$$\frac{1}{(n+a)(n+b)} = \int_0^{\infty} e^{-nx} \left(\frac{e^{-ax} - e^{-bx}}{b-a}\right) dx,$$

giving a starting point for the steps above.

We must justify changing the order of summation and integration. Since the integrands are all nonnegative for $0 < x < \infty$, we can apply the monotone convergence theorem (see, for instance, Folland [**3**, p. 49]) to switch the sum and the integral and obtain

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{(n+a)(n+b)} &= \sum_{n=1}^{\infty} \int_0^{\infty} e^{-nx} \left(\frac{e^{-ax} - e^{-bx}}{b-a}\right) dx \\ &\stackrel{M.C.T.}{=} \int_0^{\infty} \left(\frac{e^{-ax} - e^{-bx}}{b-a}\right) \sum_{n=1}^{\infty} e^{-nx}\, dx \\ &= \frac{1}{b-a} \int_0^{\infty} (e^{-ax} - e^{-bx}) \left(\frac{e^{-x}}{1-e^{-x}}\right) dx \\ &= \frac{1}{b-a} \int_0^1 \frac{u^a - u^b}{1-u}\, du. \end{aligned} \tag{2}$$

As a first example, take $a = 0$ and $b = 1/2$; then (2) yields

$$\sum_{n=1}^{\infty} \frac{1}{n(n+\frac{1}{2})} = 2\int_0^1 \frac{1-u^{\frac{1}{2}}}{1-u}\, du = 2\int_0^1 \frac{1}{1+u^{\frac{1}{2}}}\, du = 4(1-\ln 2).$$

For another example, take $a = 0$ and b a positive integer; then (2) yields the popular telescoping series identity

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n(n+b)} &= \frac{1}{b}\int_0^1 \frac{1-u^b}{1-u}\, du = \frac{1}{b}\int_0^1 \left(1 + u + u^2 + \cdots + u^{b-1}\right) du \\ &= \frac{1}{b}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{b}\right). \end{aligned}$$

We refer the reader to Efthimiou's paper [**2**] to see a rich application of this technique to other series.

A more general technique Efthimiou's technique can be generalized to series of the form $\sum_{n=1}^{\infty} u_n v_n$ where it is convenient to write only v_n as a Laplace transform integral. Again, the series can be written as a sum of integrals, but this time there is a factor of u_n before each integral. If the order of summation and integration can be interchanged (again, with proper justification!), we will need to find an explicit sum for the series $h(x) = \sum_{n=1}^{\infty} u_n e^{-nx}$. If all this works out, we will have

$$\sum_{n=1}^{\infty} u_n v_n = \sum_{n=1}^{\infty} u_n \int_0^{\infty} e^{-nx} f(x)\, dx$$
$$\stackrel{?}{=} \int_0^{\infty} f(x) \left(\sum_{n=1}^{\infty} u_n e^{-nx} \right) dx = \int_0^{\infty} f(x) h(x)\, dx.$$

For example, consider

$$\sum_{n=1}^{\infty} \frac{r^n}{an+b}, \quad \text{where} \quad r \in [-1, 1), \quad a > 0, \quad \text{and} \quad b \geq 0. \tag{3}$$

The trick will be to recognize $1/(an+b)$ as being equal to the Laplace transform integral $= \int_0^{\infty} e^{-nx} (e^{-b/ax})/a\, dx$.

As long as $r \neq -1$, the partial sums of $\sum_{n=1}^{\infty} r^n e^{-nx} (e^{-b/ax})/a$ are dominated above by

$$\sum_{n=1}^{\infty} \left| r^n e^{-nx} \left(\frac{1}{a} e^{-\frac{b}{a}x} \right) \right| = \frac{1}{a} e^{-\frac{b}{a}x} \left(\frac{|r|\, e^{-x}}{1 - |r|\, e^{-x}} \right), \quad \text{where}$$
$$\int_0^{\infty} \frac{1}{a} e^{-\frac{b}{a}x} \left(\frac{|r|\, e^{-x}}{1 - |r|\, e^{-x}} \right) dx \leq \frac{1}{a} \int_0^{\infty} \left(\frac{|r|\, e^{-x}}{1 - |r|\, e^{-x}} \right) dx < \infty.$$

This time, we apply the Lebesgue dominated convergence theorem (again, see Folland [**3**, p. 53]) to switch the order of summation and integration to obtain

$$\sum_{n=1}^{\infty} \frac{r^n}{an+b} = \sum_{n=1}^{\infty} \int_0^{\infty} r^n e^{-nx} (\frac{1}{a} e^{-\frac{b}{a}x})\, dx$$
$$\stackrel{D.C.T.}{=} \int_0^{\infty} \left(\frac{1}{a} e^{-\frac{b}{a}x} \right) \sum_{n=1}^{\infty} r^n e^{-nx}\, dx$$
$$= \frac{1}{a} \int_0^{\infty} e^{-\frac{b}{a}x} \left(\frac{r e^{-x}}{1 - r e^{-x}} \right) dx = \frac{1}{a} \int_0^1 \frac{r u^{\frac{b}{a}}}{1 - ru}\, du.$$

We have established this formula for all $r \in (-1, 1)$; but since both $\sum_{n=1}^{\infty} r^n / an + b$ and (4) exist for $r = -1$, they must also be equal by Abel's theorem (see, for instance, Buck [**1**, p. 279]).

We offer two examples. When $a = 1$ and $b = 0$, (4) yields

$$\sum_{n=1}^{\infty} \frac{r^n}{n} = \int_0^1 \frac{r}{1 - ru}\, du = \ln \left(\frac{1}{1-r} \right),$$

and when $r = -1$, $a = 1$, and $b = \frac{1}{2}$, it gives

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n + \frac{1}{2}} = - \int_0^1 \frac{u^{\frac{1}{2}}}{1+u}\, du = \frac{\pi}{2} - 2.$$

As one more example of this technique, consider

$$\sum_{n=1}^{\infty} \frac{r^n}{(n+a)(n+b)}, \quad \text{where} \quad r \in [-1, 1], \quad \text{and} \quad b > a > -1. \tag{4}$$

We already know how to write $1/((n+a)(n+b))$ as a Laplace transform integral. As long as $r \in (-1, 1)$, the estimation of partial sums is so similar to what we did before that the details are left to the reader. Once more, we apply the Lebesgue dominated convergence theorem to switch the sum and the integral, obtaining

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{r^n}{(n+a)(n+b)} &= \sum_{n=1}^{\infty} r^n \int_0^{\infty} e^{-nx} \left(\frac{e^{-ax} - e^{-bx}}{b-a} \right) dx \\
&\overset{D.C.T.}{=} \int_0^{\infty} \left(\frac{e^{-ax} - e^{-bx}}{b-a} \right) \sum_{n=1}^{\infty} r^n e^{-nx}\, dx \\
&= \frac{1}{b-a} \int_0^{\infty} (e^{-ax} - e^{-bx}) \left(\frac{re^{-x}}{1 - re^{-x}} \right) dx \\
&= \frac{r}{b-a} \int_0^1 \frac{u^a - u^b}{1 - ru}\, du.
\end{aligned} \tag{5}$$

This argument works for $r \in (-1, 1)$, but since the infinite sum and (6) both exist for $r = \pm 1$, they must be equal for $r = \pm 1$ by Abel's theorem as before. So, for instance, when $a = 0$ and $b = 1$, (6) yields

$$\sum_{n=1}^{\infty} \frac{r^n}{n(n+1)} = r \int_0^1 \frac{1-u}{1-ru}\, du = 1 + \left(\frac{1-r}{r} \right) \ln(1-r).$$

Closing remark We do not mean to suggest that the closed form expressions for the series discussed in this paper are new. There are various other ways to evaluate them. Rather, it has been our intention to give exposure to a nice technique for evaluating certain series. It is our hope that you add this technique to your toolbox of tricks for series.

Exercises Enjoy!

1. Let $b \in \{1, 2, 3, \ldots\}$. Follow the steps below to give an alternate proof of the telescoping series identity

$$\sum_{n=1}^{\infty} \frac{1}{n(n+b)} = \frac{1}{b} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{b} \right) :$$

   (i) Use the Laplace transform integral $1/n + b = \int_0^{\infty} e^{-nx} (e^{-bx})\, dx$ to show that

$$\sum_{n=1}^{\infty} \frac{1}{n(n+b)} = \int_0^1 u^{b-1} \ln \left(\frac{1}{1-u} \right) du.$$

   (ii) Show that $\int_0^1 u^{b-1} \ln(\frac{1}{1-u}) du = \frac{1}{b}(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{b})$.

2. Find exact values for the following series, using techniques described in this paper. Check your answers with your favorite computer software package!

(i) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+5)}$ (iv) $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n(n+1)}$

(ii) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n4^n}$ (v) $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n(n+2)}$

(iii) $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2n+1}$ (vi) $\displaystyle\sum_{n=1}^{\infty} \frac{(1/2)^n}{n(n+k)}$, where $k \in \{1, 2, 3, \ldots\}$

REFERENCES

1. R. C. Buck, *Advanced Calculus*, 3rd ed., McGraw-Hill, New York, 1978.
2. C. Efthimiou, Finding exact values for infinite sums, this MAGAZINE **72** (1999), 45–51.
3. G. Folland, *Real Analysis: Modern Techniques and Their Applications*, John Wiley & Sons, New York, 1984.

Answers to selected exercises from Wetzel's article *Fits and Covers*

Exercise 1 The triangle T fits into D precisely when

$$d \geq \begin{cases} \dfrac{2abc}{\sqrt{(a+b+c)(-a+b+c)(a-b+c)(a+b-c)}} & \text{if } b^2+c^2 \geq a^2 \\ a & \text{if } b^2+c^2 \leq a^2. \end{cases}$$

Exercise 6

(a) The disk of diameter $\sqrt{a^2+b^2}$.

(b) The disk of diameter $2s$.

(c) The disk of diameter $p/2$.

Exercise 8 The disk of diameter 1 whose center lies at the midpoint of the curve obviously covers the curve; but so does the disk with center at the midpoint of the line segment that joins the endpoints (cf. Exercise 6(c)).

Exercise 11 Its sides are $1/(2\sqrt{3})$ and $1/\sqrt{6}$.

PROBLEMS

ELGIN H. JOHNSTON, *Editor*
Iowa State University

Assistant Editors: RĂZVAN GELCA, Texas Tech University; ROBERT GREGORAC, Iowa State University; GERALD HEUER, Concordia College; VANIA MASCIONI, Ball State University; PAUL ZEITZ, The University of San Francisco

Proposals

To be considered for publication, solutions should be received by May 1, 2004.

1681. *Proposed by Mihai Manea, Princeton University, Princeton, NJ.*

Let p be a prime number. Prove that the polynomial

$$x^{p-1} + 2x^{p-2} + 3x^{p-3} + \cdots + (p-1)x + p$$

is irreducible in $\mathbb{Z}[x]$.

1682. *Proposed by Paul Bracken, University of Texas, Edinburg, TX.*

Let n be a positive integer. Prove that

$$\sum_{j=1}^{n} \frac{1}{\binom{n}{j}} = \frac{n+1}{2^n} \sum_{j=0}^{n-1} \frac{2^j}{j+1}.$$

1683. *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada.*

For integer $n \geq 2$ and nonnegative real numbers $x_1, x_2, \ldots, x_n$, define

$$A_n = (x_1^2 + x_2^2)(x_2^2 + x_3^2) \cdots (x_n^2 + x_1^2)/2^n,$$
$$B_n = (x_1x_2 + x_2x_3 + \cdots + x_nx_1)^n/n^n.$$

(a) Determine all n, if any, such that $A_n \geq B_n$ for all choices of x_k's.
(b) Determine all n, if any, such that $B_n \geq A_n$ for all choices of x_k's.

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 Elgin Johnston, Problems Editor, Department of Mathematics, Iowa State University, Ames IA 50011, or mailed electronically (ideally as a LaTeX file) to ehjohnst@iastate.edu. All communications should include the readers name, full address, and an e-mail address and/or FAX number.

1684. *Proposed by Ethan S. Brown, Massachusetts Institute of Technology, Cambridge, MA., and Christopher J. Hillar, University of California, Berkeley, CA.*

Let S be the set of all words of length n in two letters, say a and b. Define an equivalence relation on S as follows: given a word W, the reverse of W, the complement of W (that is, change all as to bs and all bs to as) and the reverse of the complement are all equivalent to W. Find the number of equivalence classes of S that do not contain any palindromes.

1685. *Proposed by Michel Bataille, Rouen, France.*

Find all continuous functions $f : (0, \infty) \longrightarrow (0, \infty)$ such that

$$f(a) + f(b) + f(c) = f(a)f(b)f(c)$$

whenever $ab + bc + ca = 1$.

Quickies

Answers to the Quickies are on page 404.

Q935. *Proposed by Götz Trenkler, University of Dortmund, Dortmund, Germany.*

Let A and B be two $n \times n$ idempotent matrices with complex entries and satisfying

$$A + B + AB + BA = 0.$$

What can be said about A and B?

Q936. *Proposed by Mowaffaq Hajja, Yarmouk University, Irbid, Jordan, and Peter Walker, American University of Sharjah, United Arab Emirates.*

Let A be the family of all acute triangles and let T_k be the family of triangles whose side lengths a, b, c satisfy

$$a + b \geq kc, \qquad b + c \geq ka, \qquad \text{and} \qquad c + a \geq kb.$$

Prove that there is no k for which $T_k = A$, and find the smallest k for which $T_k \subseteq A$.

Solutions

One Fibonacci or Two? **December 2002**

1657. *Proposed by Elias Lampakis, Messina, Greece.*

Let f_j denote the jth Fibonacci number ($f_1 = f_2 = 1$ and $f_{n+2} = f_{n+1} + f_n$.) Prove that for each positive integer n, there is a positive integer k and integers $a_1, a_2, \ldots, a_k \in \{1, 2\}$ such that $n = \sum_{j=1}^{k} a_j f_j$.

Solution by JPV Abad, San Francisco, CA., Walnut Creek, CA., and Charlottesville, VA.

We prove that if $s_1, s_2, \ldots$ is a sequence of positive integers with $s_1 = 1$ and $1 + \sum_{j=1}^{m} s_j \geq s_{m+1}$, then for each positive integer n, there is a positive integer k and integers $a_1, a_2, \ldots, a_k \in \{1, 2\}$ with $n = \sum_{j=1}^{k} a_j s_j$. This will establish the assertion in the problem statement because $f_1 = 1$ and $1 + \sum_{j=1}^{m} f_j = 1 + (f_{m+2} - 1) =$

$f_{m+1} + f_m \geq f_{m+1}$.

We first use induction to prove that if n is a nonnegative integer with $n \leq \sum_{j=1}^{m} s_j$, then there are integers $b_1, b_2, \ldots, b_m \in \{0, 1\}$ with $\sum_{j=1}^{m} b_j s_j = n$. This is clearly true for $n = 0, 1$. Assume that the result is true for all nonnegative $n \leq \sum_{j=1}^{r} s_j$. If $\sum_{j=1}^{r+1} s_j \geq n \geq 1 + \sum_{j=1}^{r} s_j$, then $\sum_{j=1}^{r} s_j \geq n - s_{r+1} \geq 0$, so by the induction hypothesis $n = \sum_{j=1}^{r} b_j s_j + s_{r+1}$ for some $b_1, b_2, \ldots, b_r \in \{0, 1\}$.

Now let n be a positive integer. Find k such that $\sum_{j=1}^{k} s_j \leq n \leq -1 + \sum_{j=1}^{k+1} s_j$. Then $n - \sum_{j=1}^{k} s_j \leq \sum_{j=1}^{k} s_j$, so there are integers $b_1, \ldots, b_k \in \{0, 1\}$ with

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

Also solved by Tsehaye Andebrhan (Eritrea), Kristin Bailey, Roy Barbara (Lebanon), Michel Bataille (France), D. Bednarchak, David Bendall, Robert Bernstein, Jany C. Binz (Switzerland), Jean Bogaert (Belgium), Keith Brandt and Margaret Richey, J. L. Brown Jr., Bucknell Problems Group, Mario Catalani (Italy), Eddie Cheng, John Christopher, Wenchang Chu and Gabriella Zammillo (Italy), Con Amore Problem Group (Denmark), David Constantine, Chip Curtis, Knut Dale (Norway), M. N. Deshpande (India), Charles R. Diminnie, Daniele Donini (Italy), Tim Flood, Ovidiu Furdui, Natalio H. Guersenzvaig (Argentina), Tracy D. Hamilton, Peter Hohler (Switzerland), Enkel Hysnelaj (Australia), Tom Jager, Casey Jefferies and David Flaspohler, Richard Johnsonbaugh, Lenny Jones, Stephen Kaczkowski, Achim Kehrein (Germany), Jason Lee, Kenneth Levasseur, Kathleen E. Lewis, Carl Libis, S. C. Locke, Bill Mixon, Gert V. Morzé (Germany), Valerian M. Nita, Chris Nolen and Dana Steele and Crystal Mack and Sahar Rashidi, Giray Ökten, Jeff Ovall, Gary Phelps, Rob Pratt, B. Ravikumar, Phillip P. Ray, Rolf Richberg (Germany), Nelisa Roach and Farley Mawyer, Jeremy Rouse, David Rowe, Ossama A. Saleh and Stan Byrd, Edward Schmeichel, R. P. Sealy (Canada), Harry Sedinger, Heinz-Jürgen Seiffert (Germany), Jorge Nuno Silva (Portugal), Achilleas Sinefakopoulos, Nicholas C. Singer, Helen Skala, Skidmore College Problem Group, Richard M. Smith, Ronald L. Smith, John Spellmann and Carrroll Bandy, Philip D. Straffin, Christopher N. Swanson, H. T. Tang, Nicola Tarasca (Italy), R. S. Tiberio, Dave Trautman, Daniel Treat, Xiaoshen Wang, William P. Wardlaw, Darren D. Wick, Michael Woltermann, Naveed Zaman, Li Zhou, and the proposer. There was one solution with no name.

A Polynomial Identity **December 2002**

1658. *Proposed by William Gasarch, Department of Computer Science, University of Maryland, College Park, MD.*

Let n and d be integers with $0 \leq d < n$. Find all polynomials P of degree d such that

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} P(k) = \sum_{k=0}^{\infty} (-1)^k \frac{P(k)}{k!}.$$

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

We first prove that the expression on the left side of the equation is zero. To establish this it suffices to prove that $\sum_{k=0}^{n} (-1)^k \binom{n}{k} k^m = 0$ for $0 \leq m < n$. The sum is equal to $(-1)^n n!\, S_2(m, n)$, where $S_2(m, n)$ is the Stirling number of the second kind, defined as the number of ways of partitioning a set of m elements into n nonempty sets. Thus, if $m < n$, then $S_2(m, n) = 0$. Hence $\sum_{k=0}^{n} (-1)^k \binom{n}{k} k^m = 0$.

It now suffices to characterize those polynomials P for which $\sum_{k=0}^{\infty} (-1)^k \frac{P(k)}{k!} = 0$. This sum can be readily evaluated if P is expressed as a linear combination of falling factorials of x, defined by

$$\begin{aligned} &(x)_0 = 1, \\ &(x)_m = x(x-1)\cdots(x-m+1), \qquad m = 1, 2, \ldots. \end{aligned} \tag{1}$$

Assume that P is a polynomial of degree $d < n$ and that $P(x) = a_0(x)_0 + a_1(x)_1 + \cdots + a_d(x)_d$. Because

$$\sum_{k=0}^{\infty}(-1)^k\frac{(k)_m}{k!} = \sum_{k=m}^{\infty}(-1)^k\frac{1}{(k-m)!} = \sum_{j=0}^{\infty}(-1)^{j+m}\frac{1}{j!} = \frac{(-1)^m}{e},$$

it follows that

$$\sum_{k=0}^{\infty}(-1)^k\frac{P(k)}{k!} = \frac{1}{e}\sum_{m=0}^{d}(-1)^m a_m.$$

Thus a necessary and sufficient condition that the identity in question be satisfied is

$$P(x) = a_0(x)_0 + a_1(x)_1 + \cdots + a_d(x)_d \quad \text{with} \quad a_0 - a_1 + a_2 - \cdots + (-1)^d a_d = 0.$$

Also solved by Tsehaye Andebrhan (Eritrea), Michel Bataille (France), Jean Bogaert (Belgium), Mario Catalani (Italy), John Christopher, Con Amore Problem Group (Denmark), Chip Curtis, Knut Dale (Norway), Daniele Donini (Italy), Marty Getz and Dixon Jones, Julien Grivaux (France), Natalio H. Guersenzvaig (Argentina), James C. Hickman, Elias Lampakis (Greece), McDaniel College Problems Group, Jeremy Rouse, Achilleas Sinefakopoulos, Nicholas C. Singer, Albert Stadler (Switzerland), Chu Wenchang and Di Claudio Leontina Veliana (Italy), Li Zhou, and the proposer.

A Cyclic Pentagon **December 2002**

1659. *Proposed by Erwin Just (Emeritus), and Norman Schaumberger (Emeritus), Bronx Community College of the City of New York, Bronx, NY.*

Assume that pentagon $A_1A_2A_3A_4A_5$ is cyclic and that $\overline{A_1A_2}$ is parallel to $\overline{A_3A_4}$. Prove that

$$\frac{(A_1A_3)^2 - (A_2A_3)^2}{(A_3A_4)^2} = \frac{(A_5A_2)^2 - (A_5A_1)^2}{(A_5A_3)^2 - (A_5A_4)^2}.$$

I. *Solution by Philip D. Straffin, Beloit College, Beloit, WI.*

Let $s_{ij} = A_iA_j$. Ptolemy's Theorem applied to the cyclic quadrilateral $A_2A_3A_4A_5$ gives

$$s_{24}s_{35} = s_{23}s_{45} + s_{34}s_{25}, \tag{1}$$

and applied to the cyclic quadrilateral $A_1A_3A_4A_5$ gives

$$s_{14}s_{35} = s_{13}s_{45} + s_{34}s_{15}. \tag{2}$$

Because $A_1A_2A_3A_4$ is cyclic and A_1A_2 is parallel to A_3A_4, we have $s_{24} = s_{13}$ and $s_{14} = s_{23}$. Substituting these into (1) and (2) results in the two equations

$$s_{13}s_{35} - s_{23}s_{45} = s_{34}s_{25}$$
$$s_{23}s_{35} - s_{13}s_{45} = s_{34}s_{15}.$$

Subtracting the square of the second equation from the square of the first gives

$$(s_{13}^2 - s_{23}^2)(s_{35}^2 - s_{45}^2) = s_{34}^2(s_{25}^2 - s_{15}^2),$$

which is equivalent to the equality to be proved.

II. *Solution by Jany C. Binz, Bolligen, Switzerland.*

We may assume that the vertices of the pentagon lie on the unit circle in the complex plane, and rotate the figure so that $A_1 = p$, $A_2 = \overline{p}$, $A_3 = \overline{q}$, $A_4 = q$, and $A_5 = r$. Then

$$\frac{(A_1A_3)^2 - (A_2A_3)^2}{(A_3A_4)^2} = \frac{(\overline{q}-p)(q-\overline{p}) - (\overline{q}-\overline{p})(q-p)}{(q-\overline{q})(\overline{q}-q)} = \frac{p-\overline{p}}{q-\overline{q}},$$

and

$$\frac{(A_5A_2)^2 - (A_5A_1)^2}{(A_5A_3)^2 - (A_5A_4)^2} = \frac{(\overline{p}-r)(p-\overline{r}) - (p-r)(\overline{p}-\overline{r})}{(\overline{q}-r)(q-\overline{r}) - (q-r)(\overline{q}-\overline{r})} = \frac{p-\overline{p}}{q-\overline{q}}.$$

This proves the desired equality. The result does not depend on r, that is, A_5 can be any point on the unit circle except for the midpoint of arc A_1A_2 or the midpoint of arc A_3A_4. Nonconvex or degenerate pentagons are admissible.

Also solved by Herb Bailey, Michel Bataille (France), Jean Bogaert (Belgium), Chip Curtis, M. N. Deshpande (India), Daniele Donini (Italy), Ragnar Dybvik (Norway), Ovidiu Furdui, Marty Getz and Dixon Jones, Enkel Hysnelaj (Australia), Victor Kutsenok, Neela Lakshmanan (India), Elias Lampakis (Greece), Ralph Rush, Achilleas Sinefakopoulos, SMSU Problem Solving Group, Li Zhou, and the proposer.

A Triangle Inequality **December 2002**

1660. *Proposed by Michel Bataille, Rouen, France.*

In triangle ABC, let $a = BC$, $b = CA$, and $c = AB$. Prove that

$$\frac{b+c}{a^2}\cos A + \frac{c+a}{b^2}\cos B + \frac{a+b}{c^2}\cos C \geq \frac{9}{a+b+c}.$$

Solution by Li Zhou, Polk Community College, Winter Haven, FL.

Without loss of generality, we may assume that $c \geq b \geq a$. Then $\frac{1}{a^2} \geq \frac{1}{b^2} \geq \frac{1}{c^2}$ and $(b+c)\cos A \geq (c+a)\cos B \geq (a+b)\cos C$. Hence, by Chebyshev's Inequality,

$$\begin{aligned}
&\frac{b+c}{a^2}\cos A + \frac{c+a}{b^2}\cos B + \frac{a+b}{c^2}\cos C \\
&\quad \geq \frac{(b+c)\cos A + (c+a)\cos B + (a+b)\cos C}{3}\left(\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2}\right) \\
&\quad = \frac{a+b+c}{3}\left(\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2}\right),
\end{aligned}$$

where we have used the facts that $b\cos A + a\cos B = c$, $c\cos B + b\cos C = a$, and $a\cos C + c\cos A = b$. Finally, by the AM-GM-HM inequalities,

$$\left(\frac{a+b+c}{3}\right)^2 \geq \left(\sqrt[3]{abc}\right)^2 = \sqrt[3]{a^2b^2c^2} \geq \frac{3}{\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2}}.$$

It follows that

$$\frac{a+b+c}{3}\left(\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2}\right) \geq \frac{9}{a+b+c},$$

completing the proof.

Also solved by Tsehaye Andebrhan (Eritrea), Herb Bailey, Roy Barbara (Lebanon), Jean Bogaert (Belgium), Mario Catalani (Italy), Chip Curtis, Eugene Curtin and John W. Spellmann, Knut Dale (Norway), Daniele Donini (Italy), Ovidiu Furdui, Marty Getz and Dixon Jones, Julien Grivaux (France), Enkel Hysnelaj (Australia), Stephen Kaczkowski, Parvis Khalili, D. Kipp Johnson, Murray S. Klamkin (Canada), Neela Lakshmanan (India), Kee-Wai Lau (China), Vivek Kumar Mehra (India), Phillip P. Ray, Rolf Richberg (Germany), Achilleas Sinefakopoulos, and the proposer.

A Product of Projectors **December 2002**

1661. *Proposed by Götz Trenkler, Dortmund, Germany.*

Let P and Q be two $n \times n$ complex orthogonal projectors, that is, $P = P^* = P^2$ and $Q = Q^* = Q^2$. Prove that PQ is an orthogonal projector if and only if PQ is normal, that is, if and only if $(PQ)^*PQ = PQ(PQ)^*$.

Solution by Tom Jager, Calvin College, Grand Rapids, MI.

It is clear that if PQ is an orthogonal projector, then PQ is normal.

If PQ is normal, then there exist a unitary matrix U and a diagonal matrix D such that $PQ = UDU^*$. Now

$$\begin{aligned}(PQ)^3 &= PQ(PQP)Q = PQ\big(PQ(PQ)^*\big)Q \\ &= PQ\big((PQ)^*PQ\big)Q = PQ(QPQ)Q = (PQ)^2.\end{aligned}$$

Thus 0 and 1 are the only possible eigenvalues for PQ. It follows that $D^2 = D$ and $D^* = D$, so that $(PQ)^2 = PQ$ and $(PQ)^* = PQ$.

Also solved by Morgan Balwanz and Cara Lee and Ann McDonald and Takumi Nishjuchi and Deborah Stewart, Michel Bataille (France), David Bendall, Eric Brawner, Minh Can, Adam Coffman, Luz M. DeAlba, Daniele Donini (Italy), David E. Manes, Nicholas C. Singer, Daniel Treat, Xiaoshen Wang, Li Zhou, and the proposer.

Answers

Solutions to the Quickies from page 400.

A935. Both A and B are zero matrices. This can be seen as follows. Because $A^2 = A$ and $B^2 = B$, we have

$$(A + B)^2 = A + B + AB + BA = 0.$$

Hence, $A + B$ is nilpotent, so the eigenvalues, and consequently the trace, of $A + B$ are zero. Thus

$$0 = \text{tr}(A + B) = \text{tr}(A) + \text{tr}(B) = \text{rank}(A) + \text{rank}(B),$$

where the last equality follows because A and B are idempotent. It follows that $\text{rank}(A) = \text{rank}(B) = 0$, so $A = 0 = B$.

A936. We may restrict our attention to triangles whose longest side has length 2. We may further assume that any triangle has vertices ABC where $A = (-1, 0)$, $B = (1, 0)$, and C is in the upper half plane and in the intersection of the disks of radius 2 centered at A and B. Such a triangle is acute if and only if C lies outside of the circle $\mathcal{C}$ with diameter AB. Also, a triangle is in T_k if and only if C lies outside the ellipse E with foci A, B and vertices $(k, 0)$, $(-k, 0)$. Because the $\mathcal{C}$ and E cannot coincide, we cannot have $T_k = \mathcal{A}$. However, $\mathcal{C}$ and E are tangent when $k^2 - 1 = 1$, that is, when $k = \sqrt{2}$. This is the smallest value of k for which $T_k \subseteq \mathcal{A}$.

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.

Zanganeh, Lila Azam, Not Moliére! Ah, nothing is sacred, *New York Times* (6 September 2003) B9, B11; (http://www.nytimes.com/2003/09/06/arts/theater/06CORN.html . Holmes, David I., and Judit Kardos, Who was the author? An introduction to stylometry, *Chance* 16 (2) (2003) 5–8. Rudman, Joseph, Cherry picking in nontraditional authorship attribution studies, 26–32.

A book recently published in France calls into question the authorship by Molière of the plays *Le Tartuffe*, *Le Misanthrope*, and others, ascribing them instead to Pierre Corneille (author of *Le Cid*). Sacre bleu! The basis of the claim by Dominique Labbé in *Corneille in the Shadow of Molière* is "infallible statistical evidence," using relative frequency of word usage. In a special-topic issue of *Chance*, Holmes and Kardos sketch the history and place of *stylometry*, "the statistical analysis of literary style," including notable success in identifying the author of disputed Federalist papers. Other articles in the issue use mainly principal components analysis to pin down the author of *The Royal Book of Oz* and of letters attributed to Civil War Gen. George Pickett. Finally, a response by Rudman points out a frequent flaw in studies of this nature—"cherry picking," the selection of data favorable to one's thesis and the neglect of unfavorable evidence. Rudman identifies potential pitfalls in the selections of the texts, style markers, statistical tests, control texts, stopping point, and ex post facto analysis, concluding that practitioners of stylometry should be better educated and more collaborative.

Robinson, Sara, Are medical students meeting their (best possible) match? *SIAM News* 36 (3) (April 2003) 1, 8–9. How much can matching theory improve the lot of medical residents?, 36 (6) (July-August 2003) 4–5.

Until recently, medical students were matched to hospital residencies using rankings and the infamous algorithm analyzed by David Gale and L.S. Shapley for finding a stable matching in a bipartite graph ("College admissions and the stability of marriage," *American Mathematical Monthly* 69 (1962) 9–14). Despite contrary claims by the National Resident Matching Program (NRMP), "the Match" gave each medical student the worst possible (in the student's estimation) hospital, and each hospital the best possible students (in its ranking), of any stable matching. Medical residents sued, not on the basis of fraudulent claims or the algorithm itself (which was adjusted in the 1990s), but alleging that NRMP's monopoly restrains competition and lowers salaries for residents. Robinson's second article summarizes studies on matching auctions, which include price competition. A new wrinkle: School systems need a fair way to allocate students to schools under "school choice"—particularly since the No Child Left Behind federal law allows a student at a "failing" school to transfer to another school—when not all students can be assigned to schools of their choice. The school systems probably don't realize that they are dealing with a sophisticated algorithmic problem, and they need expert advice from mathematicians and economists; Boston, for example, uses a natural-sounding algorithm that encourages lying about preferences and hence may produce unstable matches.

BBC/WGBH Boston, Nova: *Infinite Secrets: The Genius of Archimedes*, 2003; 55-min film. Directed and produced by Liz Tucker. $19.95 (at `http://www.pbs.org/wgbh/nova/archimedes`). Kaufman, Sarah, On "Nova," "secrets" only hinted at, *Washington Post* (30 September 2003) C7; `http://www.washingtonpost.com/wp-dyn/articles/A20063-2003Sep29.html`. Netz, Reviel, Ken Saito, and Natalie Tchernetska, A new reading of *Method* Proposition 14: Preliminary evidence from the Archimedes Palimpsest (Part 1), *SCIAMVS: Sources and Commentaries in Exact Sciences* 2 (2001) 1–29; (Part 2), 3 (2002) 109–125. Beck, Ellen, Text shows Archimedes did consider infinity, UPI (11 November 2002); `http://www.upi.com/view.cfm?StoryID=20021111-054914-9241r`.

A new film about the rediscovered Archimedes Palimpsest hypes exaggerated speculation ("a book that could have changed the history of the world") about how early integration of Archimedes' ideas into mainstream mathematics would have accelerated technological progress. Reviewer Kaufman has it right: "[I]n overselling it, the program misses a subtler, more poetic point. The drama lies in viewing the book not as a missed opportunity, but as a long-obscured portrait of the past, a vision of human experience hidden under 2,000 years of dust." A shorter but better film is *The Archimedes Palimpsest* from the Walters Art Gallery (1999; 30 min, $14.95, `www.thewalters.org`). Netz et al. report on new discoveries from careful examination of the Palimpsest, including a "previously unsuspected application" of a lemma "with a proportion involving infinitely many objects"; they claim that Archimedes stated that two particular infinite sets have an equal number of members.

Cipra, Barry, *What's Happening in the Mathematical Sciences*, vol. 5, American Mathematical Society, 2002; vi + 95 pp, $19(P). ISBN 0–8218–2904–1.

This new installment (now biennial instead of annual) in an acclaimed series features attractive presentations and understandable explanations of advances in mathematics, on such topics as modular forms, protein folding, Kepler's sphere-packing conjecture, finitude of the universe, traffic jams, small-world networks, and the three-body problem.

Albert, Jim, *Teaching Statistics Using Baseball*, Mathematical Association of America, 2002; xi + 289 pp, $45(P) ($36 for members). ISBN 0–88385–727–8.

This book provides examples and exercises that apply probability and statistics to baseball, "the most statistical of all sports." The book can be used either as a source of examples and applications or as a "framework" for an introductory course in statistics for sports-interested students. The 40 case studies are grouped by statistical topic, including a final chapter on modeling baseball using Markov chains.

Pook, Les, *Flexagons Inside Out*, Cambridge University Press, 2003; xi + 170 pp, $90, $35(P). ISBN 0–521–81970–9, 0–521–52574–8.

Flexagons are polygons folded from strips of paper, which display different faces when flexed along the folds. They were first discovered in 1939 by Arthur H. Stone and then investigated by a "Flexagon Committee" at Princeton that included Stone, Bryant Tuckerman, Richard Feynman, and John W. Tukey; Feynman diagrams in atomic physics arose from his diagrams for flexagons. Martin Gardner popularized flexagons in his column on recreational mathematics in *Scientific American*. This book summarizes the mathematics behind flexagons, at a level that can be appreciated by recreational mathematicians.

Feeman, Timothy G., *Portraits of the Earth: A Mathematician Looks at Maps*, American Mathematical Society, 2002; xiii + 123 pp, $26(P). ISBN 0–8218–3255–7.

The book arose from development of a course called "Cartographiometry," on the interface between mathematics and cartography; its development was supported by the NSF's Mathematics Across the Curriculum program. One-variable calculus is sufficient for most of the author's purposes, together with trigonometry; scale factors in two directions substitute for multivariable calculus. The dozen chapters average four exercises apiece, four student projects are described, and an accompanying Maple graphics package for drawing maps is available at the Maple Website.

NEWS AND LETTERS

Acknowledgments

In addition to our Associate Editors, the following referees have assisted the MAGAZINE *during the past year. We thank them for their time and care.*

Bailey, Herbert R., *Rose-Hulman Institute of Technology, Terre Haute, IN*

Barksdale, Jr., James B., *Western Kentucky University, Bowling Green, KY*

Barria, Jose, *Santa Clara University, Santa Clara, CA*

Beauregard, Raymond A., *University of Rhode Island, Kingston, RI*

Beeman, William O., *Brown University, Providence, RI*

Beezer, Robert A., *University of Puget Sound, Tacoma, WA*

Beineke, Lowell W., *Indiana-Purdue University, Ft. Wayne, IN*

Borwein, Jonathan M., *Simon Fraser University, Burnaby, Canada*

Bressoud, David M., *Macalester College, St. Paul, MN*

Brunson, Barry W., *Western Kentucky University, Bowling Green, KY*

Bullington, Grady D., *University of Wisconsin, Oshkosh, Oshkosh, WI*

Burton, David M., *University of New Hampshire, Durham, NH*

Callan, C. David, *University of Wisconsin, Madison, WI*

Chakerian, G. Don, *University of California, Davis, Davis, CA*

DeTemple, Duane W., *Washington State University, Pullman, WA*

Dodge, Clayton W., *University of Maine, Orono, ME*

Dunbar, Steven R., *University of Nebraska-Lincoln, Lincoln, NE*

Dunne, Edward G., *American Mathematical Society, Providence, RI*

Edgar, Hugh M., *Vancouver, Canada*

Eisenberg, Bennett, *Lehigh University, Bethelem, PA*

Elderkin, Richard H., *Pomona College, Claremont, CA*

Ensley, Douglas E., *Shippensburg University, Shippensburg, PA*

Feroe, John A., *Vassar College, Poughkeepsie, NY*

Fisher, Evan, *Lafayette College, Easton, PA*

Fisher, J. Chris, *University of Regina, Regina, Canada*

Flint, Donna L., *South Dakota State University, Brookings, SD*

Frantz, Marc, *Indiana University, Bloomington, IN*

Fung, Maria G., *Western Oregon University, Monmouth, OR*

Gordon, Russell A., *Whitman College, Walla Walla, WA*

Grabiner, Judith V., *Pitzer College, Claremont, CA*

Green, Euline I., *Abilene Christian University, Abilene, TX*

Grosshans, Frank D., *West Chester University of Pennsylvania, Baltimore, MD*

Guichard, David R., *Whitman College, Walla Walla, WA*

Guy, Richard K., *University of Calgary, Calgary, Canada*

Hayashi, Elmer K., *Wake Forest University, Winston-Salem, NC*

Hoffoss, Diane, *University of San Diego, San Diego, CA*

Howard, Fredric T., *Wake Forest University, Winston-Salem, NC*

Hull, Thomas, *Merrimack College, North Andover, MA*

Jafari, Farhad, *University of Wyoming, Laramie, WY*

Janke, Steven J., *Colorado College, Colorado Springs, CO*

Jepsen, Charles H., *Grinnell College, Grinnell, IA*

Johnson, Diane L., *Humboldt State University, Arcata, CA*

Johnson, Warren P., *Bates College, Lewiston, ME*

Jones, Michael A., *Montclair State University, Upper Montclair, NJ*

Jordan, James H., *Washington State University, Pullman, WA*

Larson, Dean S., *Gonzaga University, Spokane, WA*

Lautzenheiser, Roger G., *Rose-Hulman Institute of Technology, Terre Haute, IN*

Leise, Tanya L., *Rose-Hulman Institute of Technology, Terre Haute, IN*

Lo Bello, Anthony J., *Allegheny College, Meadville, PA*

Lynch, Mark, *Millsaps College, Jackson, MS*

Mackin, Gail, *Northern Kentucky University, Highland Heights, KY*

Mackiw, George B., *Loyola College in Maryland, Baltimore, MD*

Manvel, Bennet, *Colorado State University, Ft. Collins, CO*

Martin, Greg, *The University of British Columbia, Vancouver, Canada*

Mazur, Mark S., *Duquesne University, Pittsburgh, PA*

McDermot, Richard F., *Allegheny College, Meadville, PA*

McGuire, Linda E., *Muhlenberg College, Allentown, PA*

Meda, Gowri, *Forest Grove, OR*

Merrill, Kathy D., *Colorado College, Colorado Springs, CO*

Michaels, John G., *State University of New York At Brockport, Brockport, NY*

Monson, Barry, *University of New Brunswick, Canada*

Morics, Steven W., *University of Redlands, Redlands, CA*

Neidinger, Richard D., *Davidson College, Davidson, NC*

Nelsen, Roger B., *Lewis and Clark College, Portland, OR*

Neudauer, Nancy Ann, *Pacific University, Forest Grove, OR*

Oakley, Patricia A., *Goshen College, Goshen, IN*

Ostrov, Daniel N., *Santa Clara University, Santa Clara, CA*

Pedersen, Jean J., *Santa Clara University, Santa Clara, CA*

Pfaff, Thomas J., *Ithaca College, Ithaca, NY*

Pfiefer, Richard E., *San Jose State University, San Jose, CA*

Pinsky, Mark A., *Northwestern University, Evanston, IL*

Riley, Jr., John H., *Bloomsburg University of Pennsylvania, Bloomsburg, PA*

Rodger, Chris, *Auburn University, Montgomery, AL*

Ross, Kenneth A., *University of Oregon, Eugene, OR*

Ross, Peter, *Santa Clara University, Santa Clara, CA*

Roth, Richard L., *University of Colorado, Boulder, CO*

Rouse, Jeremy, *Harvey Mudd College, Claremont, CA*

Salwach, Chester J., *Lafayette College, Easton, PA*

Samuels, Stephen M., *Purdue University, West Lafayette, IN*

Saxe, Karen, *Macalester College, St. Paul, MN*

Schaefer, Edward F., *Santa Clara University, Santa Clara, CA*

Schattschneider, Doris J., *Moravian College, Bethlehem, PA*

Schumer, Peter D., *Middlebury College, Middlebury, VT*

Scott, Richard A., *Santa Clara University, Santa Clara, CA*

Shahriari, Shahriar, *Pomona College, Claremont, CA*

Sherman, Gary, *Rose-Hulman Institute of Technology, Terre Haute, IN*

Shifrin, Theodore, *University of Georgia, Athens, GA*

Simoson, Andrew J., *King College, Bristol, TN*

Sloughter, Daniel C., *Furman University, Greenville, SC*

Somer, Lawrence E., *Catholic University of America, Washington, DC*

Stone, Alexander P., *University of New Mexico, Albuquerque, NM*

Szabó, Tamás, *Weber State University, Ogden, UT*

Tikoo, Mohan Lal, *Southeast Missouri State University, Cape Girardeau, MO*

Tucker, Thomas W., *Colgate University, Hamilton, NY*

Villalobos, Cristina, *The University of Texas-Pan American, Edinburg, TX*

Wagon, Stan, *Macalester College, St. Paul, MN*

Walden, Byron, *Santa Clara University, Santa Clara, CA*

Whitehead, Tamsen, *Santa Clara University, Santa Clara, CA*

Wilson, Julia M., *SUNY-Fredonia, Fredonia, NY*

Worner, Tamara S., *Wayne State College, Wayne, NE*

Zaman, Naveed, *West Virginia State College, Institute, WVA*

Zorn, Paul, *St. Olaf College, Northfield, MN*

Zullo, Holly S., *Carroll College, Helena, MT*

Index to Volume 76

AUTHORS

TITLES

PROBLEMS

The letters P, Q, and S refer to Proposals, Quickies, and Solutions, respectively; page numbers appear in parentheses. For example, P1665(68) refers to Proposal 1655, which appears on page 68.

I would also like to acknowledge those who have refereed proposals for the Problem Column: Steven Dunbar, Richard Gibbs, Sunder Sethuraman, and Ananda Weerasinghe.

ted States Postal Service

tatement of Ownership, Management, and Circulation

ublication Title: Mathematics Magazine	2. Publication Number: 0 0 2 5 - 5 7 0 x	3 Filing Date: 9/4/03
ssue Frequency: Bimonthly, except July and August	5. Number of Issues Published Annually: 5	6. Annual Subscription Price: $20
Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4)*: MAA, 1529 18th St., NW, Washington, DC 20036		Contact Person: H. Waldman Telephone: 202-387-5200

Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*

Same

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

lisher *(Name and complete mailing address)*

MAA, 1529 18th St., NW, Washington, DC 20036-1385

tor *(Name and complete mailing address)*

Frank Farris, Santa Clara University, Santa Clara, CA 95053

naging Editor *(Name and complete mailing address)*

Harry Waldman

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

ll Name	Complete Mailing Address
The Mathematical ASSN of America	1529 18th St., NW, Washington, DC 20036

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

ll Name	Complete Mailing Address

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

Form 3526, October 1999 *(See Instructions on Reverse)*

13. Publication Title: Mathematics Magazine			14. Issue Date for Circulation Data Below: June, 2003	
15		Extent and Nature of Circulation	Average no. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Da
a. Total Number of Copies *(Net press run)*			14,000	15,000
b Paid and/or Requested Circulation	(1)	Paid/Requested Outside-County Mail Subscriptions Stated on Form 3541. *(Include advertiser's proof and exchange copies)*	12,500	13,500
	(2)	Paid In-County Subscriptions Stated on Form 3541 *(Include advertiser's proof and exchange copies)*	0	0
	(3)	Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Non-USPS Paid Distribution	0	0
	(4)	Other Classes Mailed Through the USPS	0	0
c Total Paid and/or Requested Circulation *[Sum of 15b. (1), (2),(3),and (4)]*		▶	12,500	13,500
d Free Distribution by Mail *(Samples, complimentary, and other free)*	(1)	Outside-County as Stated on Form 3541	0	0
	(2)	In-County as Stated on Form 3541	0	0
	(3)	Other Classes Mailed Through the USPS	1,000	1,000
e. Free Distribution Outside the Mail *(Carriers or other means)*			0	0
f. Total Free Distribution *(Sum of 15d. and 15e.)*		▶	1,000	1,000
g Total Distribution *(Sum of 15c. and 15f)*		▶	13,500	14,500
h. Copies not Distributed			500	500
i. Total *(Sum of 15g. and h.)*		▶	14,000	15,000
j. Percent Paid and/or Requested Circulation *(15c. divided by 15g. times 100)*			90%	90%

16. Publication of Statement of Ownership
☒ Publication required. Will be printed in the December issue of this publication. ☐ Publication not required

17. Signature and Title of Editor, Publisher, Business Manager, or Owner: Donald J. Albers — Date: 9/4/03

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

Instructions to Publishers

1. Complete and file one copy of this form with your postmaster annually on or before October 1. Keep a copy of the completed fo for your records.
2. In cases where the stockholder or security holder is a trustee, include in items 10 and 11 the name of the person or corporation whom the trustee is acting. Also include the names and addresses of individuals who are stockholders who own or hold 1 perce or more of the total amount of bonds, mortgages, or other securities of the publishing corporation. In item 11, if none, check the box. Use blank sheets if more space is required.
3. Be sure to furnish all circulation information called for in item 15. Free circulation must be shown in items 15d, e, and f.
4. Item 15h., Copies not Distributed, must include (1) newsstand copies originally stated on Form 3541, and returned to the publis (2) estimated returns from news agents, and (3), copies for office use, leftovers, spoiled, and all other copies not distributed.
5. If the publication had Periodicals authorization as a general or requester publication, this Statement of Ownership, Managemen and Circulation must be published; it must be printed in any issue in October or, if the publication is not published during Octob the first issue printed after October.
6. In item 16, indicate the date of the issue in which this Statement of Ownership will be published.
7. Item 17 must be signed.

Failure to file or publish a statement of ownership may lead to suspension of Periodicals authorization.

PS Form 3526, October 1999 *(Reverse)*

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW
Washington, DC 20036