Vol. 81, No. 3, June 2008

MATHEMATICS MAGAZINE

Helaman Ferguson's Four Canoes

- Bernoulli Numbers and Polynomials
- Somewhat More than Governors Need to Know about Trigonometry
- π to Thousands of Digits from Vieta's Formula

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 Farris, Editor, Mathematics Magazine, Department of Mathematics, Santa Clara University, 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 credit: The cover image shows Helaman Ferguson's sculpture, entitled Four Canoes, outside of the Science and Engineering Center on the campus of the University of St. Thomas in St. Paul, MN. Why is it four canoes? Our lead article by Melissa Loe and Jenny Shepard explains that title and delves into the sculpture's mathematical and physical properties.

AUTHORS

Melissa Shepard Loe received her doctorate from the University of Minnesota and is an Associate Professor of Mathematics at the University of St. Thomas. Her professional interests include mathematics education, topology, and geometry. She enjoys working with K-12 teachers, has conducted workshops in geometry, and has acted as a consultant for middle school mathematics curricula. Most of her spare time is filled by her four daughters' activities, but when she has time to herself, she enjoys skiing, volleyball, reading, and playing the piano.

Jenny Merrick Borovsky received her BA in Mathematics from Gustavus Adolphus College. She is currently the K-12 Mathematics Department Chair at St. Paul Academy and Summit School and teaches in the middle school. Her interest in *Four Canoes* began during a summer workshop that was taught by Professor Loe. Jenny was inspired to recreate the hexagonal tiling using Geometer's Sketchpad and, in the process, learned much more. Her other interests include coaching athletics, cooking, and quilting.

Tom M. Apostol, Professor Emeritus at Caltech, is best known for his textbooks on calculus, analysis, and analytic number theory, (translated into 5 languages), and for creating *Project MATHEMATICS!*, a series of award-winning video programs that bring mathematics to life with computer animation (translated into Hebrew, Portuguese, French, and Spanish). He has published 90 articles and research papers, 40 of them since he retired in 1992. In 2000 he was elected a Corresponding Member of the Academy of Athens. He contributed two chapters to the *Digital Library of Mathematical Functions:* NT: *Functions of Number Theory*, and ZE: *Zeta Functions*.

Skip Garibaldi learned trigonometry before dropping out of high school and Galois theory while working on his PhD at U C San Diego. These days he works on Galois cohomology as an associate professor at Emory University. He is proud of the fact that a villain on the TV show Aqua Teen Hunger Force is named after Emory.

Vol. 81, No. 3, June 2008

MATHEMATICS MAGAZINE

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

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

Advertising correspondence should be addressed to

MAA Advertising
1529 Eighteenth St. NW
Washington DC 20036

Phone: (866) 821-1221
Fax: (202) 387-1208
E-mail: advertising@maa.org

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

Change of address, missing issue inquiries, and other subscription correspondence:

MAA Service Center, *maahq@maa.org*

All at the address:

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

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

Postmaster: Send address changes to Membership/ Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036-1385.

Printed in the United States of America

ARTICLES

The Mathematics of Helaman Ferguson's *Four Canoes*

MELISSA SHEPARD LOE
University of St. Thomas
St. Paul, MN 55105
msloe@stthomas.edu

JENNY MERRICK BOROVSKY
St. Paul Academy and Summit School
St. Paul, MN 55105
jborovsky@spa.edu

Helaman Ferguson's massive sculpture, entitled *Four Canoes*, resides outside of the Science and Engineering Center on the campus of the University of St. Thomas in St. Paul, MN. (See [1] for pictures of the artwork and installation process). The sculpture consists of two linked granite "donuts", one red and the other black. Each measures six feet in diameter, and weighs more than three tons. These donuts rest on granite pedestals rising two feet above thirty jagged granite hexagons that tile the ground beneath the sculpture. So, why is it called *Four Canoes*? What does it have to do with mathematics? What is the significance of the tiling? What rules govern placement of the individual tiles? Is it periodic? Why don't the donuts wobble or fall off the pedestals? By combining different mathematical approaches, this paper will attempt to answer these questions.

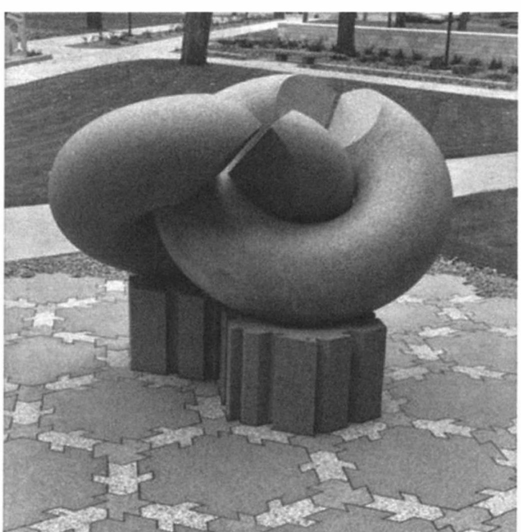


Figure 1 Helaman Ferguson's *Four Canoes*. (Photo courtesy of University Relations, University of St. Thomas.)

Möbius Bands, Klein Bottles, and Canoes

Understanding *Four Canoes* requires some knowledge of Möbius bands and Klein bottles. The reader who is unfamiliar with these can find ample information in recreational mathematics books and websites; for example, see [**2**] or [**3**].

Helaman Ferguson's model of a Möbius band, which differs from the usual construction, but is better suited to understanding *Four Canoes*, is in cross-cap form as in [**4**]: Imagine a strip of paper, with the ends labeled "a" to be identified, as in FIGURE 2. Now curve the strip, as though you were creating a cylinder, bringing the a-edges near each other at the top of the curve. Rather than twisting one edge 180 degrees to identify the a-edges as in the usual construction, imagine making a crease at the center of each a-edge, folding downward in an inverted v-shape. Sew these two inverted v's together across each other, tip-to-tip and tail-to-tail. The Möbius band now has a self-intersection, called a *cross-cap*, along the seam (see FIGURE 3).

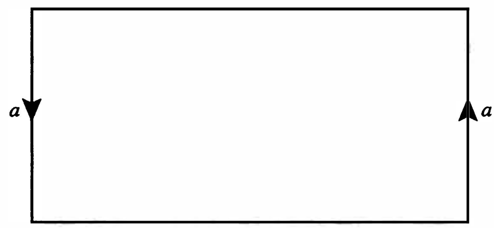


Figure 2 Planar model of a Möbius band.

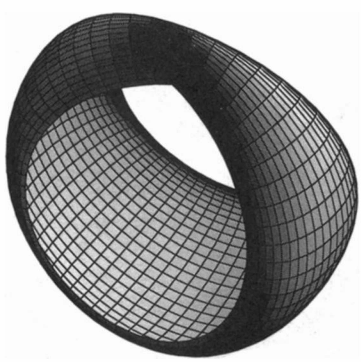


Figure 3 A Möbius band: a cylinder with a cross-cap.

An alternative way to think of the Möbius band in cross-cap form is to start with a canoe, like the one pictured in FIGURE 4. A canoe can be constructed from a (more or less) rectangular sheet of material. The bow and stern are the a-edges, and after folding along the keel or center line, each has been sewn (in non-Möbius fashion) to itself. Imagine stretching and curling the canoe of FIGURE 4 upward and around at the ends until the bow and stern meet. Cut the bow and stern apart along the seams and re-sew the forward port side to the aft starboard side and vice versa, again constructing a surface like the one in FIGURE 3. We have simply traded fore and aft, port and starboard for the tips and tails of the a-arrows.

FIGURE 5 shows the standard planar model of a *Klein bottle*. Sewing together the edges marked "b" (respecting the direction of the arrows), we get a cylinder with two

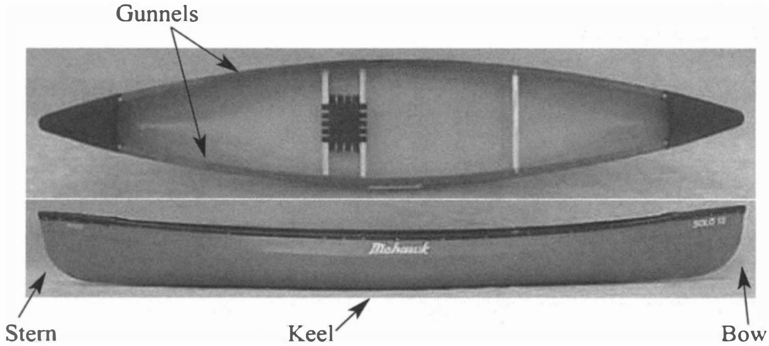


Figure 4 Canoe. (Photo courtesy of the Mohawk Canoe Company.)

oppositely oriented boundaries marked "a". In order to complete the construction by sewing the a-edges together, we have to allow the bottle to intersect and pass through itself. (See [**2**] or [**3**] for more about Klein bottles.)

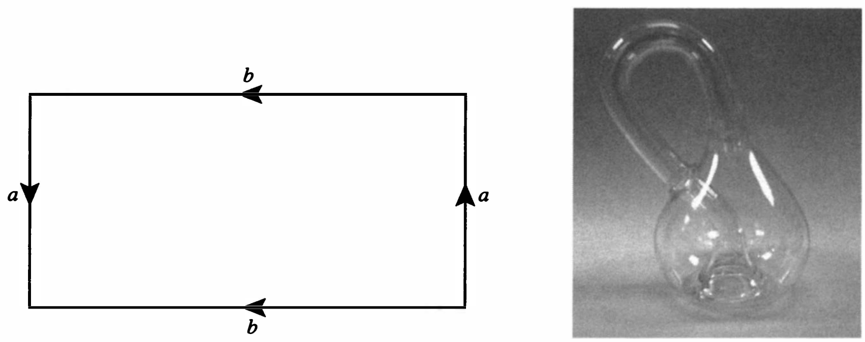


Figure 5 (a) Planar Klein bottle. (b) Glass Klein bottle. (Photo courtesy of Acme Klein Bottles.)

A second way to create a Klein bottle is by sewing two Möbius bands together along their simple closed curve boundaries. You'll get a pinched part when you're near the end of the sewing, where the Möbius bands need to pass through each other in order to complete the construction. If you *could* do this in three dimensions, you'd get a Klein bottle. FIGURE 6 shows the sewing process. To convince yourself that FIGURE 6 represents a Klein bottle, imagine sewing together the edges marked "b_2" to get a cylinder. The boundaries of this cylinder are the two edges to be sewn together, one oriented clockwise, one counterclockwise, just as in FIGURE 5.

The well-known anonymous limerick describes it nicely:

A mathematician named Klein
Thought the Möbius strip was divine
And he said, "If you glue
The edges of two,
You'll get a weird bottle like mine!"

Finally, to understand the name, *Four Canoes*, imagine that two flexible rubber canoes are stacked "gunnels-to-gunnels". Bend and curve them into a donut shape,

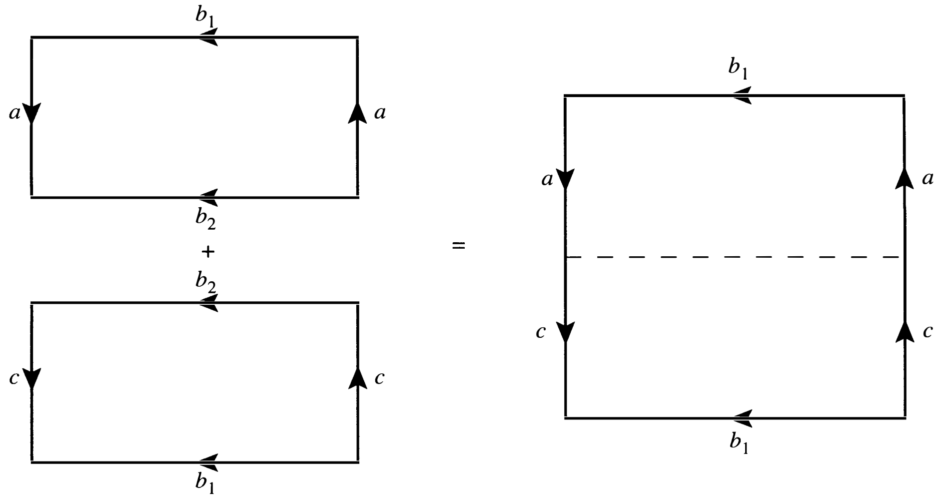


Figure 6 A Klein bottle is two Möbius bands sewn together along their boundaries.

stretching the bottom one while contracting the upper one in a complementary fashion. The keel of one canoe stays outermost, and the keel of the other one is innermost, their gunnels still touching. To transform these canoes into Möbius bands in cross-cap form, we need to unlace the bows and sterns, and cross-sew them as we did earlier in FIGURE 3. Lastly, sew these two canoes to each other along their gunnels to get a Klein bottle in double cross-cap form. This is consistent with our earlier view of the Klein bottle as two Möbius bands sewn together. Finally, the name of the sculpture fits: each "pinched donut" is a Klein bottle, each Klein bottle is two canoes. Hence we have *Four Canoes*. (For Ferguson's description and explanation of *Four Canoes*, see [4].)

The Tiling

Shifting our focus to the tiles beneath the sculpture, as depicted in FIGURE 7, consider a hexagon instead of a rectangle. With edges labeled as in FIGURE 8a, our hexagon also represents a Klein bottle. (As in FIGURES 5a and 6, sew together the b-edges to

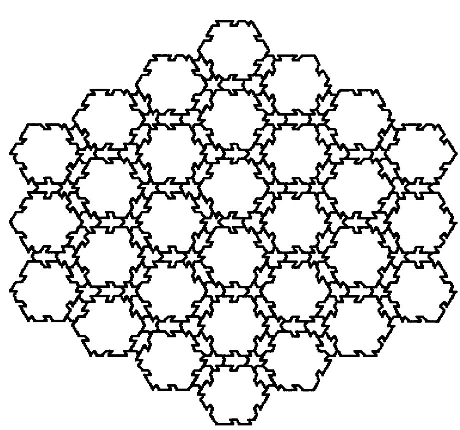


Figure 7 Helaman Ferguson's patch of thirty tiles.

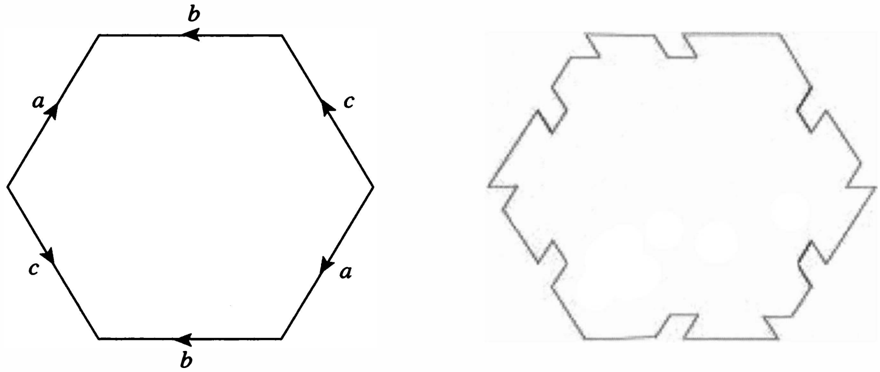


Figure 8 (a) Labeled hexagon Klein bottle. (b) Notched hexagon Klein bottle.

get a cylinder whose boundary edges are oppositely oriented). FIGURE 8b is the outline of a granite tile from the base of Ferguson's sculpture. The rhombuses and triangles cut from the sides of these hexagons replace the arrows and labels from FIGURE 8a. These cuts indicate the identification of edges that transforms each hexagon tile into a Klein bottle. The cuts also create a "right side" and a "wrong side", so that if a tile is flipped over, it doesn't match any right-side-up tile. (Without the triangular cuts, the hexagon has no right or wrong side). Ferguson's tiles are all rotations of one another; all are "face-up".

In addition to the 30 jagged hexagons, the tiling includes 42 corner pieces (one wherever three hexagons meet), and 71 edge-arrows between edges of adjacent tiles. The edge-arrows are all identical, but eight differently shaped corner pieces are needed. The shape of the corner piece required at a given corner depends on the number of edge-arrows pointing inward at that corner and on how many of the three hexagons meeting there have a triangle cut from an edge.

All triangles and rhombuses cut from the edges of the hexagon are necessary for the tile to represent a Klein bottle. However, the rule for placement of the tiles, stated below, depends only on the rhombus cuts, so in our effort to understand the tiling, we'll eliminate the triangle cuts to make the pictures less cluttered.

The single rule governing the placement of tiles in FIGURE 7 is: *two tiles can be adjacent along an edge whenever an edge-arrow is created by the pair of rhombus cuts (see* FIGURE *10a), but not when a zig-zag is created (see* FIGURE *10b).*

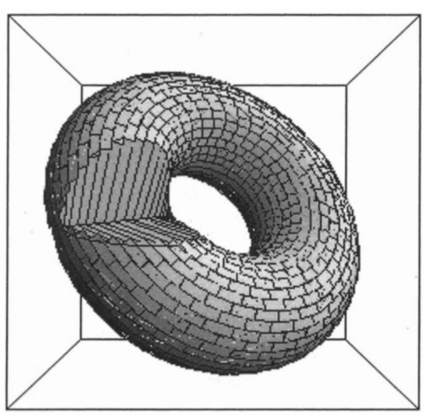


Figure 9 A pinched donut.

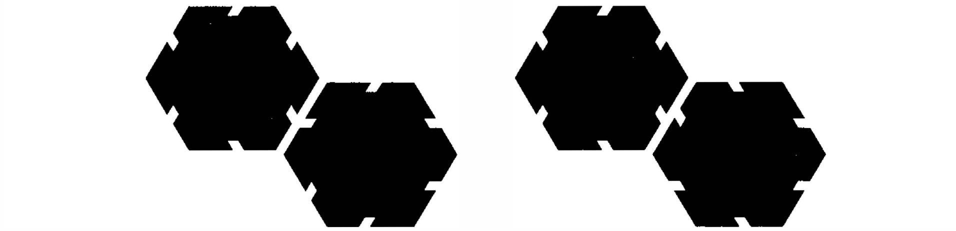


Figure 10 (a) Legal placement of tiles. (b) Illegal placement of tiles.

Ferguson had two objectives in placing the tiles: following the rule he had created, and making the finished patch of tiles fit together in an aesthetically pleasing way [**5**]. He was not concerned with the periodicity of the patch. Consequently, there is no apparent pattern to the rotation of each tile. Notice that in applying Ferguson's rule an arrangement of tiles could occur in which no tile might legally fill a given hole in the tiling (for example, FIGURE 11). In this case, Ferguson simply rotated a tile or two, so that the hole could be filled in [**5**].

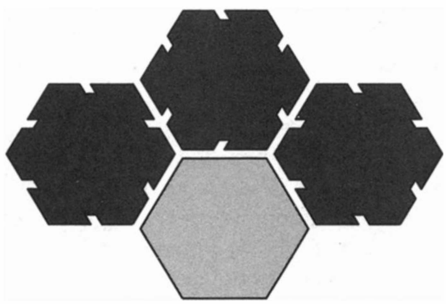


Figure 11 An impossible hole to fill.

The *Four Canoes* tiling appears to be non-periodic, so following the installation, Ferguson wondered whether the tiles admitted a periodic tiling [**5**]. A *periodic tiling* is one in which there exist two non-parallel translations which, when operating on a small patch or *unit cell* of the tiling, will generate the entire pattern. A set of tiles that does not admit a periodic tiling is called *aperiodic*, and such sets are difficult to come by. Penrose tiles are a well-known aperiodic tiling. For more on Penrose tiles, see [**6**].

THEOREM 1. *Helaman Ferguson's Klein bottle tiles can be arranged to tile the plane periodically.*

Proof. Allow only rotations of 0°, 120°, and 240°, as in FIGURE 12. We use a different color for each of the three rotations of the Klein bottle tile. ■

FIGURE 12 exhibits a periodic tiling of the plane whose symmetry group is $p31m$. See [**7**] for more on symmetry groups of tilings of $\mathbf{R}^2$. See [**8**] for constructions of tilings for each of the seventeen planar symmetry groups.

The unit cell of the periodic tiling in FIGURE 12 is surprisingly small: a rhombus or hexagon with the same area as three hexagon tiles! We've included two such unit cells in FIGURE 12, but others are possible of course. Two adjacent sides of the rhombus unit cell provide the translations that generate the pattern.

There are other ways to create the entire pattern from a small patch. If we allow *all* symmetry operations (reflections, rotations and glide reflections as well as transla-

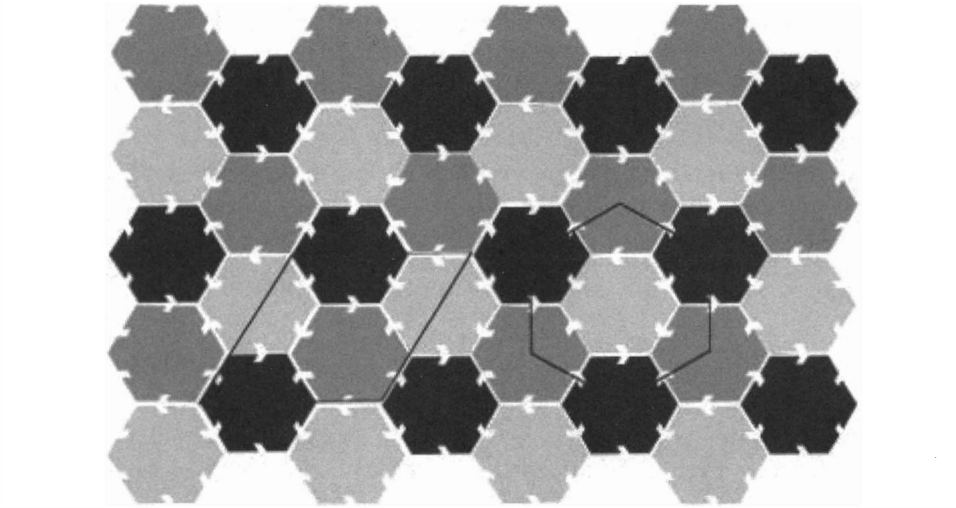


Figure 12 Two unit cells in a periodic tiling by Helaman Ferguson's Klein bottle tiles.

tions), then just half of the rhombus—a triangular patch bounded by the mirror lines in FIGURE 13—will do. Such a patch is called a *fundamental region* for the tiling, and is always no bigger than the unit cell. The fundamental region can have any shape that joins with itself to produce a periodic pattern. M. C. Escher's tessellations provide a rich source of tilings with irregularly shaped fundamental regions [**9**].

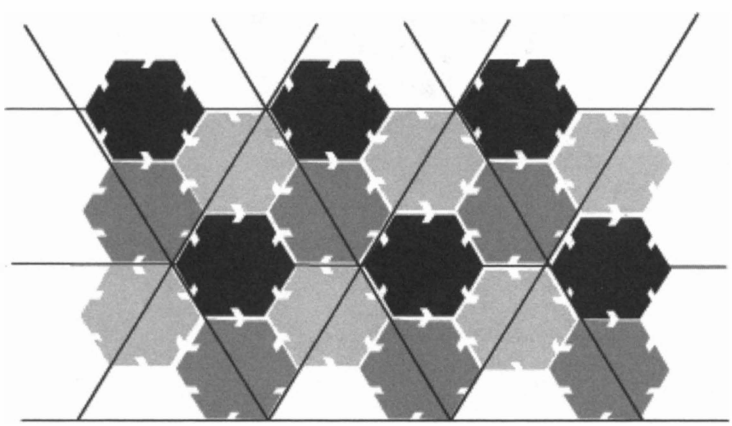


Figure 13 The lines of symmetry in the periodic tiling.

It is still unknown whether Ferguson's patch of thirty tiles (FIGURE 7) can be continued to periodically (or aperiodically) fill the plane (or the sphere) [**5**].

we look at some simple geometry to determine the distance between centers of adjacent hexagons (the dimensions of the hexagon tiles). Second, we use multivariable calculus to guarantee the secure fit of each pinched torus into a "bowl" in the pedestal supporting it.

Dimensions of the tiles. Each donut measures six feet in diameter, is two feet thick, and has a two-foot-diameter-hole in the middle. Each rests on the center of a pedestal, which is just a prism on a Klein bottle hexagon tile (see FIGURE 1).

The distance between adjacent hexagon centers can be determined easily from the following two facts:

1. Each two-foot-thick donut passes through the two-foot-diameter hole in the other donut, so they fit together snugly, at right angles (see FIGURE 14). The linked donuts are tilted 45 degrees from a vertical (or horizontal) plane.
2. In the tilted position, each donut rests on its shell, exactly one foot below the center circle deformation retract of the donut; i.e, the circle that is the very core of the donut (FIGURE 15).

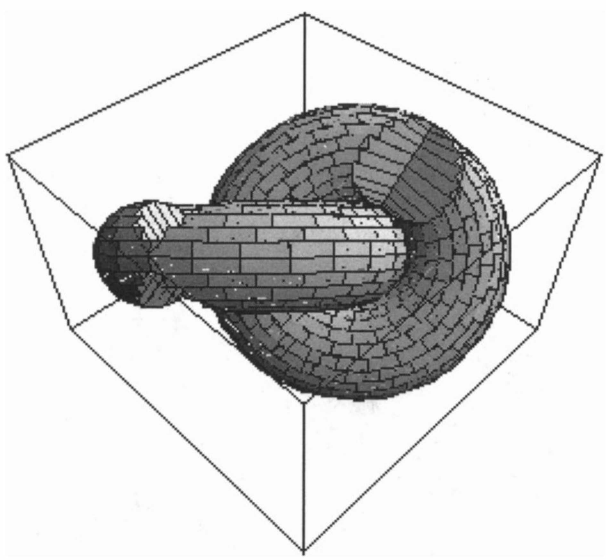


Figure 14 *Mathematica* plot of the linked Klein bottles.

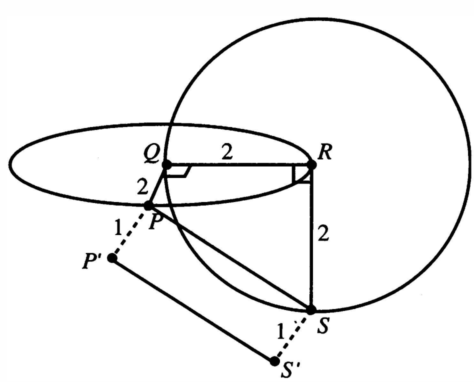


Figure 15 The core of Figure 14 before tilting 45 degrees.

The first fact gives us the linked circles in FIGURE 15 that represent the donut-cores. The points P and S are each on a circle that forms the core of a donut. Point Q is on the core of the vertical donut, and at the center of the hole of the horizontal donut. Similarly, point R is on the core of the horizontal donut, and at the center of the hole of the vertical donut. Observe that the two-foot segments PQ and QR are orthogonal, as are QR and RS. These form three edges of a cube with P and S at opposite corners. Thus the distance from P to S is $2\sqrt{3}$ feet.

Now the point P' (respectively S') is exactly one foot from P (respectively S) on the shell of the donut where it would rest if the tilted sculpture were placed on a flat surface. So the distance from P' to S' is precisely the distance between centers of adjacent hexagons in the patch of tiles. Finally, we use the second fact to conclude that this distance, the distance between the centers of adjacent tiles in the tiling, is also $2\sqrt{3}$ feet.

Before we can determine the exact size of a hexagon tile, we need to specify the width of an edge-arrow, that is, the width of the gap between adjacent tiles. Ferguson chose to make the width of an arrow equal to one-sixth the width of a hexagon tile, so the width of a hexagon (from side to opposite side) is $12\sqrt{3}/7$ feet, and the width of an edge-arrow is $2\sqrt{3}/7$ feet. From the side-to-side width of the hexagon, we easily calculate that the distance between its opposite vertices is $24/7$ feet. Indeed, upon measuring the tiles, we find that these dimensions are accurate!

Stability. Finally, we turn to the pedestals supporting the granite Klein bottles. If the pedestals were flat on top, each Klein bottle would contact its pedestal at a single point, and the sculpture would wobble dangerously. To prevent this, a "bowl" was cut into the top of each pedestal to hold the sculpture securely. What shape does this bowl have? We seek the equation of the curve that sweeps out the bowl.

Consider the point on the torus where the 45 degree tilted torus would rest if it were placed on a horizontal plane. Any vertical plane slicing through the tilted torus at this point will intersect the torus in a closed curve. Take the vertical plane that is also perpendicular to a small circular cross section of the torus, as in FIGURE 15. Because of the symmetry of the torus, this plane intersects the torus in a symmetric, kidney bean shaped curve. If we rotate this curve to cut out a bowl from the top of the pedestal, this bowl will contact the torus all along the kidney bean curve. The bowl's depth can be adjusted to give the desired stability.

To find the equation of the curve, we parameterize the torus and slicing plane, and find the curve of their intersection.

Consider the torus, lying on its side, given in cylindrical coordinates $[r, \theta, z]$ by

$$(r-2)^2 + z^2 = 1.$$

Changing to Cartesian coordinates (x, y, z), we have

$$x^2 + y^2 + z^2 - 4\sqrt{x^2 + y^2} + 4 = 1.$$

We'll use the "slicing plane" given by $x + z = 2$, at a 45 degree angle to the torus (see FIGURE 16). Notice that the intersection of the plane with the torus is indeed a curve shaped like a kidney bean.

Next, we perform two transformations to tilt the plane vertically, and put the origin of the coordinate system at the center of the small (vertical) circular cross section of the torus: First use the translation $x \to x + 2$. Now the torus is given by:

$$(x+2)^2 + y^2 + z^2 - 4\sqrt{(x+2)^2 + y^2} + 4 = 1.$$

and the slicing plane is $x + z = 0$.

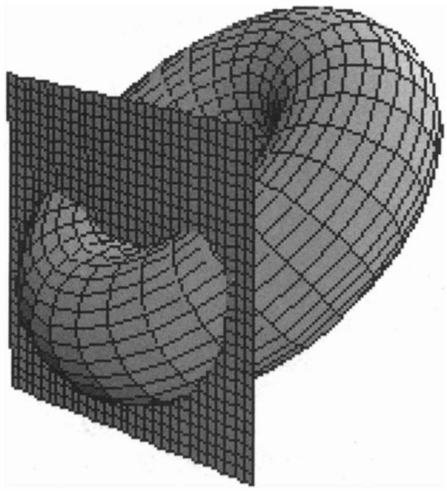


Figure 16 Torus and slicing plane.

Secondly, rotate the coordinate system, so that the slicing plane is given by $x = 0$. Now the torus and plane are both tilted 45 degrees. The rotation matrix is

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad \text{where } \theta = \pi/4.$$

Thus, the torus is given by

$$\left(\frac{\sqrt{2}}{2}(x-z)+2\right)^2 + y^2\left(\frac{\sqrt{2}}{2}(x+z)+2\right)^2 - 4\sqrt{\left(\frac{\sqrt{2}}{2}(x-z)+2\right)^2 + y^2} + 4 = 1,$$

and the plane is given by

$$\frac{\sqrt{2}}{2}(x-z) + \frac{\sqrt{2}}{2}(x+z) = 0; \quad \text{i.e., } x = 0.$$

Using *Mathematica*, we solve these equations simultaneously to obtain the equation for their intersection curve:

$$z^2 + y^2 - 2\sqrt{2}z - 4\sqrt{\frac{z^2}{2} + y^2 - 2\sqrt{2}z + 4} + 7 = 0.$$

This curve is plotted in the yz–plane in FIGURE 17 (again, with *Mathematica*), and looks exactly as we expected it to look from FIGURE 16.

The bowl at the top of each pedestal was achieved by rotating the lower part of this curve about a vertical axis through the pedestal's center to sweep out the required volume. The sculpture rests snugly in the two bowls and is stable enough to withstand the college students (and professors) who occasionally climb on the canoes to pose for photographs.

Acknowledgments. The authors gratefully acknowledge the help and encouragement of Helaman Ferguson in the preparation of this paper, the expert LaTeX assistance provided by the author's husband, Brian Loe, and the helpful suggestions of the referees. The authors modified and used *Mathematica* source code found in Alfred Gray's book *Modern Differential Geometry of Curves and Surfaces* to produce some of the figures.

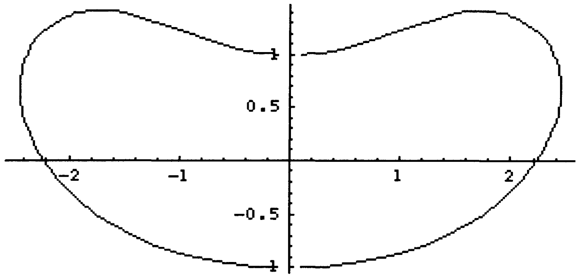


Figure 17 The kidney bean curve.

REFERENCES

1. Helaman Ferguson, *Helasculpt*, `http://www.helasculpt.com/gallery/kleinbottlinksthomas`.
2. Cut the Knot, `http://www.cut-the-knot.org/do_you_know/paper_strip.shtml`.
3. Stephen Barr, *Experiments in Topology*, Dover Publications, New York, 1989.
4. Helaman Ferguson, Duality in mathematical sculpture: linking Klein bottles in granite, *International Journal of Shape Modeling* **5**(1) (1999) 55–68.
5. Helaman Ferguson, personal communications.
6. Martin Gardner, Extraordinary nonperiodic tiling that enriches the theory of tiles, *Scientific American,* January 1977, 110–121.
7. Branko Grünbaum and G. C. Shephard, *Tilings and Patterns*, W. H. Freeman and Company, New York, 1987.
8. Peter Stevens, *Handbook of Regular Patterns: an Introduction to Symmetry in Two Dimensions*, MIT Press, Cambridge, MA, 1980.
9. Doris Schattschneider, *Visions of Symmetry*, W. H. Freeman, New York, 1990.
10. Alfred Gray, *Modern Differential Geometry of Curves and Surfaces*, CRC Press, Boca Raton, FLA, 1993.

A Primer on Bernoulli Numbers and Polynomials

TOM M. APOSTOL
Project MATHEMATICS!
California Institute of Technology
Pasadena, CA 91125
apostol@caltech.edu

A large literature exists on Bernoulli numbers and Bernoulli polynomials, much of it in widely scattered books and journals. This article serves as a brief primer on the subject, bringing together basic results (most of which are well known), together with proofs, in a manner readily accessible to those with a knowledge of elementary calculus. Some new formulas are also derived.

Background

Bernoulli numbers and polynomials are named after the Swiss mathematician Jakob Bernoulli (1654–1705), who introduced them in his book *Ars Conjectandi*, published posthumously (Basel, 1713). They first appeared in a list of formulas (reproduced in FIGURE 1) for summing the pth powers of n consecutive integers, for $p = 1$ to $p = 10$. Bernoulli uses the symbol $\int$, an elongated S, to indicate summation. In modern notation his first three examples are equivalent to the familiar relations

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

$$\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1),$$

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

A general formula for the sum of pth powers (not explicitly stated by Bernoulli), can be written as

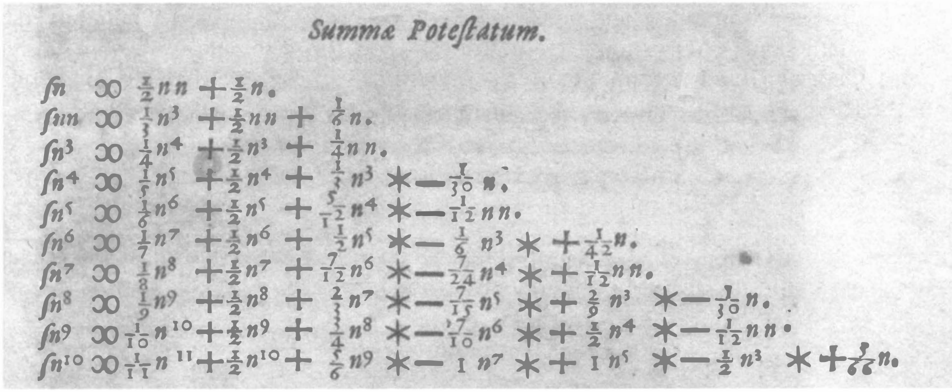
$$\sum_{k=1}^{m-1} k^p = \frac{B_{p+1}(m) - B_{p+1}}{p+1}, \quad p \ge 1,\ m \ge 2, \tag{1}$$

where $B_n(x)$ is a polynomial in x of degree n, now called a Bernoulli polynomial, given by

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} B_k x^{n-k}, \quad n \ge 0, \tag{2}$$

and where the B_k are rational numbers called Bernoulli numbers. They can be defined recursively as follows:

$$B_0 = 1, \quad \text{and} \quad \sum_{k=0}^{n-1} \binom{n}{k} B_k = 0 \quad \text{for } n \ge 2. \tag{3}$$



Summæ Poteſtatum.

$\int n \; \infty \; \frac{1}{2}nn + \frac{1}{2}n.$

$\int nn \; \infty \; \frac{1}{3}n^3 + \frac{1}{2}nn + \frac{1}{6}n.$

$\int n^3 \; \infty \; \frac{1}{4}n^4 + \frac{1}{2}n^3 + \frac{1}{4}nn.$

$\int n^4 \; \infty \; \frac{1}{5}n^5 + \frac{1}{2}n^4 + \frac{1}{3}n^3 \ast - \frac{1}{30}n.$

$\int n^5 \; \infty \; \frac{1}{6}n^6 + \frac{1}{2}n^5 + \frac{5}{12}n^4 \ast - \frac{1}{12}nn.$

$\int n^6 \; \infty \; \frac{1}{7}n^7 + \frac{1}{2}n^6 + \frac{1}{2}n^5 \ast - \frac{1}{6}n^3 \ast + \frac{1}{42}n.$

$\int n^7 \; \infty \; \frac{1}{8}n^8 + \frac{1}{2}n^7 + \frac{7}{12}n^6 \ast - \frac{7}{24}n^4 \ast + \frac{1}{12}nn.$

$\int n^8 \; \infty \; \frac{1}{9}n^9 + \frac{1}{2}n^8 + \frac{2}{3}n^7 \ast - \frac{7}{15}n^5 \ast + \frac{2}{9}n^3 \ast - \frac{1}{30}n.$

$\int n^9 \; \infty \; \frac{1}{10}n^{10} + \frac{1}{2}n^9 + \frac{3}{4}n^8 \ast - \frac{7}{10}n^6 \ast + \frac{1}{2}n^4 \ast - \frac{1}{12}nn.$

$\int n^{10} \; \infty \; \frac{1}{11}n^{11} + \frac{1}{2}n^{10} + \frac{5}{6}n^9 \ast - 1\,n^7 \ast + 1\,n^5 \ast - \frac{1}{2}n^3 \ast + \frac{5}{66}n.$

Figure 1 Reproduction of the list of formulas on p. 97 of *Ars Conjectandi,* in which Bernoulli numbers and Bernoulli polynomials first appeared in print. The open ∞ symbol in the second column was used in that era as an equals sign. The large asterisk in later columns indicates zero coefficients of missing powers. *Courtesy of the Archives, California Institute of Technology.*

Using this recursion with $n = 2, 3, \ldots$, we quickly obtain the following:

$$B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}, \quad B_5 = 0, \quad B_6 = \frac{1}{42},$$
$$B_7 = 0, \quad B_8 = -\frac{1}{30}, \quad B_9 = 0, \quad B_{10} = \frac{5}{66}.$$

Equation (2) shows that $B_n = B_n(0)$ for $n \geq 0$. The sum in (3) can also be written in the form

$$B_n = \sum_{k=0}^{n} \binom{n}{k} B_k, \quad n \geq 2, \tag{4}$$

which, in view of (2), reveals that

$$B_n = B_n(1), \quad n \geq 2. \tag{5}$$

Bernoulli numbers with even subscripts ≥ 2 alternate in sign, and those with odd subscripts ≥ 3 are zero. A knowledge of Bernoulli numbers, in turn, quickly gives explicit formulas for the first few Bernoulli polynomials:

$$B_0(x) = 1, \quad B_1(x) = x - \frac{1}{2}, \quad B_2(x) = x^2 - x + \frac{1}{6},$$
$$B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x, \quad B_4(x) = x^4 - 2x^3 + x^2 - \frac{1}{30},$$
$$B_5(x) = x^5 - \frac{5}{2}x^4 + \frac{5}{3}x^2 - \frac{1}{6}x.$$

Bernoulli proudly announced [**5**, p. 90] that with the help of the last entry in his list of formulas in FIGURE 1 it took him less than half of a quarter of an hour to find that the tenth powers of the first 1000 numbers being added together will yield the sum

91,409,924,241,424,243,424,241,924,242,500.

Today Bernoulli numbers and polynomials play an important role in many diverse areas of mathematics, for example in the Euler-Maclaurin summation formula [**2**], in number theory [**1**], and in combinatorics [**4**]. One of the most remarkable connections is to the Riemann zeta function $\zeta(s)$, defined for $s > 1$ by the infinite series

$$\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}. \tag{6}$$

Leonhard Euler (1707–1783) discovered that when s is an even integer, the sum can be expressed in terms of Bernoulli numbers by the formula

$$\zeta(2n) = (2\pi)^{2n} \frac{|B_{2n}|}{2(2n)!}. \tag{7}$$

In particular,

$$\zeta(2) = \frac{\pi^2}{6}, \quad \zeta(4) = \frac{\pi^4}{90}, \quad \zeta(6) = \frac{\pi^6}{945},$$
$$\zeta(8) = \frac{\pi^8}{9450}, \quad \zeta(10) = \frac{\pi^{10}}{93555}. \tag{8}$$

The Basel problem

The problem of evaluating $\zeta(2)$ in closed form has an interesting history, which apparently began in 1644 when Pietro Mengoli (1625–1686) asked for the sum of the reciprocals of the squares. The problem became widely known in 1689 when Jakob Bernoulli wrote, "If somebody should succeed in finding what till now withstood our efforts and communicate it to us we shall be much obliged to him." By the 1730s it was known as the Basel problem and had defied the best efforts of many leading mathematicians of that era. Euler solved the problem around 1735 in response to a challenge by Jakob's younger brother Johann Bernoulli (1667–1748), who was Euler's teacher and mentor. Euler soon obtained the more general formula in (7). Sadly, Jakob did not live to see young Euler's triumphant discovery and its surprising connection with Bernoulli numbers. For a proof of (7) see [**1**, p. 266]. To date, no simple closed form analogous to (7) is known for $\zeta(n)$ for any odd power $n \geq 3$.

Generating functions

There are alternative methods for introducing Bernoulli numbers and polynomials. One of the most useful was conceived by Euler, who observed that they occur as coefficients in the following power series expansions:

$$\frac{z}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n, \quad |z| < 2\pi, \tag{9}$$

and

$$\frac{ze^{xz}}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!} z^n, \quad |z| < 2\pi. \tag{10}$$

The parameters x and z can be real or complex. The functions on the left are called generating functions for the Bernoulli numbers and polynomials.

Basic properties deduced from the generating functions

The use of generating functions leads to simple and direct proofs of many basic properties of Bernoulli numbers and polynomials, the most important of which are derived here. For example, to deduce (2) from (9) and (10), write

$$\sum_{n=0}^{\infty} \frac{B_n(x)}{n!} z^n = \frac{z}{e^z - 1} \cdot e^{xz} = \left(\sum_{n=0}^{\infty} \frac{B_n}{n!} z^n\right) \cdot \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} z^n\right). \tag{11}$$

Multiply the two power series on the right, using the fact that the product of two convergent power series

$$A(z) = \sum_{n=0}^{\infty} a(n) z^n \quad \text{and} \quad B(z) = \sum_{n=0}^{\infty} b(n) z^n,$$

is another power series given by

$$A(z)B(z) = \sum_{n=0}^{\infty} c(n) z^n, \quad \text{where} \quad c(n) = \sum_{k=0}^{n} a(k) b(n-k).$$

Equating coefficients of z^n in (11) we obtain

$$\frac{B_n(x)}{n!} = \sum_{k=0}^{n} \frac{B_k}{k!} \frac{x^{n-k}}{(n-k)!},$$

which is equivalent to (2). Incidentally, (9) and (10) also show that

$$B_n = B_n(0), \quad n \geq 0. \tag{12}$$

To deduce that

$$B_{2n+1} = 0 \quad \text{for} \quad n \geq 1, \tag{13}$$

rewrite (9) in the form

$$\frac{z}{e^z - 1} + \frac{z}{2} = 1 + \sum_{n=2}^{\infty} \frac{B_n}{n!} z^n.$$

Now observe that the left member is an even function of z (it is unchanged when z is replaced by $-z$), hence the right member is even and therefore contains no odd powers of z, and we get (13).

The power-sum formula (1) and its extension (16), are immediate consequences of the following:

Difference equation:

$$B_n(x+1) - B_n(x) = nx^{n-1}, \quad n \geq 1. \tag{14}$$

To prove (14), use (10) in the identity

$$z \frac{e^{(x+1)z}}{e^z - 1} - z \frac{e^{xz}}{e^z - 1} = ze^{xz}$$

to obtain

$$\sum_{n=0}^{\infty} \frac{B_n(x+1) - B_n(x)}{n!} z^n = \sum_{n=0}^{\infty} \frac{x^n}{n!} z^{n+1}.$$

Equating coefficients of z^n gives (14). Taking $x = 0$ in (14) we find a companion to (12):

$$B_n(0) = B_n(1), \quad n \geq 2. \tag{15}$$

Replace x by $x + k$, and n by $p + 1$ in (14), then sum on k to obtain the following extension of (1):

$$\sum_{k=0}^{m-1} (x+k)^p = \frac{B_{p+1}(m+x) - B_{p+1}(x)}{p+1}, \quad p \geq 1, \ m \geq 1. \tag{16}$$

When $x = 0$, (16) reduces to (1), and when $x = a/d$, $d \neq 0$, it implies

$$\sum_{k=0}^{m-1} (a+dk)^p = d^p \frac{B_{p+1}(m+a/d) - B_{p+1}(a/d)}{p+1}, \quad p \geq 1, \ m \geq 1. \tag{17}$$

In particular, if a and d are integers, (17) provides a formula for the sum of the pth powers of m integers in arithmetic progression.

Symmetry relation:

$$B_n(1-x) = (-1)^n B_n(x), \quad n \geq 0. \tag{18}$$

This follows at once from the identity

$$z\frac{e^{(1-x)z}}{e^z - 1} = -z\frac{e^{-zx}}{e^{-z} - 1}.$$

Take $x = 0$ in (18) to find $B_{2n+1}(1) = -B_{2n+1}(0)$, so by (15) and (10) this gives another proof of (13). When $x = 1/2$ in (18) we get

$$B_{2n+1}\left(\frac{1}{2}\right) = 0, \quad n \geq 0. \tag{19}$$

Now replace x by $-x$ in (18) and use (14) to obtain

$$(-1)^n B_n(-x) = B_n(x) + nx^{n-1}, \quad n \geq 1. \tag{20}$$

Addition formula:

$$B_n(y+x) = \sum_{k=0}^{n} \binom{n}{k} B_k(y) x^{n-k}, \quad n \geq 0. \tag{21}$$

This is an immediate consequence of the identity

$$\frac{ze^{(y+x)z}}{e^z - 1} = \frac{ze^{yz}}{e^z - 1} \cdot e^{xz}.$$

Taking $y = 0$ in (21) gives (2).

Raabe's multiplication formula:

$$B_n(mx) = m^{n-1}\sum_{k=0}^{m-1} B_n\left(x + \frac{k}{m}\right), \quad n \geq 0,\ m \geq 1. \tag{22}$$

This states that a sum of Bernoulli polynomials of degree n at equally spaced values of the argument is another Bernoulli polynomial of degree n. To prove (22), equate coefficients of z^n in the identity

$$\sum_{n=0}^{\infty} \frac{m^{1-n}B_n(mx)}{n!} z^n = \sum_{n=0}^{\infty} \frac{\sum_{k=0}^{m-1} B_n(x+\frac{k}{m})}{n!} z^n, \tag{23}$$

which can be proved as follows. From (10) we have

$$\sum_{n=0}^{\infty} B_n\left(x+\frac{k}{m}\right)\frac{z^n}{n!} = \frac{ze^{(x+\frac{k}{m})z}}{e^z-1} = \frac{ze^{xz}}{e^z-1}e^{kz/m}.$$

Now sum both members on k for $k = 0, 1, \ldots, m-1$. The sum of exponentials is a geometric sum given by

$$\sum_{k=0}^{m-1} e^{kz/m} = \frac{e^z-1}{e^{z/m}-1},$$

if $z \neq 0$. (The restriction $z \neq 0$ is not serious because (23) holds trivially when $z = 0$.) Therefore

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{\sum_{k=0}^{m-1} B_n(x+\frac{k}{m})}{n!} z^n &= \frac{ze^{xz}}{e^z-1}\sum_{k=0}^{m-1} e^{kz/m} = \frac{ze^{xz}}{e^{z/m}-1}\\ &= m\frac{(z/m)e^{(mx)(z/m)}}{e^{z/m}-1} = m\sum_{n=0}^{\infty}\frac{B_n(mx)}{n!}\left(\frac{z}{m}\right)^n\\ &= \sum_{n=0}^{\infty}\frac{m^{1-n}B_n(mx)}{n!}z^n,\end{aligned}$$

which implies (23) and hence (22).

Another application of Bernoulli's power-sum formula

Equation (17) extends (1) to the sum of the pth powers of m integers in arithmetic progression. It is natural to ask if there is a formula similar to (1) when the sum on the left is extended over those integers relatively prime to m. We will show that

$$\sum_{k=1}^{m}{}' k^p = \sum_{d|m} \mu(d)d^p\frac{B_{p+1}(m/d) - B_{p+1}}{p+1}, \quad p \geq 1,\ m \geq 1, \tag{24}$$

where $\sum'$ indicates that the sum is extended over k relatively prime to m. In (24), $\mu(d)$ is the well known Möbius function of elementary number theory, which enters naturally because of the following formula [**1**, p. 25] that selects numbers relatively prime to m:

$$\sum_{d|(k,m)} \mu(d) = \begin{cases} 1 & \text{if } (k,m)=1 \\ 0 & \text{otherwise.} \end{cases}$$

Thus, we have

$$\sum_{k=1}^{m}{}' k^p = \sum_{k=1}^{m-1} \sum_{d|m \text{ and } d|k} \mu(d) k^p.$$

Now write $k = rd$, and the foregoing equation becomes

$$\sum_{k=1}^{m}{}' k^p = \sum_{d|m} \mu(d) d^p \sum_{1 \le r < m/d} r^p = \sum_{d|m} \mu(d) d^p \frac{B_{p+1}(m/d) - B_{p+1}}{p+1},$$

where in the last step we used (1) with m replaced by m/d. This proves (24).

By using (2) to expand the Bernoulli polynomial in (24) in powers of m/d, we can also write

$$\sum_{k=1}^{m}{}' k^p = \frac{1}{p+1} \sum_{r=1}^{p+1} \binom{p+1}{r} m^r B_{p+1-r} \sum_{d|m} \mu(d) d^{p-r}. \tag{25}$$

The dependence on the Möbius function can be removed by invoking Theorem 2.8 of [**1**], which implies

$$\sum_{d|m} \mu(d) d^{p-r} = \prod_{q|m} (1 - q^{p-r}),$$

where the product is taken over all prime divisors q of m. Therefore we have an alternative form of (25):

$$\sum_{k=1}^{m}{}' k^p = \frac{1}{p+1} \sum_{r=1}^{p+1} \binom{p+1}{r} m^r B_{p+1-r} \prod_{q|m} (1 - q^{p-r}). \tag{26}$$

When $m = 1000$ and $p = 10$ the product contains only the primes 2 and 5, and (26) gives

36,366,968,829,066,536,008,898,579,270,000

for the sum of the tenth powers of those integers up to 1000 that are relatively prime to 1000. This sum is about 39.8% of Bernoulli's value mentioned earlier for the sum extended over all integers up to 1000, which is not too surprising because exactly 40% of the numbers less than 1000 are relatively prime to 1000. Like Bernoulli, the author did this calculation by hand, but unlike Bernoulli it took him more than a quarter of an hour.

Properties involving calculus

Differentiate each member of (10) with respect to x and equate coefficients of z^n to get the following:

Derivative formula:

$$B_n'(x) = n B_{n-1}(x), \quad n \ge 1, \tag{27}$$

which is also a consequence of (2). This leads to another proof of addition formula (21). Repeated differentiation of (27) gives

$$B_n''(x) = n(n-1)B_{n-2}(x), \dots, B_n^{(k)}(x) = k!\binom{n}{k}B_{n-k}(x). \tag{28}$$

On the other hand, for each fixed y the Taylor expansion of the polynomial $B_n(y+x)$ in powers of x is given by

$$B_n(y+x) = \sum_{k=0}^{n} \frac{1}{k!}B_n^{(k)}(y)x^k = \sum_{k=0}^{n}\binom{n}{k}B_{n-k}(y)x^k = \sum_{k=0}^{n}\binom{n}{k}B_k(y)x^{n-k},$$

which is (21).

Now replace n by $n+1$ in (27) and integrate to obtain the following:

Integration formula:

$$\int_x^y B_n(t)\,dt = \frac{B_{n+1}(y) - B_{n+1}(x)}{n+1}, \quad n \geq 0. \tag{29}$$

This, together with (14), implies

$$\int_x^{x+1} B_n(t)\,dt = x^n, \quad n \geq 0, \tag{30}$$

which, when $x = 0$, gives

$$\int_0^1 B_n(t)\,dt = 0, \quad n \geq 1. \tag{31}$$

From (29) we find the recursion relation

$$B_n(x) = n\int_0^x B_{n-1}(t)\,dt + B_n(0), \quad n \geq 1. \tag{32}$$

Recursion formulas for Bernoulli polynomials

Equation (32) suggests another method for defining the Bernoulli polynomials and Bernoulli numbers recursively. Define $b_0 = 1$, and $b_0(x) = 1$. Guided by (32) and (31), define

$$b_n(x) = n\int_0^x b_{n-1}(t)\,dt + b_n, \tag{33}$$

where the constant b_n is chosen so that

$$\int_0^1 b_n(t)\,dt = 0. \tag{34}$$

It is easily verified that the functions $b_n(x)$ and constants b_n obtained in this manner are exactly the same as the Bernoulli polynomials $B_n(x)$ and the Bernoulli numbers B_n. For example, using (33) with $n = 1$ we find

$$b_1(x) = \int_0^x b_0(t)\,dt + b_1 = x + b_1,$$

whereas (34) requires

$$\int_0^1 b_1(t)\,dt = \frac{1}{2} + b_1 = 0.$$

This gives $b_1 = -\frac{1}{2} = B_1$, and $b_1(x) = x - \frac{1}{2} = B_1(x)$. Similarly, using (33) with $n = 2$ we find

$$b_2(x) = 2\int_0^x b_1(t)\,dt + b_2 = x^2 - x + b_2,$$

while (34) requires

$$\int_0^1 b_2(t)\,dt = \frac{1}{3} - \frac{1}{2} + b_2 = 0,$$

so $b_2 = \frac{1}{6} = B_2$, and $b_2(x) = x^2 - x + \frac{1}{6} = B_2(x)$.

By induction we find $b_n = B_n$ and $b_n(x) = B_n(x)$ for all $n \geq 0$. In other words, the recursion formulas

$$B_n(x) = n\int_0^x B_{n-1}(t)\,dt + B_n, \qquad \int_0^1 B_n(t)\,dt = 0, \qquad n \geq 1, \tag{35}$$

together with $B_0 = 1$, $B_0(x) = 1$, provide an alternative method for defining Bernoulli numbers and polynomials recursively.

The following further recursion formula for Bernoulli polynomials

$$nB_n(x) - xnB_{n-1}(x) = \sum_{k=1}^{n} \binom{n}{k} kB_k x^{n-k}, \qquad n \geq 1, \tag{36}$$

is a simple consequence of the familiar Pascal triangle property of binomial coefficients,

$$\binom{n}{k} - \binom{n-1}{k} = \binom{n-1}{k-1}.$$

In fact, from (2) we have

$$\begin{aligned} B_n(x) - xB_{n-1}(x) &= \sum_{k=0}^{n} \left[\binom{n}{k} - \binom{n-1}{k}\right] B_k x^{n-k} \\ &= \sum_{k=1}^{n} \binom{n-1}{k-1} B_k x^{n-k} = \sum_{k=1}^{n} \binom{n}{k} \frac{k}{n} x^{n-k}, \end{aligned}$$

which gives (36) after multiplication by n.

Further recursion formulas for Bernoulli numbers

The defining recursion in (3) and its variation in (4) can be written in other equivalent forms. First we show that (4) is equivalent to

$$\sum_{k=1}^{n} \binom{n}{k-1} B_k = -B_n, \tag{37}$$

which is valid for $n \geq 2$ (but not for $n = 1$). Write

$$\binom{n}{k-1} = \binom{n+1}{k} - \binom{n}{k},$$

to obtain

$$\sum_{k=1}^{n} \binom{n}{k-1} B_k = \sum_{k=0}^{n} \binom{n+1}{k} B_k - \sum_{k=0}^{n} \binom{n}{k} B_k,$$

because the terms with $k = 0$ cancel. The first sum on the right is 0 by (3) and we see that (4) is equivalent to (37).

Next we show that (37), in turn, is equivalent to the following recursion which we believe is new:

$$\sum_{k=2}^{n} \binom{n}{k-2} \frac{B_k}{k} = \frac{1}{(n+1)(n+2)} - B_{n+1}, \quad n \geq 2. \tag{38}$$

To derive this, use the relation

$$\binom{n}{k-2} \frac{1}{k} = \frac{k-1}{(n+1)(n+2)} \binom{n+2}{k},$$

multiply both members of (38) by $(n+1)(n+2)$, then add the summand terms for $k = 0$, 1, and $n+1$ to both sides to get the equivalent formula

$$\sum_{k=0}^{n+1} (k-1) \binom{n+2}{k} B_k = -(n+2) B_{n+1}.$$

Because of (3), this equation, in turn, is equivalent to

$$\sum_{k=1}^{n+1} k \binom{n+2}{k} B_k = -(n+2) B_{n+1}. \tag{39}$$

Now write $k\binom{n+2}{k} = (n+2)\binom{n+1}{k-1}$, cancel the common factor $(n+2)$, and replace n by $n-1$, and (39) becomes (37). Hence (38) is equivalent to (4).

A special case of (38), obtained by Horata [**3**] states that

$$\sum_{k=0}^{n-1} \binom{2n}{2k} \frac{B_{2k+2}}{2k+2} = \frac{1}{(2n+1)(2n+2)}, \quad n \geq 1. \tag{40}$$

To get this from (38), replace n by $2n$ in (38) and use the fact that $B_k = 0$ for odd $k \geq 3$.

Horata proved (40) by a complicated method that used formulas expressing Bernoulli numbers in terms of Stirling numbers of the second kind, and showed that both members of (40) are congruent modulo p for all primes p. Our direct proof of the more general result (38) does not depend on Stirling numbers or congruences.

An alternative form of (37) can be obtained by taking $x = 1$ in (36) and using (5):

$$n^2 B_{n-1} = -\sum_{k=1}^{n-2} \binom{n}{k} k B_k, \quad n \geq 3. \tag{41}$$

To see that this is equivalent to (37), divide by n in (41), then subtract $(n-1)B_{n-1}$ from both sides to get (37) with $n-1$ in place of n.

Yet another recursion can be obtained by integrating the product $xB_n(x)$ in two ways. On the one hand, from (2) we have

$$\int_0^x tB_n(t)\,dt = \sum_{k=0}^{n}\binom{n}{k}B_k\int_0^x t^{n+1-k}dx = \sum_{k=0}^{n}\binom{n}{k}\frac{B_k}{n+2-k}x^{n+2-k}.$$

Now calculate the same integral using integration by parts and (27) to obtain

$$\int_0^x tB_n(t)\,dt = \frac{1}{n+1}\int_0^x tB'_{n+1}(t)\,dt = \frac{1}{n+1}\left\{xB_{n+1}(x) - \int_0^x B_{n+1}(t)\,dt\right\}$$
$$= \frac{1}{n+1}\left\{xB_{n+1}(x) - \frac{B_{n+2}(x) - B_{n+2}(0)}{n+2}\right\}.$$

Equating the two results we find the polynomial identity

$$\sum_{k=0}^{n}\binom{n}{k}\frac{B_k}{n+2-k}x^{n+2-k} = \frac{1}{n+1}\left\{xB_{n+1}(x) - \frac{B_{n+2}(x) - B_{n+2}(0)}{n+2}\right\}.$$

When $x = 1$ this simplifies to

$$\sum_{k=0}^{n}\binom{n}{k}\frac{B_k}{n+2-k} = \frac{B_{n+1}}{n+1}, \quad n \ge 1. \tag{42}$$

We leave it as a challenge to the reader to find another proof of (42) as a direct consequence of (3) without the use of integration.

Concluding remarks

In the three centuries after Jakob Bernoulli introduced his power-sum formulas, the polynomials and numbers that bear his name have been generalized in many different directions and have spawned hundreds of papers. You can get an idea of the many important areas of mathematics that have been influenced by these elementary topics by searching for *Bernoulli numbers* or *Bernoulli polynomials* on the world wide web.

The author is grateful to an anonymous referee who pointed out that that an English translation of Bernoulli's *Ars Conjectandi* was published in 2005 by The Johns Hopkins University Press under the title *The Art of Conjecturing, together with Letter to a Friend on Sets in Court Tennis*, by Dudley Sylla, and that collateral material related to the early history of (1) can be found in a paper by D. E. Knuth, *Johann Faulhaber and Sums of Powers*, in *Mathematics of Computation* **61** (1993), 277–294, and in another Johns Hopkins publication, *Pascals Arithmetic Triangle*, by A. W. F. Edwards (2002), especially Chapter 10.

Summary of basic formulas (as numbered in the text)

Defining relations:

$$B_n(x) = \sum_{k=0}^{n}\binom{n}{k}B_k x^{n-k}, \quad n \ge 0 \tag{2}$$

$$B_0 = 1, \quad B_n = \sum_{k=0}^{n}\binom{n}{k}B_k, \quad n \ge 2 \tag{4}$$

Special values:

$$B_n = B_n(1), \quad n \geq 2, \qquad B_n = B_n(0), \quad n \geq 0 \tag{12}$$

$$B_{2n+1} = 0, \quad n \geq 1 \tag{13}$$

Generating functions:

$$\frac{z}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n, \quad |z| < 2\pi \tag{9}$$

$$\frac{ze^{xz}}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!} z^n, \quad |z| < 2\pi \tag{10}$$

Difference equation:

$$B_n(x+1) - B_n(x) = nx^{n-1}, \quad n \geq 1 \tag{14}$$

Symmetry relation:

$$B_n(1-x) = (-1)^n B_n(x), \quad n \geq 0 \tag{18}$$

Addition formula:

$$B_n(y+x) = \sum_{k=0}^{n} \binom{n}{k} B_k(y) x^{n-k}, \quad n \geq 0 \tag{21}$$

Raabe's multiplication formula:

$$B_n(mx) = m^{n-1} \sum_{k=0}^{m-1} B_n\left(x + \frac{k}{m}\right), \quad n \geq 0,\ m \geq 1 \tag{22}$$

Derivative formulas:

$$B_n'(x) = nB_{n-1}(x), \quad n \geq 2 \tag{27}$$

$$B_n''(x) = n(n-1)B_{n-2}(x), \ldots, B_n^{(k)}(x) = k!\binom{n}{k} B_{n-k}(x) \tag{28}$$

Integration formulas:

$$\int_x^y B_n(t)\,dt = \frac{B_{n+1}(y) - B_{n+1}(x)}{n+1}, \quad n \geq 1 \tag{29}$$

$$\int_x^{x+1} B_n(t)\,dt = x^n, \quad n \geq 1 \tag{30}$$

$$\int_0^1 B_n(t)\,dt = 0, \quad n \geq 1 \tag{31}$$

Alternative recursion formulas:

$$B_n(x) = n\int_0^x B_{n-1}(t)\,dt + B_n, \qquad \int_0^1 B_n(t)\,dt = 0, \quad n \geq 1 \tag{35}$$

with $B_0 = 1$, $B_0(x) = 1$,

$$\sum_{k=2}^{n}\binom{n}{k-2}\frac{B_k}{k} = \frac{1}{(n+1)(n+2)} - B_{n+1}, \quad n \geq 2 \tag{38}$$

$$n^2 B_{n-1} = -\sum_{k=1}^{n-2}\binom{n}{k}kB_k, \quad n \geq 3 \tag{41}$$

$$\sum_{k=0}^{n}\binom{n}{k}\frac{B_k}{n+2-k} = \frac{B_{n+1}}{n+1}, \quad n \geq 0 \tag{42}$$

Power-sum formulas:

$$\sum_{k=1}^{m-1} k^p = \frac{B_{p+1}(m) - B_{p+1}}{p+1}, \quad p \geq 1,\ m \geq 2 \tag{1}$$

$$\sum_{k=0}^{m-1}(a+dk)^p = d^p\frac{B_{p+1}(m+a/d) - B_{p+1}(a/d)}{p+1}, \quad p \geq 1,\ m \geq 2 \tag{17}$$

$$\sum_{k=1}^{m}{}' k^p = \sum_{d|m}\mu(d)d^p\frac{B_{p+1}(m/d) - B_{p+1}}{p+1}, \quad p \geq 1,\ m \geq 1 \tag{24}$$

$$\sum_{k=1}^{m}{}' k^p = \frac{1}{p+1}\sum_{r=1}^{p+1}\binom{p+1}{r}m^r B_{p+1-r}\prod_{q|m}(1-q^{p-r}) \quad p \geq 1,\ m \geq 1 \tag{26}$$

REFERENCES

1. T. M. Apostol, *Introduction to Analytic Number Theory*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1976.
2. T. M. Apostol, An elementary view of Euler's summation formula, *American Mathematical Monthly* **106** (1999) 409–418.
3. K. Horata, On congruences involving Bernoulli numbers and irregular primes, II, *Report Fac. Sci. and Technol. Meijo Univ.* **31** (1991) 1–8.
4. J. Riordan, *Combinatorial Identities*, Robert E. Krieger, Huntington, NY, 1979. Reprint of the 1968 original with corrections.
5. D. E. Smith, *A Source Book in Mathematics*, McGraw-Hill, New York, 1929.

Erratum

Authors Chungwu Ho and Seth Zimmerman have written to point out an error on page 14 of their paper "On Infinitely Nested Radicals," this MAGAZINE, Vol. 81, February 2008: The gaps mentioned for the set S_2 do not exist. Gaps exist only for sets S_a with $a \geq 3$.

Somewhat More than Governors Need to Know about Trigonometry[1]

SKIP GARIBALDI
Department of Mathematics & Computer Science
Emory University
Atlanta, GA 30322
skip@member.ams.org

In school, I had to memorize the values of sine and cosine at the angles 0, 30, 45, 60, and 90 degrees. This always made me wonder: Why those angles and not others? After all, there is *some* angle whose sine is, say, 3/4. Why not include it in the table? More generally: What "nice" angles have "nice" values for sine and cosine?

The main goal of this paper is to find answers to these questions "organically" by making naive guesses at the answers and seeing if these guesses are correct. By successively refining our naive guesses, we end up with essentially complete answers. The second goal of this paper is to highlight the connections between trigonometry and Galois theory by using results from a standard course on Galois theory to answer these questions. These things are well known to experts; we aim to popularize them.

The most a governor needs to know

The third goal of this article is to help out former Governor Jeb Bush of Florida. At a July 2004 appearance to promote state-wide annual testing of students in public schools, a high school student asked him: "What are the angles in a 3-4-5 triangle?" The governor responded "I don't know. 125, 90.... and whatever remains on 180." [**14**] Aside from not noticing that 125 and 90 add up to more than 180, it's not such a bad answer. In his favor, the governor remembered that the interior angles add up to 180, something I'm not confident *my* governor would remember. Also, he got one of the three angles right. Presumably he remembered that there is a right triangle that has sides 3-4-5 and applied the side-side-side congruence theorem to conclude that every 3-4-5 triangle is a right triangle.

Unfortunately, the story gets worse. The high school student then replied: "It's 30-60-90," which we all know is wrong, because the cosine of 30 degrees is $\sqrt{3}/2$, not 4/5.

And the story gets yet worse. The AP reporter asserted that the correct answer was 90, 53.1, and 36.9 degrees, which—although not really wrong—does not address the question of whether or not the governor should have known these angles. A later story [**16**] did better. It quoted a retired math professor who said "I don't think those are very well known angles" and "I wouldn't expect many mathematicians to know that."

A good answer to the student's question is provided by the following theorem:

THE GOVERNOR'S THEOREM. *If a right triangle has integer side lengths, then the acute angles are irrational, when measured in degrees.*

Looking at the trig table we all memorized, we notice that all the angles are integers when measured in degrees, so no one—much less a governor—should be expected to answer the student's question.

[1]This is the text for the address at the 2006 MAA State Dinner for Georgia, held at Mercer University in April.

How to prove the theorem? One way is to get it as a corollary of Lehmer's Theorem, which we will prove in the next section.

Remarks. The Governor's Theorem appears as problem 239b in [**15**]. The theorem is also true if the word "degrees" is replaced by "radians". That is a standard consequence of the Lindemann-Weierstrass Theorem, and is Exercise 1 in §4.13 of [**8**].

Degrees

Since all the angles in the standard trig table are rational when measured in degrees, below we only consider angles of that type.

Naively, we might want to fill our trig tables with angles θ such that θ (measured in degrees), $\sin\theta$, and $\cos\theta$ are all rational. But the Governor's Theorem—still to be proved—tells us that there are no such θ's with $0 < \theta < 90°$. So we should include more angles in our trig table.

For example, our standard trig table includes familiar friends like 45 degrees, for which sine and cosine are both $1/\sqrt{2}$, which is not rational. Rather, $1/\sqrt{2}$ is *algebraic*, i.e., there is a nonzero polynomial $f(x)$ with rational coefficients such that $f(1/\sqrt{2}) = 0$, namely $f(x) = 2x^2 - 1$.

We might ask: What rational angles are such that $\sin\theta$ and $\cos\theta$ are algebraic? The hope would be that these angles θ would make up a nice trig table. But this doesn't work, because $\sin\theta$ and $\cos\theta$ are *always* algebraic when θ is rational. Gauss said that this was well known in Article 337 of his *Disquisitiones* [**4**], and indeed we can see it by using multiple-angle formulas (or reading the rest of this section).

So we need a finer notion of "nice" to pick out angles θ that one would want in a trig table. Recall that the *degree* of an algebraic number is defined to be the minimum of $\deg f(x)$ as $f(x)$ varies over the nonzero rational polynomials such that $f(r)$ is zero. For example, rational numbers have degree 1 whereas $1/\sqrt{2}$ and $\sqrt{3}/2$ have degree 2.

Write θ as $360k/n$ where k and n are relatively prime natural numbers. The famous number-theorist D.H. Lehmer proved in [**10**]:

LEHMER'S THEOREM. *If $n \geq 3$, then the degree of $\cos(360k/n)$ is $\phi(n)/2$. If n is positive and $\neq 4$, then the degree of $\sin(360k/n)$ is*

$$\begin{cases} \phi(n) & \text{if } \gcd(n, 8) = 1 \text{ or } 2 \\ \phi(n)/4 & \text{if } \gcd(n, 8) = 4 \\ \phi(n)/2 & \text{if } \gcd(n, 8) = 8. \end{cases}$$

In the theorem, the symbol ϕ denotes Euler's ϕ-function. Recall that $\phi(1)$ is defined to be 1 and that $\phi(mn) = \phi(m)\phi(n)$ for relatively prime numbers m and n. Finally, for p a prime we have:

$$\phi(p^e) = p^{e-1}(p-1).$$

The first few values of ϕ are given by the table

n	1	2	3	4	5	6	7	8
$\phi(n)$	1	1	2	2	4	2	6	4

Let us check Lehmer's Theorem on 30 degrees. We can write 30 as 360/12. The theorem says that $\cos(30) = \sqrt{3}/2$ has degree $\phi(12)/2 = 2$—which is true—and that $\sin(30) = 1/2$ has degree $\phi(12)/4 = 1$—which is also true. Also, the excluded values of n are genuine exclusions. For cosine, we exclude $n = 1, 2$, corresponding to multiples of 180 degrees. The cosines of such angles are 0 or ± 1, hence of degree 1 (and not $1/2$). For sine, we exclude $n = 4$, corresponding to 90 and 270 degrees, for which sine is 0, hence of degree 1 (and not $1/2$).

Note that Lehmer's Theorem implies Gauss's assertion that $\sin\theta$ and $\cos\theta$ are algebraic for every rational angle θ.

Note also the asymmetry between sine and cosine in Lehmer's Theorem. The author finds this unsettling, in light of the complementary angle formula $\sin\theta = \cos(90° - \theta)$.

We apply Lehmer's Theorem to draw up a table. In each row of the following table, we specify a degree d (less than 8). For each degree, we list the denominators n such that $\sin(360k/n)$ has degree d or $\cos(360k/n)$ has degree d. To construct such a table, one applies Lehmer's Theorem to all the natural numbers n such that sine or cosine of $360/n$ can have degree at most 7; finding the list of such n's is an exercise with the definition of ϕ given above. (Lehmer's article gives a similar table, but its entries for sine are incorrect because the theorem for sine is not stated correctly there; Niven's book [**11**] has the correct version of the theorem—reproduced above—but no table. Also, in Lehmer's cosine table there is a 36 that should be a 30.)

TABLE 1: Denominators n such that $\sin 360k/n$ or $\cos 360k/n$ have degree $d < 8$.

d	sine	cosine
1	1, 2, 4, 12	1, 2, 3, 4, 6
2	3, 6, 8, 20	5, 8, 10, 12
3	28, 36	7, 9, 14, 18
4	5, 10, 16, 24, 60	15, 16, 20, 24, 30
5	44	11, 22
6	7, 9, 14, 18, 52, 84	13, 21, 26, 28, 36, 42
7	none	none

The $d = 1$ row of Table 1 says that—for θ a rational acute angle—$\sin\theta$ is rational if and only if θ is 30 degrees, and $\cos\theta$ is rational if and only if θ is 60 degrees. (For a proof of this using trig identities, see [**13**].) This proves the Governor's Theorem.

The reasoning in the previous paragraph also shows that no rational angle has a sine of $3/4$. This answers the question posed in the introduction, albeit in an unsatisfying way.

Now that we understand how to apply Lehmer's Theorem, we prove it. Following the stated goal of this paper, we use Galois theory. A proof that appears more concrete but is really the same can be found in Section 3.4 of [**11**]. We write θ for $360k/n$.

Put

$$z := e^{2\pi ik/n} = \cos\theta + i\sin\theta,$$

a primitive nth root of unity in the complex numbers. Then

$$\cos\theta = \frac{z + \bar{z}}{2},$$

where $\bar{z}$ denotes the complex conjugate of z. Since k and n are relatively prime, there is some natural number ℓ such that $k\ell$ is congruent to $n - k \pmod{n}$. Therefore,

$$\bar{z} = e^{-2\pi ik/n} = (e^{2\pi ik/n})^{\ell} = z^{\ell}.$$

In particular, $\bar{z}$ belongs to $\mathbb{Q}(z)$. (Alternatively, we can remember the fact that $\mathbb{Q}(z)$ is a Galois extension of $\mathbb{Q}$.) Now z satisfies the equation

$$z^2 - 2z\cos\theta + 1 = 0$$

because $z\bar{z} = 1$, so the dimension of $\mathbb{Q}(z)$ over $\mathbb{Q}(\cos\theta)$ is 1 or 2. On the other hand, these two field extensions are not the same: z cannot belong to $\mathbb{Q}(\cos\theta)$ because z is not a real number (because n is at least 3) and $\mathbb{Q}(\cos\theta)$ consists of real numbers. That is, the dimension $[\mathbb{Q}(z) : \mathbb{Q}(\cos\theta)]$ is 2. Since the dimension of $\mathbb{Q}(z)$ over $\mathbb{Q}$ is $\phi(n)$—see e.g. [**2**, §13.6, Cor. 42]—we obtain:

$$[\mathbb{Q}(\cos\theta) : \mathbb{Q}] = \frac{[\mathbb{Q}(z) : \mathbb{Q}]}{[\mathbb{Q}(z) : \mathbb{Q}(\cos\theta)]} = \frac{\phi(n)}{2}.$$

This proves the part of Lehmer's Theorem regarding cosine.

The result for sine can be derived from the result for cosine using the complementary angle formula

$$\sin\left(360\frac{k}{n}\right) = \cos\left(360\frac{k}{n} - 90\right) = \cos\left(360\frac{4k-n}{4n}\right).$$

This part of the proof is identical to the one on p. 38 of Niven's book, so we just sketch it. First, observe that since n is not 4, the fraction $(4k-n)/(4n)$, when put in lowest terms, has denominator at least 3, so we may indeed apply the cosine result. Next, divide the proof into cases depending on the highest power of 2—say 2^e—dividing n. For example, if $e = 1$, then $(4k-n)/(4n)$ in lowest terms has denominator $2n$, and $\sin(\theta)$ has degree $\phi(2n)/2$. But $n = 2m$ for some m odd, so

$$\frac{\phi(2n)}{2} = \frac{\phi(4m)}{2} = \frac{\phi(4)\phi(m)}{2} = \phi(m) = \phi(2)\phi(m) = \phi(n).$$

The remaining cases are left to the reader. This completes the proof of Lehmer's Theorem.

For later use, we note that $\mathbb{Q}(z)$ is Galois over $\mathbb{Q}$ with abelian Galois group, as we learned in our Galois theory course. Consequently, every subfield of $\mathbb{Q}(z)$ is also Galois over $\mathbb{Q}$ with abelian Galois group, including $\mathbb{Q}(\cos\theta)$ and any subfield of it.

Remark. The proof of the cosine part of Lehmer's Theorem amounts to the observation that $\mathbb{Q}(\cos\theta)$ is the maximal real subfield of $\mathbb{Q}(z)$, i.e., the intersection of $\mathbb{Q}(z)$ with the real numbers.

An expanded trig table

Our motivation is to answer the question: Why does the standard trig table include exactly the angles 0, 30, 45, 60, and 90 degrees? If you want to pick angles to put in a trig table and you know Table 1, it would make sense to include exactly those rational angles θ such that $\sin\theta$ and $\cos\theta$ both have degree at most d for some choice of d. If you do this with $d = 2$, you find $n = 1, 2, 3, 4, 6, 8, 12$ corresponding to angles that are multiples of 360, 180, 120, 90, 60, 45, or 30 degrees. Surprise! You find the standard trig table. This explains why the standard trig table includes exactly the angles that it does.

Taking $d = 3$, we find no new angles by Table 1.

Taking $d = 4$, we find new denominators $n = 5, 10, 16, 20, 24$, corresponding to angles that are multiples of 72, 36, $22\frac{1}{2}$, 18, and 15 degrees respectively. An expanded trig table including these angles is given in Table 2 below.

TABLE 2: An expanded trig table including 72° and all rational angles θ between 0° and 45° such that $\sin\theta$ and $\cos\theta$ have degree ≤ 4.

θ (in degrees)	$\sin\theta$	$\cos\theta$
0	0	1
15	$\frac{-1+\sqrt{3}}{2\sqrt{2}}$	$\frac{1+\sqrt{3}}{2\sqrt{2}}$
18	$\frac{-1+\sqrt{5}}{4}$	$\frac{\sqrt{5+\sqrt{5}}}{2\sqrt{2}}$
$22\frac{1}{2}$	$\frac{\sqrt{2-\sqrt{2}}}{2}$	$\frac{\sqrt{2+\sqrt{2}}}{2}$
30	$1/2$	$\sqrt{3}/2$
36	$\frac{\sqrt{5-\sqrt{5}}}{2\sqrt{2}}$	$\frac{1+\sqrt{5}}{4}$
45	$1/\sqrt{2}$	$1/\sqrt{2}$
	$\vdots$	
72	$\frac{\sqrt{5+\sqrt{5}}}{2\sqrt{2}}$	$\frac{-1+\sqrt{5}}{4}$

This table warrants some remarks. First: How does one compute $\sin\theta$ and $\cos\theta$ for the new angles θ? The only difficult one turns out to be $\theta = 72°$; once $\sin 72°$ and $\cos 72°$ have been computed, all the other entries can be filled in using the half-angle and complementary-angle formulas and known entries from the standard table. To compute $\sin 72°$ and $\cos 72°$, we follow the proof of Lehmer's Theorem and put

$$z := \cos 72° + i \sin 72° = e^{2\pi i/5}.$$

Note that $z^5 = 1$, i.e., z is a 5th root of unity. Since 5 is of the form $2^{2^1} + 1$, it is a *Fermat prime*, and Gauss gave a general method for computing pth roots of unity for such primes p, see [**4**, §VII], [**3**, §§20–27], or [**17**, Ch. 12]. An explicit form of z can be found in almost any book on Galois theory where this method is presented. (Alternatively, one can use trigonometric formulas as in [**6**, pp. 39, 40] or geometry as in [**1**].)

Second, some of the entries are noteworthy. For example, the half-angle formula gives

$$\sin 15° = \sqrt{\frac{1-\cos 30°}{2}} = \frac{\sqrt{2-\sqrt{3}}}{2},$$

but in the table we find $(-1+\sqrt{3})/2\sqrt{2}$. Which is it? Indeed they are equal, since they are positive real numbers whose squares agree. The second version is in the table due to the author's general aversion to "nested" square roots.

We can similarly calculate $\sin 22\frac{1}{2}°$ by applying the half-angle formula to $\cos 45°$ and finding

$$\sin 22\frac{1}{2}^\circ = \frac{\sqrt{2-\sqrt{2}}}{2}.$$

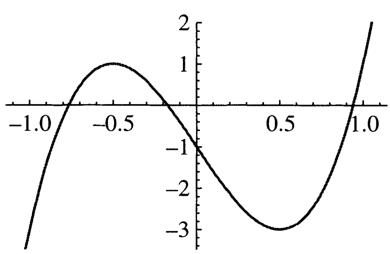
This nested expression is the one listed in the table. This is not due to some laziness on the author's part, but rather to mathematical reality: this expression cannot be rewritten to eliminate the nesting. Indeed, in the previous paragraph, the field $\mathbb{Q}((-1+\sqrt{3})/2\sqrt{2})$ is (exercise!) the same as $\mathbb{Q}(\sqrt{2},\sqrt{3})$, a Galois extension of $\mathbb{Q}$ whose Galois group is the Klein four-group (exercise!). Here the field $\mathbb{Q}(\sqrt{2-\sqrt{2}}/2)$ is also a Galois extension of $\mathbb{Q}$ of degree 4, but with Galois group $\mathbb{Z}/4\mathbb{Z}$. (This assertion is Exercise 14 in §14.2 in **[2]**.) The reader who verifies these assertions about the Galois groups will see that they imply that $\sin 22\frac{1}{2}^\circ$ cannot be rewritten so as to remove the nested square root.

Bibliographic references. Hoehn **[5]** gives a nice geometric derivation of sine and cosine for the angles 15° and 75°. Generally, one can find $\sin\theta$ and $\cos\theta$ whenever θ is an integer multiple of 3 degrees once one knows the values for $\theta = 36°$ and various trig identities, see e.g. **[18]**. For the general question of unnesting square roots, see **[9]**.

Can we expand the table a little further?

Continuing the procedure from the previous section, we could consider angles θ such that $\sin\theta$ and $\cos\theta$ have degree ≤ 5. By Table 1, no new angles θ are found. (In general, for every odd number $d \geq 3$, the angles θ such that $\sin\theta$ and $\cos\theta$ have degree $\leq d$ are exactly the same angles as those whose sine and cosine have degree $\leq d-1$. Proving this claim is an exercise combining Lehmer's Theorem and the definition of ϕ which we leave to the reader.)

If we consider those angles with degree ≤ 6, there are lots of new angles, but also an ugly phenomenon. Consider the case $n = 18$, corresponding to the angle $\theta = 20°$. Table 1 says that $\cos\theta$ has degree 3,[2] and it is a root of the polynomial $8x^3 - 6x - 1$, as can be seen by the triple-angle formula. (This fact is familiar from a course on Galois theory, because it is part of the standard proof that the angle 20° cannot be constructed with straightedge and compass.) The graph of this polynomial is



It clearly has three real roots. However, Cardano's formula for finding the roots of a cubic polynomial gives the three seemingly complex roots[3]

$$x = \sqrt[3]{\frac{1+\sqrt{-3}}{16}} + \sqrt[3]{\frac{1-\sqrt{-3}}{16}}.$$

[2]We only include 20° at the degree ≤ 6 stage because sin(20°) has degree 6.

[3]Each complex number has three cube roots, so the displayed expression a priori gives 9 complex numbers. The roots of the polynomial $8x^3 - 6x - 1$ are the values of x where the sum of the two cube roots is a real number.

This is an example of the famous "casus irreducibilis" of the cubic formula, and it is well known that one cannot avoid using complex numbers in writing down these roots (see e.g., [**2**, §14.7] or keep reading this article). Since cos(20°) is so ugly, we shouldn't include it in our table.

In fact, this example is typical. Write $\theta = 360k/n$ where k and n are relatively prime integers. We have:

THEOREM. *The numbers* $\cos\theta$ *and* $\sin\theta$ *can be written using only rational numbers, addition, subtraction, multiplication, division, and roots of* positive *numbers if and only if* $\phi(n)$ *is a power of* 2.

In the example of 20 degrees above, $\phi(18)$ is 6, which is not a power of 2. When we prove the theorem below, we will have in particular proved that cosine of 20 degrees cannot be written using only field operations and roots of positive numbers.

The theorem says that the only candidates for angles θ to be added to our trig table are those $\theta = 360k/n$ where $\phi(n)$ is a power of 2. Such natural numbers n are precisely those n such that regular n-gon is constructible with compass and straightedge—proved by Gauss and Wantzel. Gauss listed the 38 such $n \leq 300$ at the end of [**4**], and in principle the list is known for $n < 2^{2^{22}} + 1$ (approximately 2×10^{1262611}), see [**12**, Seq. A003401].

In the next section, we will investigate what new entries could be added to our trig table. But first we prove the theorem stated above. Write E for the field consisting of real numbers that can be written as in the statement of the theorem, i.e., that can be written using only rational numbers, addition, subtraction, multiplication, division, and roots of positive real numbers.

We first prove the "if" direction, i.e., we suppose that $\phi(n)$ is a power of 2. This is the easier implication, and the proof proceeds just as for constructibility of a regular n-gon. By the observation just after the proof of Lehmer's Theorem, the extension $\mathbb{Q}(\cos\theta)$ of $\mathbb{Q}$ is Galois with abelian Galois group of order $\phi(n)/2$. Therefore, there is a chain of fields

$$\mathbb{Q} = K_0 \subset K_1 \subset \cdots \subset K_r = \mathbb{Q}(\cos\theta)$$

such that each K_{i+1} is a 2-dimensional extension of K_i. It follows that K_{i+1} is obtained from K_i by adjoining the square root of some element $a_i \in K_i$. Since K_{i+1} is contained in $\mathbb{Q}(\cos\theta)$ and hence in $\mathbb{R}$, the element a_i is positive. We conclude that $\cos\theta$ belongs to E. The pythagorean identity $\sin^2\theta + \cos^2\theta = 1$ implies that

$$\sin\theta = \pm\sqrt{1 - \cos^2\theta},$$

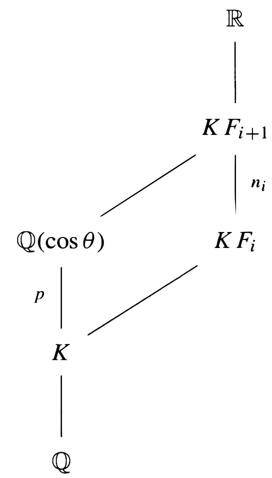
hence $\sin\theta$ also belongs to E. This completes the proof of the "if" direction.

To prove the more difficult "only if" direction, we suppose that $\phi(n)$ is not a power of 2, i.e., is divisible by an odd prime p. Note that this implies that n is at least 7. For sake of contradiction, suppose that $\cos\theta$ or $\sin\theta$ belongs to E. In the latter case, the pythagorean identity implies that $\cos\theta$ belongs to E, so in any case we have that $\cos\theta$ belongs to E. Lehmer's Theorem gives that p divides the dimension of $\mathbb{Q}(\cos\theta)$ over $\mathbb{Q}$. We observed after the proof of that theorem that $\mathbb{Q}(\cos\theta)$ is Galois over $\mathbb{Q}$ with abelian Galois group, so there exists a Galois extension K of $\mathbb{Q}$ contained in $\mathbb{Q}(\cos\theta)$ such that the dimension of $\mathbb{Q}(\cos\theta)$ over K is p.

Since $\cos\theta$ belongs to E, there is a tower of fields

$$\mathbb{Q} = F_0 \subset F_1 \subset \cdots \subset F_r \subset \mathbb{R}$$

such that F_r contains $\cos\theta$ and each F_{i+1} is obtained from F_i by adjoining a real n_ith root of some positive $a_i \in F_i$. By inserting additional terms in this tower if necessary, we may assume that n_i is a prime number for all i. Fix i such that the compositum KF_i does not contain $\cos\theta$, but KF_{i+1} does.



The field $KF_i(\cos\theta)$ properly contains KF_i and (since K is Galois over $\mathbb{Q}$) the dimension of $KF_i(\cos\theta)$ over KF_i divides the dimension of $K(\cos\theta)$ over K, which is the prime p. So $KF_i(\cos\theta)$ has dimension p over KF_i. On the other hand, $KF_{i+1} = KF_i(\sqrt[n_i]{a_i})$ is a proper extension of KF_i of degree dividing the prime n_i; since $KF_i(\cos\theta)$ is contained in KF_{i+1}, we conclude that $n_i = p$ and

$$KF_i(\cos\theta) = KF_{i+1} = KF_i(\sqrt[p]{a_i})$$

is Galois over KF_i. It follows that KF_{i+1} contains a full set of pth roots of unity. But this is impossible because p is odd and KF_{i+1} consists of real numbers. This contradicts the assumption that $\cos\theta$ or $\sin\theta$ belongs to E and completes the proof of the theorem.

Remarks. (i) One reader of this paper asked what the correct theorem would be if the word "positive" were removed from the statement of the theorem. Or, to restate the question: What rational angles θ have values of $\cos\theta$ and $\sin\theta$ that are solvable by radicals? In the proof of Lehmer's Theorem, we saw that $\cos\theta$ generates an abelian extension of $\mathbb{Q}$, so the answer is "all rational angles θ" by Galois's criterion.

(ii) The proof of the difficult direction of the theorem did not involve the result about constructibility of regular n-gons, despite the similarity in the two statements. Indeed, the restriction that $\phi(n)$ be a power of 2 arose here because $\phi(n)$ is—up to multiplication by a power of 2—the degree of a real Galois extension $Q(\cos\theta)$ of Q obtained by taking roots of positive real numbers. This implied that $Q(\cos\theta)$ can be obtained by adjoining only square roots, hence that $\phi(n)$ is a power of 2. In contrast, in the theorem about n-gons, you begin with the hypothesis that you can only take square roots.

(iii) A more general version of the proof above is given in [**7**] or with Exercises 12–14 in §14.7 of [**2**]. Our proof is much simpler because the number that we assume belongs to E—namely, $\cos\theta$—generates a Galois extension of $\mathbb{Q}$.

What the next few numbers look like

By the previous section, our trig table should only include angles $360k/n$ where $\phi(n)$ is a power of 2. If we want to expand Table 2, how should we do it?

We have already considered the denominators $n = 1, 2, \ldots, 6, 8, 10, 12, 16, 20, 24$. The smallest natural number n not in that list and with $\phi(n)$ a power of 2 is $n = 15$, corresponding to angles that are multiples of 24 degrees. To compute $\sin 24^\circ$ and $\cos 24^\circ$, we can use the familiar trick from the construction of regular n-gons. Namely, factor 15 as $3 \cdot 5$ and recall that we know $e^{2\pi i/3}$ and $e^{2\pi i/5}$ because we know the values of sine and cosine for 120° and 72°, so we can explicitly compute

$$e^{2\pi i/3} \cdot e^{2\pi i/5} = e^{2\pi i \cdot 8/15}.$$

Moreover, $2 \cdot 8 = 15 + 1$, so

$$(e^{2\pi i \cdot 8/15})^2 = e^{2\pi i/15}.$$

Performing these computations and extracting the real part, we find:

$$8\cos 24^\circ = (1+\sqrt{5}) + (\sqrt{5}-1)\sqrt{\frac{3}{2}(5+\sqrt{5})} = (1+\sqrt{5}) + 2\sqrt{\frac{3}{2}(5-\sqrt{5})}.$$

The next n for which the corresponding θ does not appear in Table 2 is $n = 17$, corresponding to the angle $\theta = 21\frac{3}{17}^\circ$. Seventeen is a Fermat prime, and the values of sine and cosine can be computed by Gauss's algorithm. Gauss himself gave the following explicit formula for $\cos\theta$ in Article 365 of [**4**]:

$$\begin{aligned} 16\cos 21\frac{3}{17}^\circ &= -1 + \sqrt{17} + \sqrt{34 - 2\sqrt{17}} \\ &\quad + 2\sqrt{17 + 3\sqrt{17} - \sqrt{34 - 2\sqrt{17}} - 2\sqrt{34 + 2\sqrt{17}}} \end{aligned}$$

These numbers look pretty ugly! Another strategy is to apply a friendly trig identity like a half-angle formula to entries in the table. Doing this with 18 degrees, we find for example that

$$\cos 9^\circ = \frac{1}{2}\sqrt{2 + \sqrt{\frac{5+\sqrt{5}}{2}}}.$$

The reader is encouraged to continue along these lines until their personal limits of expression complexity are attained.

Summary

We observed that we can construct a trig table for each natural number d by including precisely those rational angles θ such that $\cos\theta$ and $\sin\theta$ are algebraic of degree at most d. For $d = 1$, the table only included multiples of 90°. For $d = 2$, we got the standard trig table consisting of multiples of 30° and 45°. For $d = 4$, we found a larger table, and we exhibited a portion of it in Table 2. Unfortunately, for $d \geq 6$, it is impossible to write the cosine of some angles without using complex numbers. Further investigation revealed that, without using complex numbers, we can only write down $\cos\theta$ and $\sin\theta$ if $\phi(n)$ is a power of 2, where n is the denominator of θ. The proof of this last fact was different from the similar-sounding result about constructibility of regular n-gons with straightedge and compass.

Acknowledgments. I thank Steve Edwards, Julia Garibaldi, Fred Helenius, and an anonymous referee for their comments and for pointing out references.

REFERENCES

1. B. Bradie, Exact values for the sine and cosine of multiples of 18°—a geometric approach, *College Math. J.* **33** (2002) 318–319.
2. D. S. Dummit and R. M. Foote, *Abstract Algebra*, 3rd ed., Wiley, 2004.
3. H. M. Edwards, *Galois Theory*, Springer, 1984.
4. C. F. Gauss, *Disquisitiones Arithmeticae*, trans. by A. A. Clarke, rev. by W. C. Waterhouse, Springer, 1986.
5. L. Hoehn, Proof without words, *College Math. J.* **35** (2004) 282.
6. H. E. Huntley, *The Divine Proportion: A Study in Mathematical Beauty*, Dover, 1970.
7. I. M. Isaacs, Solution of polynomials by real radicals, *Amer. Math. Monthly* **92** (1985) 571–575.
8. N. Jacobson, *Basic Algebra 1*, 2nd ed., W. H. Freeman, 1985.
9. S. Landau, How to tangle with a nested radical, *Math. Intelligencer* **16**(2) (1994) 49–55.
10. D. H. Lehmer, A note on trigonometric algebraic numbers, *Amer. Math. Monthly* **40** (1933) 165–166.
11. I. Niven, *Irrational Numbers*, Carus Mathematical Monographs, vol. 11, Mathematical Association of America, 1956.
12. N. Sloane et al., Online encyclopedia of integer sequences, online at `http://www.research.att.com/~njas/sequences/Seis.html`
13. J. M. H. Olmsted, Rational values of trigonometric functions, *Amer. Math. Monthly* **52** (1945) 507–508.
14. M. Schneider, Teenager stumps Bush with pop math quiz, Associated Press, 6 July 2004.
15. D. O. Shklarsky, N. N. Chentzov, and I. M. Yaglom, *The USSR Olympiad Problem Book: Selected Problems and Theorems of Elementary Mathematics*, Dover, 1994.
16. St. Petersburg Times (FL), Query stumps governor, but it wasn't easy, was it?, 7 July 2004.
17. J.-P. Tignol, *Galois' Theory of Algebraic Equations*, World Scientific, 2001.
18. Wikipedia, Exact trigonometric constants, available at `http://en.wikipedia.org/wiki`, version 12 November 2006.

NOTES

π to Thousands of Digits from Vieta's Formula

RICK KREMINSKI
Texas A & M University – Commerce
Commerce, TX 75429
kremin@boisdarc.tamu-commerce.edu

Introduction

Two infinite products for π, Wallis's and Vieta's, are well-known and striking (and, surprisingly, related—see e.g. [**1**]):

$$\frac{\pi}{2} = \frac{2\cdot 2}{1\cdot 3}\frac{4\cdot 4}{3\cdot 5}\frac{6\cdot 6}{5\cdot 7}\cdots\frac{2n\cdot 2n}{(2n-1)\cdot(2n+1)}\cdots \qquad (\textit{Wallis})$$

$$\frac{\pi}{2} = \frac{2}{\sqrt{2}}\frac{2}{\sqrt{2+\sqrt{2}}}\frac{2}{\sqrt{2+\sqrt{2+\sqrt{2}}}}\frac{2}{\sqrt{2+\sqrt{2+\sqrt{2+\sqrt{2}}}}}\cdots \qquad (\textit{Vieta})$$

While both formulas are mysterious and beautiful, the convergence of Wallis's formula is painfully slow. Vieta's is much better, although Vieta himself was able to approximate π only to 9 digits past the decimal point in 1593. In this article, we'll describe a way to accelerate the convergence of Vieta that is completely accessible to calculus students, and how we stumbled upon it through experimentation. We were somewhat astonished when we used our accelerated (Vieta) to approximate π to over 300,000 digits.

Derivations of (Wallis) and (Vieta) appear in many places; we briefly highlight a few. Typical derivations of (Wallis) often involve comparing the results of integrating $\sin^{2n+1}(x)$ and $\sin^{2n}(x)$ by parts, as in calculus texts such as [**2**]. Other derivations require the use of the infinite product for $\sin(z)$, which is not fully accessible to the beginning student (although streamlined proofs exist, e.g., [**3**]). The main disappointment with (Wallis) is that the convergence is slow. For instance, doubling the partial product of the first fifty terms yields an approximation to π of $3.12607\ldots$, not very satisfying (i.e., accurate to within $1.6\cdot 10^{-2}$). A faster product for π which is an acceleration of (Wallis) (and which uses Euler's tranformation of series) appears in [**4**].

By contrast, Vieta's formula is easier to derive, even to those with a mathematical sophistication at the level of beginning calculus. Vieta's original argument is reproduced in translation in [**5**]; an example of a recent examination is [**6**]. A typical derivation today uses trigonometric identities, and the sole calculus fact that the limit as $u\to 0$ of $\sin(u)/u = 1$:

$$\begin{aligned}\sin(x) &= 2\cos\left(\frac{x}{2}\right)\sin\left(\frac{x}{2}\right)\\ &= 2^2\cos\left(\frac{x}{2}\right)\cos\left(\frac{x}{4}\right)\sin\left(\frac{x}{4}\right)\end{aligned}$$

$$= 2^3 \cos\left(\frac{x}{2}\right)\cos\left(\frac{x}{4}\right)\cos\left(\frac{x}{8}\right)\sin\left(\frac{x}{8}\right)$$

$$\vdots$$

$$= 2^n \left(\prod_{k=1}^{k=n} \cos\left(\frac{x}{2^k}\right)\right)\sin\left(\frac{x}{2^n}\right) \tag{1}$$

Divide each side by x, then take the limit as $n \to \infty$ to get

$$\frac{\sin(x)}{x} = \left(\prod_{k=1}^{\infty}\cos\left(\frac{x}{2^k}\right)\right)\lim_{n\to\infty}\frac{\sin(x/2^n)}{x/2^n} = \prod_{k=1}^{\infty}\cos\left(\frac{x}{2^k}\right).$$

Set $x = \pi/2$, then repeatedly apply the half-angle formula

$$\cos\left(\frac{\theta}{2}\right) = \sqrt{\frac{1+\cos(\theta)}{2}};$$

taking the reciprocal yields (Vieta).

Unlike the convergence of (Wallis), doubling the partial product of the first fifty terms of (Vieta) approximates π accurate to about 30 digits (i.e., surprisingly accurate to within $1.1 \cdot 10^{-30}$). In this article we'll see how to bound carefully the error in (Vieta), and use the form of the error term to accelerate dramatically the convergence of the partial products. For instance, from just the first fifty terms of (Vieta), we will obtain π accurate to over 900 digits (i.e., accurate to within $3.3 \cdot 10^{-908}$). From fewer than 1200 terms, we found π to over 300,000 digits. And all this comes from an idea familiar to many, although in a rather different context. Namely, our acceleration algorithm is (arguably) one of the purest examples of exploiting a certain form of error term—essentially the same form that is embodied in the widely used Richardson extrapolation and Romberg integration algorithms.

Before we proceed, you may wonder how (Vieta) compares with familiar series expansions for π, such as Gregory's series

$$\pi = 4 - \frac{4}{3} + \frac{4}{5} - \frac{4}{7} + \cdots + (-1)^{k-1}\frac{4}{2k-1} + \cdots$$

and

$$\frac{\pi^2}{6} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots + \frac{1}{n^2} + \cdots.$$

Proofs of the latter without using Fourier series are increasingly accessible to calculus students, as in [**7**]. The most powerful acceleration methods for such sums have errors comparable to those arising from use of the Euler-Maclaurin summation formula, $\sum_{i=n}^{\infty} f(i) = \int_n^{\infty} f(u)du + f(n)/2 - \sum_{k=1}^{\infty}[B_{2k}/(2k)!]f^{2k-1}(n)$, where the B_j are the Bernoulli numbers and $f(x) = 1/x^2$. Note that Euler-Maclaurin is also increasingly accessible to beginning students, as in [**8**]. To illustrate its implementation, the partial sum of the first fifty terms of $\frac{1}{n^2}$ is within .02 of the value of the infinite series, a fairly unsatisfying result. But without too much trouble, we can obtain π to 100 digits from Euler-Maclaurin. Using an extra 55 correction terms, which involve the Bernoulli numbers B_2 through B_{110} and derivatives of $f(x) = 1/x^2$ through the 109th derivative, the Euler-Maclaurin estimate is accurate to within $1.6 \cdot 10^{-100}$. While impressive, compare this with the 30-digit accuracy from the first fifty terms of (Vieta), as cited above. Moreover, we will see that (Vieta), suitably accelerated, yields error less than 10^{-907} using fifty terms.[1]

Convergence and accelerated convergence of Vieta. Let p_n denote the product of the first n terms in (Vieta), i.e., $p_n = 1/[\cos(\pi/4)\cdot\cos(\pi/8)\cdots\cos(\pi/2^{n+1})]$. The key observation is that (1) applied to $x = \pi/2$ yields $p_n = 2^n\sin(\pi/2^{n+1})$. So, by the Maclaurin series for sine,

$$\frac{\pi}{2} - p_n = \frac{\pi}{2} - 2^n\left(\frac{\pi}{2^{n+1}} - \frac{\pi^3}{3!(2^{n+1})^3} + \cdots + \frac{(-1)^m\pi^{2m+1}}{(2m+1)!(2^{n+1})^{2m+1}} - \cdots\right)$$

$$= \frac{\pi}{2} - \frac{\pi}{2} + \frac{\pi^3}{3!2^3 4^n} + \cdots + \frac{(-1)^{m+1}\pi^{2m+1}2^n}{(2m+1)!2^{n+1}(2^{n+1})^{2m}} + \cdots$$

Thus we have

$$\frac{\pi}{2} = p_n + \frac{k_1}{4^n} + \frac{k_2}{16^n} + \frac{k_3}{64^n} + \frac{k_4}{256^n} + \cdots \quad \text{where } k_m = \frac{(-1)^{m+1}\pi^{2m+1}}{2^{2m+1}(2m+1)!} \tag{2}$$

Explicitly $k_1 = \pi^3/48 = 0.64596\ldots$, $k_2 = -\pi^5/3840 = -0.07969\ldots$, etc.

Before proceeding, observe that $\log_{10} 4 = .602\ldots$ so that (2) immediately hints that $|\pi - 2p_n| \sim 10^{-.6n}$, i.e., $2p_n$ approximates π to about $.6n$ digits. To illustrate, $\pi - 2p_{50} \approx 1.02\cdot 10^{-30}$ and $\pi - 2p_{100} \approx 8.04\cdot 10^{-61}$. A glance at Table 1 will show how uncanny the estimate of $.6n$ digits of accuracy really is. This phenomenon is explained as follows. Instead of using the Maclaurin series for $\sin(x)$, we'll use Taylor's theorem with remainder: $\sin(x) = x - x^3/3! + x^5\cos(c)/5!$, where c is between 0 and x. Applying this to $p_n = 2^n\sin(\pi/2^{n+1})$, similar algebra that led to (2) yields $|\pi - 2p_n| = 2|k_1|\cdot 4^{-n}|1 + \cos(c_n)\pi^2/(4\cdot 20\cdot 4^n)|$, i.e., $|\pi - 2p_n| \approx 1.292\cdot 4^{-n}|1 + \cos(c_n).1234/4^n|$, so $|\pi - 2p_n| \lesssim 1.3\cdot 10^{-.6n}$.

Next, we use (2) to develop our acceleration algorithm. It follows almost immediately:

$$\frac{\pi}{2} = p_{n+1} + \frac{k_1}{4^{n+1}} + \frac{k_2}{16^{n+1}} + \frac{k_3}{64^{n+1}} + \frac{k_4}{256^{n+1}} + \cdots \tag{3}$$

thus multiplying (3) by four yields

$$4\frac{\pi}{2} = 4p_{n+1} + 4\frac{k_1}{4^{n+1}} + 4\frac{k_2}{16^{n+1}} + 4\frac{k_3}{64^{n+1}} + 4\frac{k_4}{256^{n+1}} + \cdots \tag{4}$$

so subtracting (2) from (4), then dividing by three yields

$$\frac{\pi}{2} = \frac{4p_{n+1} - p_n}{3} + \frac{l_2}{16^n} + \frac{l_3}{64^n} + \frac{l_4}{256^n} + \cdots \tag{5}$$

where $l_2 = -4k_2/16$, $l_3 = -20k_3/64$, $l_4 = -84k_4/256$, and in general

$$l_j = ((4 - 4^j)k_j/(3\cdot 4^j) \quad \text{for } j \geq 1.$$

But then

$$\frac{\pi}{2} = \frac{4p_{n+2} - p_{n+1}}{3} + \frac{l_2}{16^{n+1}} + \frac{l_3}{64^{n+1}} + \frac{l_4}{256^{n+1}} + \cdots \tag{6}$$

[1]Note that since the Euler-Maclaurin approximation uses 105 terms, rather than just 50, a comparison to accelerated (Vieta) using 105 partial products rather than 50 might be more fair. Remarkably, accelerated (Vieta), applied to the first 105 partial products, yields π to within $7.5\cdot 10^{-3710}$.

Multiplying (6) by sixteen, then subtracting (5), and dividing the result by fifteen yields

$$\frac{\pi}{2} = \frac{16r_{n+1} - r_n}{15} + \frac{m_3}{64^n} + \frac{m_4}{256^n} + \cdots \tag{7}$$

where we have defined $r_n = \frac{4p_{n+1} - p_n}{3}$ and $m_j = ((16 - 4^j)l_j/(15 \cdot 4^j)$ for $j \geq 2$.

As an example,

n	$2p_n$	$2r_n = 2(4p_{n+1} - p_n)/3$	$2(16r_{n+1} - r_n)/15$
10	$\underline{3.141591}422\ldots$	$\underline{3.14159265358975}700\ldots$	$\underline{3.1415926535897932384626}516\ldots$
11	$\underline{3.141592}346\ldots$	$\underline{3.14159265358979}097\ldots$	
12	$\underline{3.141592}577\ldots$		

where we underlined those digits that match the digits of π. More precisely, the tabulation above shows error $7.70\ldots \cdot 10^{-8}$ for $2p_{12}$, improved to error $1.268\ldots \cdot 10^{-22}$ for the rightmost entry above.

Repeating this process produces a recursive sequence of approximations to $\pi/2$ that are increasingly accurate. Define

$$V_{k,n} = \frac{4^k V_{k-1,n+1} - V_{k-1,n}}{4^k - 1} \tag{8}$$

with $V_{0,n} = p_n$ (so e.g., $r_n = V_{1,n}$). We earlier noted that (2) implies $|\pi - 2p_n| \sim 10^{-.6n}$. The constants k_i tend to zero as $n \to \infty$; in (5) the l_i are even smaller; and in (7) the m_i are smaller still. Thus (6) and (7) lead to the more general

ACCELERATION THEOREM. *Let p_n denote the partial product of the first n terms in (Vieta), and let $V_{k,n}$ be the accelerations as defined by* (8). *Then*

$$|\pi - 2V_{k,n}| \lesssim 2C_k 10^{-.6(k+1)n}$$

where C_k tends to zero as k increases:

$$C_k < \frac{1}{4-1}\frac{1}{4^2-1}\frac{1}{4^3-1}\cdots\frac{1}{4^k-1}\frac{\pi^{2k+3}}{2^{2k+3}(2k+3)!}.$$

Sketch of proof. Re-express (5) as $\pi/2 - r_n = l_2/16^n + \cdots$; then $|\pi - 2r_n| \lesssim 2|l_2| \cdot 10^{-1.2n}$. Substituting the form for l_2, and recalling the notation that r_n is $V_{1,n}$, we see that $|\pi - 2V_{1,n}| \lesssim 2C_1 \cdot 10^{-1.2n}$ where $C_1 = |(4 - 4^2)/(3 \cdot 4^2)||k_2|$, i.e., $C_1 = |\frac{4-4^2}{3 \cdot 4^2}|\frac{\pi^5}{2^5 5!}$. Similarly, if we re-examine (7) and express $|\pi - 2V_{2,n}|$ in terms of m_3, then substitute the form for m_3, we have $|\pi - 2V_{2,n}| \lesssim 2C_2 10^{-1.8n}$ where $C_2 = |\frac{16-4^3}{15 \cdot 4^3}||\frac{4-4^3}{3 \cdot 4^3}||k_3|$, i.e., $C_2 = |\frac{16-4^3}{15 \cdot 4^3}||\frac{4-4^3}{3 \cdot 4^3}| \cdot \frac{\pi^7}{2^7 7!}$. These examples with $k = 1$ and $k = 2$ illustrate how, more generally, the inequality for C_k arises in the theorem.

As an application of the acceleration theorem, if $k = n = 35$ then the Theorem yields $C_{35} < 3.39 \cdot 10^{-471}$, which implies that $V_{35,35}$ should be accurate to more than $.6 \cdot 35 \cdot 36 + 471$ digits, i.e., to more than 1227 digits. We tabulate some values in Table 1.

Finally, we note that the approach of successively eliminating terms in the error expansion (2) as implemented in (3)–(8) is strongly reminiscent of Richardson extrapolation (used to improve computations in numerical differentiation) and Romberg integration (used to improve the trapezoidal rule in numerical integration)—see, e.g., **[9]**.

Thus, remarkably, Vieta's venerable formula can produce hundreds of thousands of digits of π.

TABLE 1

k	n	$\pi - 2p_n$	$\pi - 2V_{k,n}$
0	50	$1.0191516\ldots \cdot 10^{-30}$	$1.0191516\ldots \cdot 10^{-30}$
25	25	$1.1474627\ldots \cdot 10^{-15}$	$1.1375426\ldots \cdot 10^{-646}$
49	1	$0.31316552\ldots \cdot 10^{0}$	$3.2232174\ldots \cdot 10^{-908}$
0	70	$9.2691302\ldots \cdot 10^{-43}$	$9.2691302\ldots \cdot 10^{-43}$
35	35	$1.0943057\ldots \cdot 10^{-21}$	$1.1859835\ldots \cdot 10^{-1229}$
35	36	$2.7357643\ldots \cdot 10^{-22}$	$2.5114178\ldots \cdot 10^{-1251}$
35	37	$6.8394109\ldots \cdot 10^{-23}$	$5.3181342\ldots \cdot 10^{-1273}$
35	38	$1.7098527\ldots \cdot 10^{-23}$	$1.1261587\ldots \cdot 10^{-1294}$
69	1	$0.31316552\ldots \cdot 10^{0}$	$3.6016998\ldots \cdot 10^{-1712}$
0	100	$8.0396888\ldots \cdot 10^{-61}$	$8.03968889\ldots \cdot 10^{-61}$
50	50	$1.0191516\ldots \cdot 10^{-30}$	$4.2281275\ldots \cdot 10^{-2447}$
99	1	$0.31316552\ldots \cdot 10^{0}$	$1.3124285\ldots \cdot 10^{-3378}$
572	600	$7.5031491\ldots \cdot 10^{-362}$	$1.6782195\ldots \cdot 10^{-308441}$

Numerical investigations (or "How this all got started")

The path we actually took to discover (2) was circuitous, but may be of interest to those doing numerical work—or those who just like to number-crunch. And we believe that it is quite in the spirit of **[10]**; numerical investigations of convergence may lead to interesting acceleration algorithms. The rapidity with which a series $\sum b_n$ converges is determined by how rapidly the terms b_n being summed tend to zero. Similarly, the rapidity with which an infinite product $\prod a_n$ converges is determined by how quickly the terms a_n being multiplied tend to one. Initially we examined the actual numerical values of the terms a_n in the product:

$$\frac{\pi}{2} = (1.4142135\ldots) \cdot (1.0823922\ldots) \cdot (1.0195911\ldots) \cdot (1.0048385\ldots) \cdot (1.0012055\ldots) \cdot (1.0003012\ldots) \cdot (1.0000753\ldots) \quad (Vieta)$$

Our first observation is how the deviations from one of each factor in (Vieta) vary: each factor appears to be about 1/4th as far away from one as the previous factor. Closer inspection of a_5, a_6, a_7, and a_8 is worthwhile:

$$a_5 = 1.001205996470392602\ldots$$
$$a_6 = 1.000301272041301976\ldots$$
$$a_7 = 1.000075303831095445\ldots$$
$$a_8 = 1.000018825071775513\ldots$$

Indeed, the deviation of a_{k+1} from one does appear to be about 1/4th as large as the deviation of a_n from one:

$$dr_5 = \frac{(a_6 - 1)}{(a_5 - 1)} = 0.249811710646134\ldots$$
$$dr_6 - \frac{(a_7 - 1)}{(a_6 - 1)} = 0.249952935459968\ldots$$
$$dr_7 = \frac{(a_8 - 1)}{(a_7 - 1)} = 0.249988234352289\ldots$$

Above, we define the deviation ratio $dr_n = (a_{n+1} - 1)/(a_n - 1)$. Thus, each factor a_n appears to be of the form $a_n = 1 + \epsilon/4^n$—as opposed to, say, $1 + \epsilon/n^4$. But examining carefully the deviations from this behavior is enlightening. The three ratios dr_5, dr_6, and dr_7 deviate from .25 by .000188289..., .000047064..., and .000011765.... So, surprisingly, they too vary by about a factor of approximately 1/4; specifically, $(dr_6 - 1/4)/(dr_5 - 1/4) = .24995858\ldots$, while $(dr_7 - 1/4)/(dr_6 - 1/4) = .24998964\ldots$. (And these ratios have their own deviation from 1/4, which, also surprisingly, varies by a factor of approximately 1/4.) Thus, each factor's form, empirically, seems to be refined to

$$a_n = 1 + \epsilon_1/4^n + \epsilon_2/16^n + \epsilon_3/64^n + \epsilon_4/256^n + \cdots \qquad (9)$$

In fact, from high-precision values of a_n one can compute $4^n(a_n - 1)$ for large n and thus empirically find ϵ_1; then $16^n(a_n - 1 - \epsilon_1/4^n)$ can be computed numerically to find ϵ_2; and so on. Such computations lead to:

$$\epsilon_1 = 1.233700550136169\ldots$$
$$\epsilon_2 = 1.268347539505240\ldots$$
$$\epsilon_3 = 1.272672326564530\ldots$$
$$\epsilon_4 = 1.273175480652605\ldots$$
$$\epsilon_5 = 1.273232382729394\ldots$$

It appears that the ϵ_k are converging. Since $a_n = \sec(\pi/2^{n+1})$, via Taylor series we deduced

$$\epsilon_n = \frac{|E_{2n}|\pi^{2n}}{(2n)!2^{2n}},$$

where the E_m are the Euler numbers[2] (which are not always positive in some conventions). From known asymptotics, $|E_{2n}| \sim 8\sqrt{n/\pi}(4n/e)^{2n}$; combined with Stirling's formula, we deduced $\epsilon_n \sim 4/\pi$. We emphasize that (9) would not have been discovered had it not been for experimentation. Since there is recurrent interest in nested square roots of two (e.g., **[11]**, **[12]**), such numerical observations may be useful in their own right.

The formula in (9) led us to conjecture the form of the result in (2), but without realizing an explicit formula for the k_m. Instead, empirically

$$k_1 = 0.645964097506246\ldots$$
$$k_2 = -0.079692626246167\ldots$$
$$k_3 = 0.004681754135318\ldots$$
$$k_4 = -0.000160441184787\ldots$$

It was straightforward, albeit tedious, to determine the precise relationships between k_i and ϵ_i; e.g., $k_1 = (\pi/6)\epsilon_1$, $k_2 = (\pi/6)(\epsilon_2/5 - 4\epsilon_1^2/15)$ etc.[3] Initially we only found

[2] As an aside, [**10**, Section 2.2] discusses a curiosity regarding how the Euler numbers arise in Gregory's series $\pi = 4 - 4/3 + 4/5 - 4/7 + \cdots$.

[3] Some details may help. $e_n = \pi/2 - a_1 \cdot a_2 \cdot \cdots \cdot a_n = a_1 \cdot a_2 \cdot \cdots \cdot a_n(t_{n+1} - 1)$ where $t_{n+1} = a_{n+1} \cdot a_{n+2} \cdot \cdots$, the "tail." So $e_n = p_n(t_{n+1} - 1) = (\pi/2 - e_n)(t_{n+1} - 1) = \pi/2(t_{n+1} - 1) - e_n(t_{n+1} - 1)$. Using (0) for the factors a_{n+1}, a_{n+2}, etc, in t_{n+1}, and collecting factors yields $(t_{n+1} - 1) = \epsilon_1(\frac{1}{4^{n+1}} \cdot \frac{1}{1-1/4}) +$ (terms involving ϵ_k for $k \geq 2$); from this one obtains that $(\pi/2)\epsilon_1/3 = k_1$. By keeping track of the ϵ_2 terms, k_2 is found to be $(\pi/6)(\epsilon_2/5 - 4\epsilon_1^2/15)$, and so on. Numerically, these relations for k_1 and k_2 in terms of ϵ_1 and ϵ_2 were confirmed to all digits computed—over 250 digits in each case.

k_1, k_2, and k_3 directly in terms of ϵ_1, ϵ_2, ϵ_3. Since ϵ_k was known in terms of Euler numbers, we saw that $k_1 = \pi^3/48$; $k_2 = -\pi^5/3840$; and $k_3 = \pi^7/645120$. Entering 48, 3840, and 645120 in Neil Sloane's online integer sequence utility [13] yielded just one sequence, whose next term was 185794560; so we conjectured that $k_4 = -\pi^9/185794560$; and this was confirmed, to all digits we had computed for k_4. Using the closed form for the denominators obtained from the website, we conjectured

$$k_m = \frac{(-1)^{m+1}\pi^{2m+1}}{2^{2m+1}(2m+1)!}.$$

This allowed us to conjecture a sum for the "tail," i.e., $\pi/2 - p_n$ in (2). And this is how we discovered the (in hindsight elementary) fact that $p_n = 2^n \sin(\pi/2^{n+1})$, and led us to confirm the earlier conjectures. Thus, numerical investigation circuitously led to what eventually became the self-contained exposition in the previous section.

REFERENCES

1. T. J. Osler, The union of Vieta's and Wallis's products for pi, *American Mathematical Monthly* **106** (1999) 774–776.
2. James Stewart, *Calculus*, 5th ed., Thomson Brooks-Cole, 2003.
3. W. Pribitikin, A Simpler Proof of $\sin(\pi z) = \pi z \Pi (1 - z^2/k^2)$, *American Mathematical Monthly* **108** (2001) 767–768.
4. J. Sondow, A faster product for π and a new integral for ln $\frac{\pi}{2}$, *American Mathematical Monthly* **112** (2005) 729–734.
5. L. Berggren, J. Borwein, and P. Borwein, *Pi: A Source Book*, 3rd ed., Springer-Verlag, 2004.
6. A. Levin, A geometric interpretation of an infinite product for the lemniscate constant, *American Mathematical Monthly* **113** (2006) 510–520.
7. J. Harper, Another simple proof of $1 + 1/2^2 + 1/3^2 + \cdots = \pi^2/6$, *American Mathematical Monthly* **110** (2003) 540–541.
8. V. Lampret, The Euler-Maclaurin and Taylor formulas: twin, elementary derivations, this MAGAZINE **74** (2001) 109–122.
9. L. V. Fausett, *Numerical Methods: Algorithms and Applications*, Prentice Hall, 2003.
10. J. Borwein and D. Bailey, *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, A.K. Peters, 2004.
11. L. D. Servi, Nested square roots of 2, *American Mathematical Monthly*, **110** (2003) 326–330.
12. M. A. Nyblom, More nested square roots of 2, *American Mathematical Monthly* **112** (2005) 822–825.
13. N. Sloane, The online encyclopedia of integer sequences, `www.research.att.com/~njas/sequences/`.

Sum Kind of Asymptotic Trouble

GEORGE W. BENTHIEN, EMERITUS
Space and Naval Warfare Systems Center
San Diego, CA 92152

KEITH J. COATES
Drury University
Springfield, MO 65802
kcoates@drury.edu

Physicists and engineers (and even mathematicians) often find correct solutions to problems quickly and efficiently by not worrying too much about mathematical rigor; for example, integrating a series term by term without explicitly verifying that the conditions under which this is permissible actually obtain. However, it is equally true that a lack of rigor can lead to significant errors when solving problems in applications. In this note, we examine the derivation of constants used to solve a problem in signal processing, found in the book *Applied Analysis*, by Cornelius Lanczos. The derivation involves the asymptotic behavior of the sum of rational expressions and makes use of an approximation that seems intuitive but is incorrect, leading to a wrong value for one of the constants.

The error we examine appears in a book written by an outstanding mathematician. Born in Hungary in 1893, Cornelius Lanczos was a prolific mathematical physicist who made significant contributions in several fields of mathematics and physics, including quantum theory, relativity theory and numerical analysis [**2**]. He developed a number of powerful methods in applied mathematics, among them a matrix algorithm for calculating Fourier coefficients that was essentially the FFT—some 25 years before it was rediscovered by Tukey [**1**]. His publications total over 120 papers and eight books [**4**]. One of those books, *Applied Analysis*, originally published in 1956, has seen a number of reprintings and is still highly regarded among physicists, engineers and applied mathematicians as an extremely useful guide to many of the modern methods in the field of applied analysis. The authors of this paper were using a 1988 republication of that book [**3**] in their own work in acoustics when they discovered the error here examined.

To understand fully the mathematical context of our problem would take the better part of a semester of Fourier Analysis, but we can jump right into the problem itself, which is fairly easy to understand (you might decide to first review "big-Oh" and "little-Oh" notation by, say, peeking at [**5**]):

We have a complex-valued function ϕ of the real variable ν that is $O\left(\frac{1}{\nu}\right)$ as $\nu \to \infty$, and empirically determined real constants a_1, a_2 such that

$$\phi_1(\nu) = \phi(\nu) - \frac{ia_1}{\nu} - \frac{a_2}{\nu^2} \tag{1}$$

is $O\left(\frac{1}{\nu^3}\right)$ as $\nu \to \infty$ (note that the $O\left(\frac{1}{\nu}\right)$ term is imaginary and the $O\left(\frac{1}{\nu^2}\right)$ term is real). Some analysis leads to a needed alternate form of ϕ_1, namely equation 4-19.9 in [**3**]:

$$\phi_1(\nu) = \phi(\nu) - \frac{A_1}{\alpha + 2\pi i\nu} - \frac{A_2}{(\alpha + 2\pi i\nu)^2}, \tag{2}$$

where α, A_1 and A_2 are constants. The goal now is simply stated: to derive α, A_1 and A_2 in terms of a_1 and a_2 so that ϕ_1 in (2) remains $O\left(\frac{1}{\nu^3}\right)$ as $\nu \to \infty$.

There is freedom in the choice of α, and we need only understand here that it must be nonzero. In the derivation of A_1 and A_2, the text claims that, for large ν, (2) is the same as

$$\phi_1(\nu) = \phi(\nu) + \frac{A_1 i}{2\pi\nu} + \frac{A_2 + \alpha A_1}{(2\pi\nu)^2} \tag{3}$$

(4-19.10 in [**3**]), and so comparing with (1) and solving, we arrive at 4-19.11 in [**3**]: $A_1 = -2\pi a_1$ and $A_2 = -4\pi^2 a_2 - \alpha A_1$.

The claim (3) is rather curious. It's not difficult to show that taken literally, the claim is false: (2) and (3) are definitely not equivalent for nonzero α, that is, the ϕ_1 defined in (2) is not the same as the ϕ_1 defined in (3). But the claim was certainly intended to mean that, for large ν, the right side of (3) is an accurate approximation of the right side of (2) and so can be substituted for it. The surprise is that this statement also has problems, stemming from what is meant by the approximation's being "accurate." To understand exactly what the issue is we begin by examining (2):

$$\phi_1(\nu) - \phi(\nu) = -\frac{A_1}{\alpha + 2\pi i\nu} - \frac{A_2}{(\alpha + 2\pi i\nu)^2}$$

$$= -\frac{A_1(\alpha + 2\pi i\nu)}{(\alpha + 2\pi i\nu)^2} - \frac{A_2}{(\alpha + 2\pi i\nu)^2}$$

$$= -\frac{2\pi A_1 i\nu}{(\alpha + 2\pi i\nu)^2} - \frac{\alpha A_1 + A_2}{(\alpha + 2\pi i\nu)^2},$$

and so (2) can be rewritten as

$$\phi_1(\nu) = \phi(\nu) - \frac{2\pi A_1 i\nu}{(\alpha + 2\pi i\nu)^2} - \frac{\alpha A_1 + A_2}{(\alpha + 2\pi i\nu)^2}. \tag{2'}$$

Note that the first rational expression in (3) is imaginary and the second is real, while in (2′), the first is (for large ν) only approximately imaginary and the second only approximately real. Nevertheless, these approximations seem good: if we keep only the dominant term in the denominator of the first rational expression in (2′), we have $-\frac{2\pi A_1 i\nu}{(2\pi i\nu)^2} = \frac{A_1 i}{2\pi\nu}$, the first term on the right side of (3) (so the ratio of these two expressions approaches one for large ν). Likewise, the second rational expression in (2′) can be approximated with $-\frac{\alpha A_1 + A_2}{(2\pi i\nu)^2} = \frac{\alpha A_1 + A_2}{4\pi^2\nu^2}$, the second term on the right side of (3) (again, the ratio of these two expressions approaches one for large ν). So at first blush it does look like (3) is a legitimate substitution for (2′) for large ν. We may state the case as follows: the first rational terms in each of (2′) and (3) are close to each other, as are the second terms, so we can conclude that their sums are also close to each other.

But this is where the imprecise phrase, "are close to each other" gets us in trouble: the asymptotic "closeness" of the first terms turns out to be *of a different order* than the closeness of the second terms. This causes the order of closeness of their sums to be different from what it is apparently assumed to be, and this causes the ϕ_1 of (3) to converge to 0 more slowly than required. To see this, we examine the terms in (2′) and (3) more closely.

The first term in (3), the imaginary $\frac{A_1 i}{2\pi\nu}$, and the first term in (2′), the approximately imaginary $-\frac{2\pi A_1 i\nu}{(\alpha+2\pi i\nu)^2}$, are both of order $\mathrm{O}\left(\frac{1}{\nu}\right)$ for large ν, and since their ratio approaches one their difference is $\mathrm{o}\left(\frac{1}{\nu}\right)$. Likewise, the real second term in (3) and the approximately real second term in (2′) are both $\mathrm{O}\left(\frac{1}{\nu^2}\right)$ and their difference is $\mathrm{o}\left(\frac{1}{\nu^2}\right)$. But the $\mathrm{o}\left(\frac{1}{\nu}\right)$ difference between the purely imaginary and approximately imaginary terms contains a real $\mathrm{O}\left(\frac{1}{\nu^2}\right)$ part that cannot be ignored when computing the difference between the real and approximately real expressions:

$$-\frac{A_1 i}{2\pi\nu} - \frac{2\pi i A_1\nu}{(\alpha + 2\pi i\nu)^2} = -\frac{A_1 i}{2\pi\nu} - \left(\frac{2\pi i A_1\nu}{(2\pi i\nu)^2}\right)\left(\frac{1}{(1 + \frac{\alpha}{2\pi i\nu})^2}\right)$$

$$= \frac{A_1}{2\pi i\nu} - \frac{A_1}{2\pi i\nu}\left(1 - \frac{2\alpha}{2\pi i\nu} + \frac{3\alpha^2}{(2\pi i\nu)^2} - \cdots\right)$$

$$= -\frac{A_1\alpha}{2\pi^2\nu^2} - \text{terms of order } O\left(\frac{1}{\nu^3}\right)$$

for large ν, where we have used the identity $\frac{1}{(1+x)^2} = 1 - 2x + 3x^2 - 4x^3 + \cdots$ for $|x| < 1$. We must add this $-\frac{A_1\alpha}{2\pi^2\nu^2}$ term to the real term in (3) to give the correct quantity $\frac{-A_1\alpha + A_2}{4\pi^2\nu^2}$ describing the asymptotic behavior of the real part of $\phi_1(\nu) - \phi(\nu)$. If we don't make this correction, the $-\frac{A_1\alpha}{2\pi^2\nu^2}$ term would be absorbed into ϕ_1, making ϕ_1 of order $O\left(\frac{1}{\nu^2}\right)$. After making this correction, solving for A_1, A_2 yields

$$A_1 = -2\pi a_1 \quad \text{and} \quad A_2 = -4\pi^2 a_2 + \alpha A_1, \tag{4}$$

rather than the incorrect value of $-4\pi^2 a_2 - \alpha A_1$ for A_2 given earlier. With the substitutions of (4), the ϕ_1 of (2′) does not become identical to the ϕ_1 of (1), but it does become $O(\frac{1}{\nu^3})$ as $\nu \to \infty$ as required.

How can we keep from making the kind of error seen here? Obviously, the place to start would be to try to make rigorous the leap from (2) to (3)—an effort that should fail, leading to the realization that the leap must be wrong. Given that we already know that the imaginary rational expression in (1) is $O(\frac{1}{\nu})$ and the real expression is $O(\frac{1}{\nu^2})$, we can make explicit the apparent assumptions in **[3]** made in the assertion that (3) follows from (2):

$$\lim_{\nu\to\infty} \frac{\mathrm{im}\big(\phi_1(\nu) - \phi(\nu)\big)}{1/\nu} = \frac{A_1}{2\pi} \tag{5}$$

and

$$\lim_{\nu\to\infty} \frac{\mathrm{re}\big(\phi_1(\nu) - \phi(\nu)\big)}{1/\nu^2} = \frac{A_2 + \alpha A_1}{(2\pi)^2}. \tag{6}$$

These are now mathematically precise statements, so it's an easy exercise to show that (5) is correct and (6) is wrong. Indeed, we now know that (6) should instead read

$$\lim_{\nu\to\infty} \frac{\mathrm{re}\big(\phi_1(\nu) - \phi(\nu)\big)}{1/\nu^2} = \frac{A_2 - \alpha A_1}{(2\pi)^2}. \tag{6′}$$

In fact, even this formulation has potential problems: who's to say that in (2), $\mathrm{im}\big(\phi_1(\nu) - \phi(\nu)\big)$ doesn't have a $O\left(\frac{1}{\nu^2}\right)$ part? (5) above doesn't guarantee against it, and (6′) wouldn't find such a term if it did exist, in which case ϕ_1 would again fail to be $O\left(\frac{1}{\nu^3}\right)$.

The safest way to proceed then is to not worry about decomposing $\phi_1 - \phi$ into real and imaginary parts, but simply compute from (2)

$$\lim_{\nu\to\infty} \frac{\phi_1(\nu) - \phi(\nu)}{1/\nu} = \frac{A_1 i}{2\pi}$$

and then

$$\lim_{\nu\to\infty} \frac{\phi_1(\nu) - \phi(\nu) - \frac{A_1 i}{2\pi\nu}}{1/\nu^2} = \frac{A_2 - \alpha A_1}{(2\pi)^2}.$$

This will insure that we capture the correct coefficients for the $O\left(\frac{1}{\nu}\right)$ and $O\left(\frac{1}{\nu^2}\right)$ terms, regardless of whether they are real or imaginary.

In plain words, we must be careful when approximating a sum of terms whose orders of convergence are different; if we approximate the terms separately, it may be that the approximation of the term of lower order convergence contains a phantom term that coverges at the higher rate, and this term would not appear in our approximation of the higher order term because the approximations are being done separately. Instead, we should approximate the slower converging term first, and then subtract away precisely that approximation before attempting to approximate the higher order term(s).

Acknowledgments. The authors wish to thank the editor and an anonymous referee for very helpful suggestions in the preparation of this paper.

REFERENCES

1. James W. Cooley, Lanczos and the FFT: A discovery before its time, *Proceedings of the Cornelius Lanczos International Centenary Conference*, Philadelphia, 1994, pp. 3–9.
2. Barbara Gellai, Cornelius Lanczos: A biographical essay, *Proceedings of the Cornelius Lanczos International Centenary Conference*, Philadelphia, 1994, pp. xxi–lii.
3. Cornelius Lanczos, *Applied Analysis*, Dover, 1988.
4. Published papers and books of Cornelius Lanczos, *Proceedings of the Cornelius Lanczos International Centenary Conference*, Philadelphia, 1994, pp. lx–lxv.
5. Eric W. Weisstein, "Asymptotic notation," from *MathWorld*–A Wolfram Web Resource, www.mathworld.com/AsymptoticNotation.html

Why Are the Gergonne and Soddy Lines Perpendicular? A Synthetic Approach

ZUMING FENG
Phillips Exeter Academy
Exeter, NH 03833
zfeng@exeter.edu

In any scalene triangle the three points of tangency of the incircle together with the three vertices can be used to define three new points which are, remarkably, always collinear. This line is called the Gergonne Line. Moreover cevians through these tangent points are always concurrent at a common point that, together with the incenter, defines a second line, the Soddy Line. Why should these lines be perpendicular? Beauregard and Suryanarayan [**1**] used Eucildean coordinates to establish these results, whereas Oldknow [**2**] used trilinear coordinates. But such a geometric gem deserves a synthetic geometric proof. We shall use the classical theorems of Ceva and Menelaus to define these lines and then establish their perpendicularity by using a certain inversion.

Let ABC be a scalene triangle. Let Ω and I be its incircle and incenter, respectively. Circle Ω touches sides AB, BC, and CA at C_1, A_1, and B_1, respectively. Lines A_1B_1 and AB meet at C_2, and points A_2 and B_2 are defined analogously. As usual, we set $AB = c$, $BC = a$, $CA = b$, and $s = \frac{a+b+c}{2}$. Then it is well known that $AB_1 = AC_1 = s - a$, $BA_1 = BC_1 = s - b$, and $CA_1 = CB_1 = s - c$.

The Gergonne Line. A lovely result for identifying collinear points is Menelaus' theorem:

Let ABC be a triangle, and let P, Q, R be points on the lines BC, CA, AB, respectively. Then P, Q, R are collinear if and only if

$$\frac{BP}{PC} \cdot \frac{CQ}{QA} \cdot \frac{AR}{RB} = 1.$$

(If the lengths are directed, then the product is -1.)

Applying Menelaus' Theorem to line B_1C_1 with triangle ABC (FIGURE 1) yields

$$1 = \frac{AC_1}{C_1B} \cdot \frac{BA_2}{A_2C} \cdot \frac{CB_1}{B_1A} = \frac{BA_2 \cdot (s-a)(s-c)}{A_2C \cdot (s-b)(s-a)},$$

and so

$$\frac{BA_2}{A_2C} = \frac{s-b}{s-a}.$$

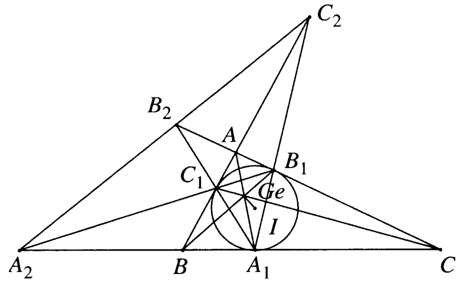


Figure 1

Likewise, we have

$$\frac{CB_2}{B_2A} = \frac{s-c}{s-b} \quad \text{and} \quad \frac{AC_2}{C_2B} = \frac{s-a}{s-c}.$$

Consequently,

$$\frac{BA_2}{A_2C} \cdot \frac{CB_2}{B_2A} \cdot \frac{AC_2}{C_2B} = 1,$$

implying that A_2, B_2, C_2 are collinear, by Menelaus' Theorem. Thus, the line passing through points A_2, B_2, and C_2 is uniquely defined (FIGURE 1). This is the Gergonne Line of triangle ABC.

The Soddy Line. A cevian of a triangle is any segment joining a vertex to a point on the opposite side. We can test cevians for concurrence by using Ceva's Theorem:

> Let AD, BE, CF be three cevians of triangle ABC. Then segments AD, BE, CF are concurrent if and only if
>
> $$\frac{AF}{FB} \cdot \frac{BD}{DC} \cdot \frac{CE}{EA} = 1.$$

Note that

$$\frac{AB_1}{B_1C} \cdot \frac{CA_1}{A_1B} \cdot \frac{BC_1}{C_1A} = \frac{(s-a)(s-c)(s-b)}{(s-c)(s-b)(s-a)} = 1.$$

By Ceva's Theorem, lines AA_1, BB_1, and CC_1 are concurrent (FIGURE 1). The point of concurrency is the Gergonne point Ge of triangle ABC. Because our triangle is scalene, I does not lie on any of the lines AA_1, BB_1, and CC_1. Hence I and Ge are distinct points. Thus there is a unique line passing through I and Ge (FIGURE 1). This line is the Soddy line of triangle ABC.

The Gergonne line and the Soddy line are perpendicular. We apply a certain inversion to show that $A_2B_2 \perp IGe$. Given a point O in the plane and a real number $r > 0$, the inversion through O with radius r maps every point P (distinct from O) to the point P' on the ray OP such that $OP \cdot OP' = r^2$. We also refer to this map as inversion through γ, the circle with center O and radius r. Key properties of inversion that will be used are (for details, please see [**3**]):

(a) Lines passing through O invert to themselves (though the individual points on the line are not all fixed, FIGURE 2, left).

(b) Lines not passing through O invert to circles through O, and vice versa (FIGURE 2, middle).

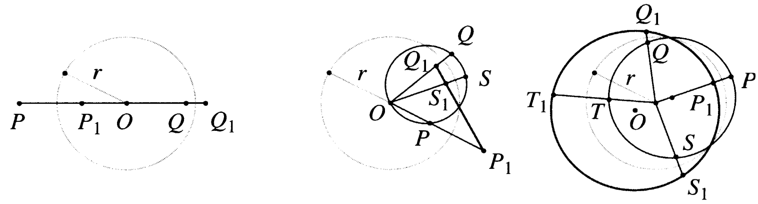


Figure 2

(c) Circles not passing through O invert to circles not passing through O (FIGURE 2, right).

(d) Inversion is a conformal map; that is, inversion preserves the angle between (the tangent lines of) any curves at their intersection points.

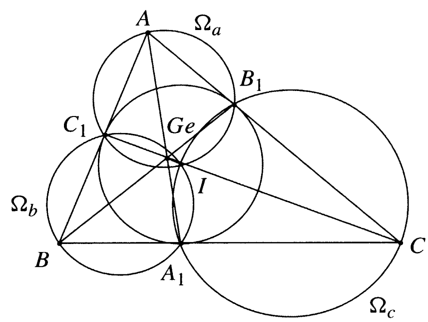


Figure 3

We consider the inversion $\mathbf{I}$ with respect to the incircle Ω (FIGURE 3). Let $\mathbf{I}(P)$ denote the image of element P under the inversion. Then $\mathbf{I}(A_1) = A_1$, $\mathbf{I}(B_1) = B_1$, and $\mathbf{I}(C_1) = C_1$. Because $\angle IB_1A = \angle IC_1A = \frac{\pi}{2}$, points A, B_1, I, C_1 lie on a circle. Let Ω_a denote this circle. We define circles Ω_b and Ω_c (FIGURE 3) analogously. By property (b), $\mathbf{I}(\Omega_a) = B_1C_1$, $\mathbf{I}(\Omega_b) = C_1A_1$, and $\mathbf{I}(\Omega_c) = A_1B_1$.

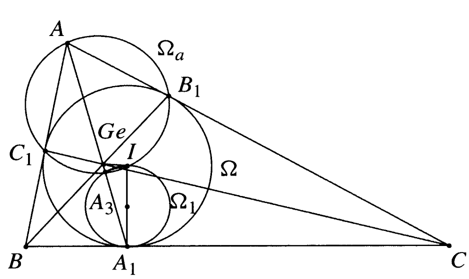


Figure 4

Let Ω_1 be the circle with segment IA_1 as a diameter (FIGURE 4). Then the image of Ω_1 under the inversion is a line, by property (b). Because Ω and Ω_1 are tangent to each other at A_1, by property (d), their images should also be tangent to each other at the image of A_1. It follows that $\mathbf{I}(\Omega_1) = BC$.

Let A_3 be the foot of the perpendicular from I to segment AA_1 (FIGURE 4). Because AI is a diameter of Ω_a and IA_1 is a diameter of Ω_1, A_3 lies on both Ω_a and Ω_1. Thus, $\mathbf{I}(A_3)$ is the intersection of $\mathbf{I}(\Omega_a) = B_1C_1$ and $\mathbf{I}(\Omega_1) = BC$; that is, $\mathbf{I}(A_3) = A_2$ (FIGURE 1).

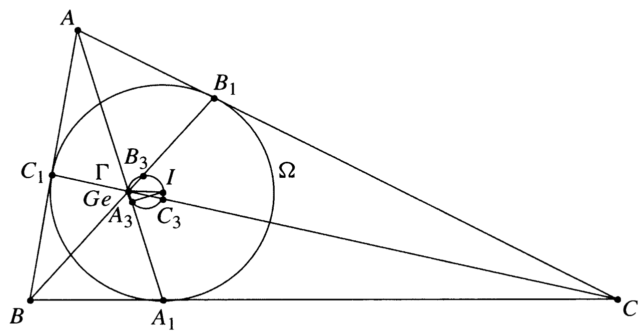


Figure 5

Points B_3 and C_3 are defined analogously and the equations $\mathbf{I}(B_3) = B_2$ and $\mathbf{I}(C_3) = C_2$ follow in similar manner (FIGURE 5). Because $\angle IA_3Ge = \angle IB_3Ge = \angle IC_3Ge = \frac{\pi}{2}$, points A_3, B_3, C_3, I, Ge lie on a circle Γ with IGe as its diameter. By property (b), $\mathbf{I}(\Gamma)$ is a line; that is, points A_2, B_2, C_2 lie on a line. (This is another proof of the existence of the Gergonne line. Furthermore, this shows that $\mathbf{I}(Ge)$ also lies on the Gergonne line.) By property (a), the image of ray IGe is ray IGe. By property (d), to show $A_2B_2 \perp IGe$ (FIGURE 1), it suffices to show that circle Γ and ray IGe are perpendicular at their intersection point Ge . But this is evident, because IGe is a diameter of circle Γ (FIGURE 5).

REFERENCES

1. R. A. Beauregard and E. R. Suryanarayan, Another Look at the Euler-Gergonne-Soddy Triangle, this MAGAZINE **76** (2003) 385–390.
2. A. Oldknow, The Euler-Gergonne-Soddy Triangle of a Triangle, *American Mathematical Monthly* **103** (1996) 319–329.
3. H. S. M. Coxeter and S. L. Greitzer, *Geometry Revisited*, New Mathematical Library, Vol. 19, Mathematical Association of America, 1967.

Golden Matrix Ring Mod p

KUNG-WEI YANG
Western Michigan University (Retired)
Kalamazoo, MI 49008

Playing with the "golden matrix ring" $\mathbb{Z}[A]$ ($A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$), we had fun proving identities involving Fibonacci numbers ($F_0 = 0$, $F_1 = 1, 1, 2, 3, 5, 8, 13, 21, \ldots$) in **[6]**. Here we return to $\mathbb{Z}[A]$ and show that if we reduce $\mathbb{Z}[A]$ modulo p, then we will get a very neat proof of one of the more remarkable properties of Fibonacci numbers: *every prime p divides some (hence infinitely many) Fibonacci numbers.*

LAW OF APPEARANCE OF p [**3**, p. 51]. *Let p be a prime* $\neq 2, \neq 5$.

(1) *If* $p = 5m \pm 1$, *then* $F_{p-1} \equiv 0 \pmod{p}$, $F_p \equiv 1 \pmod{p}$, *and* $F_{p+1} \equiv 1 \pmod{p}$.

(2) *If* $p = 5m \pm 2$, *then* $F_{p-1} \equiv 1 \pmod{p}$, $F_p \equiv -1 \pmod{p}$, *and* $F_{p+1} \equiv 0 \pmod{p}$.

Proof. We shall use the Legendre symbol in our discussion. Specifically $\left(\frac{a}{p}\right)$ is 1 if a is a quadratic residue modulo p, and it is -1 if a is a non-residue. By induction ([**6**]), we see that $A^n = F_{n-1}I + F_nA$. A quick calculation shows that $(2A - I)^2 = 5I$. So $((2A - I)^2)^{(p-1)/2} = 5^{(p-1)/2}I$, and we have $(2A - I)^p = 5^{(p-1)/2}(2A - I)$. Reducing this last equation modulo p and using Fermat's little theorem and Euler's criterion [**2**, Theorem 83, p. 69], we get $2A^p - I \equiv \left(\frac{5}{p}\right)(2A - I) \pmod{p}$. If $p = 5m \pm 1$, then the Law of Quadratic Reciprocity [**2**, Theorem 98, p. 76] assures that we have $\left(\frac{5}{p}\right) = \left(\frac{p}{5}\right) = 1$. So $A^p \equiv A \pmod{p}$, i.e., $F_{p-1}I + F_pA \equiv A \pmod{p}$. If $p = 5m \pm 2$, then, $\left(\frac{5}{p}\right)$ is $= -1$. So $A^p \equiv I - A \pmod{p}$, i.e., $F_{p-1}I + F_pA \equiv I - A \pmod{p}$. All the desired congruences follow. ∎

Since $F_3 = 2$ and $F_5 = 5$, *every* prime divides some Fibonacci numbers. Since d divides n implies F_d divides F_n [**6**, p. 133], every prime divides *infinitely* many Fibonacci numbers.

No additional work is required to extend the above Law to any generalized Fibonacci sequence ($G_{n+2} = G_{n+1} + G_n$) with initial terms G_0, G_1.

PROPOSITION. *Let* (G_n) *be a generalized Fibonacci sequence with initial terms* G_0, G_1. *Let p be a prime* $\neq 2, \neq 5$.

(1) *If* $p = 5m \pm 1$, *then* $G_{p-1} \equiv G_0 \pmod{p}$, $G_p \equiv G_1 \pmod{p}$, $G_{p+1} \equiv G_0 + G_1 \pmod{p}$.

(2) *If* $p = 5m \pm 2$, *then* $G_{p-1} \equiv G_1 - 2G_0 \pmod{p}$, $G_p \equiv G_0 - G_1 \pmod{p}$, *and* $G_{p+1} \equiv -G_0 \pmod{p}$.

Proof. All we need to note ([**6**, p. 132] or by induction) is that $G_{n-1}I + G_nA = A^n((G_1 - G_0)I + G_0A)$. If $p = 5m \pm 1$, then from the proof of the Law of appearance of p we see that

$$\begin{aligned} G_{p-1}I + G_pA &= A^p((G_1 - G_0)I + G_0A) \equiv A((G_1 - G_0)I + G_0A) \\ &= G_0I + G_1A \pmod{p}. \end{aligned}$$

If $p = 5m \pm 2$, then,

$$\begin{aligned} G_{p-1}I + G_pA &= A^p((G_1 - G_0)I + G_0A) \equiv (I - A)((G_1 - G_0)I + G_0A) \\ &= (G_1 - 2G_0)I + (G_0 - G_1)A \pmod{p}. \end{aligned}$$ ∎

If we take the initial terms to be $G_0 = L_0 = 2$, $G_1 = L_1 = 1$, then the resulting sequence is the Lucas sequence and we have the following

COROLLARY. *Let* (L_n) *be the Lucas sequence. Let p be a prime* $\neq 2, \neq 5$.

(1) *If* $p = 5m \pm 1$, *then* $L_{p-1} \equiv 2 \pmod{p}$, $L_p \equiv 1 \pmod{p}$, *and* $L_{p+1} \equiv 3 \pmod{p}$.

(2) *If* $p = 5m \pm 2$, *then* $L_{p-1} \equiv -3 \pmod{p}$, $L_p \equiv 1 \pmod{p}$, *and* $L_{p+1} \equiv -2 \pmod{p}$.

The Law of appearance of p may be found in [**2**, Theorem 180, p. 150], [**3**, IV. 19, p. 47], [**4**, Theorems A, B, C, p. 78], [**5**, Theorem, p. 68]; the corollary is in [**4**, Theorems A′, B′, C′, p. 80]. It is also worth noting that in the Law of appearance of p, if $p = 5m \pm 1$, then $A^{p-1} \equiv I \pmod{p}$ (solving [**1**, Problem 69, p. 33]), and if $p = 5m \pm 2$, then $A^{p+1} \equiv -I \pmod{p}$.

REFERENCES

1. E. B. Dynkin and V. A. Uspenskii, *Problems in the Theory of Numbers*, D. C. Heath, Boston, 1963.
2. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., Oxford University Press, Oxford, UK, 1979.
3. P. Ribenboim, *The Book of Prime Number Records*, Springer-Verlag, New York and Berlin, 1988.
4. S. Vajda, *Fibonacci and Lucas Numbers, and the Golden Section*, Ellis Horwood, Chichester, UK, 1989.
5. N. N. Vorobiev, *Fibonacci Numbers*, Birkhäuser, Boston, 2002.
6. K.-W. Yang, Fibonacci with a golden ring, this MAGAZINE **70** (1997) 131–135.

Geometric Proofs of the Weitzenböck and Hadwiger-Finsler Inequalities

CLAUDI ALSINA
Universitat Politècnica de Catalunya
E-08028 Barcelona, Spain
claudio.alsina@upc.edu

ROGER B. NELSEN
Lewis & Clark College
Portland, OR 97219
nelsen@lclark.edu

Problem 2 on the Third International Mathematical Olympiad in 1961 read [**3**]:

Let a, b, c be the sides of a triangle, and T its area. Prove:

$$a^2 + b^2 + c^2 \geq 4\sqrt{3}\,T. \tag{1}$$

In what case does equality hold?

This inequality is well known in the literature [**1, 4, 8**] as *Weitzenböck's Inequality* (sometimes spelled Weizenbock) from a paper published in 1919 by R. Weitzenböck in *Mathematische Zeitschrift*. Many analytical proofs of the inequality are known—see the above references. The "official" solution to the Olympiad problem appears to have been trigonometric, employing the identities $T = bc \sin A$, $a^2 = b^2 + c^2 - 2bc \cos A$, and $(\sqrt{3} \sin A + \cos A)/2 = \cos(A - 60°)$ [**3**]. Note that A denotes the vertex opposite the side of length a, etc.

There is a very nice geometrical interpretation of this inequality that seems to have been overlooked. If one multiplies both sides of inequality (1) by $\sqrt{3}/4$, then Weitzenböck's inequality can be written as

$$T_a + T_b + T_c \geq 3T, \tag{2}$$

where T_s denotes the area of an equilateral triangle with side length s. The situation is illustrated in FIGURE 1, where (2) states that the sum of the areas T_a, T_b, and T_c of the three shaded equilateral triangles is at least three times the area T of the white triangle.

We now present a purely geometric proof of Weitzenböck's inequality in the form given by (2). Since the proof uses the Fermat point of the original triangle, we first discuss this point. The *Fermat point* of a triangle ABC is the point F in or on the triangle for which the sum $AF + BF + CF$ is a minimum (this is also known as the solution

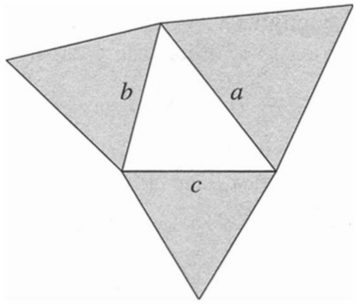


Figure 1

to Steiner's Problem). See FIGURE 2(a). When each of the angles of the triangle is smaller than 120°, the point F is the point of intersection of the lines connecting the vertices A, B, and C to the vertices of equilateral triangles constructed outwardly on the sides of the triangle, as shown in FIGURE 2(b). Furthermore, each of the six angles at F measures 60°. When one of the vertices of triangle ABC measures 120° or more, then that vertex is the Fermat point. For a variety of proofs of these rather remarkable results, see [**2, 6, 7**].

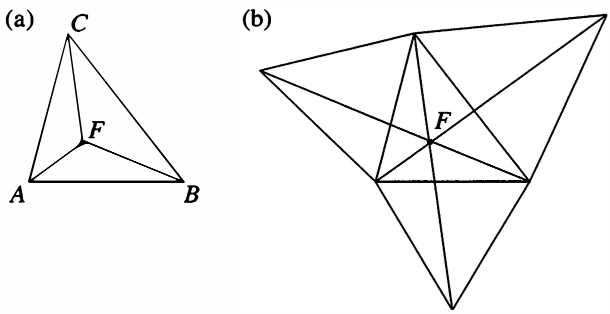


Figure 2

We are now in a position to prove (2). We first consider the case where each angle of the triangle is less than 120°. Let x, y, and z denote the lengths of the line segments joining the Fermat point F to the vertices, as illustrated in FIGURE 3(a), and note that the two acute angles in each triangle with a vertex at F sum to 60°. Hence the equilateral triangle with area T_c is the union of three triangles congruent to the dark

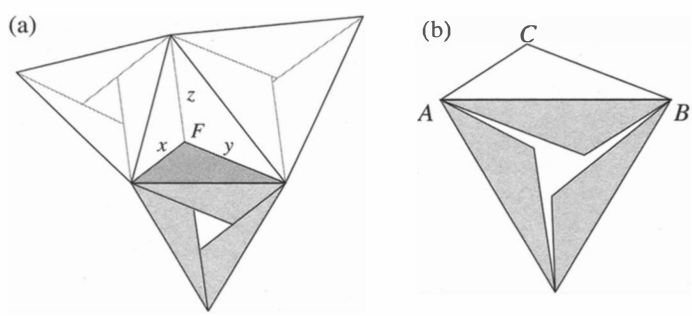


Figure 3

gray shaded triangle with side lengths x, y, and c, and an equilateral triangle with side length $|x - y|$. The same is true of the other triangles sharing the vertex F, and hence

$$T_a + T_b + T_c = 3T + T_{|x-y|} + T_{|y-z|} + T_{|z-x|}, \tag{3}$$

which establishes (2) in this case since $T_{|x-y|}$, $T_{|y-z|}$, and $T_{|z-x|}$ are each nonnegative.

When one angle (say C) measures 120° or more, then, as illustrated in FIGURE 3(b), we have

$$T_a + T_b + T_c \geq T_c \geq 3T \tag{4}$$

which completes the proof.

It follows from (3) that we have equality in (1) and (2) if and only if $x = y = z$, so that the three triangles with a common vertex at F are congruent and hence $a = b = c$, i.e., the original triangle is equilateral.

The relationship in (3) is actually stronger than the Weitzenböck inequality (1), and enables us to now prove another inequality, itself stronger than (1), the *Hadwider-Finsler Inequality* **[1, 8]**: If a, b, and c are the lengths of the sides of a triangle with area T, then

$$a^2 + b^2 + c^2 \geq 4\sqrt{3}\,T + (a-b)^2 + (b-c)^2 + (c-a)^2. \tag{5}$$

In terms of areas of triangles, (5) is equivalent to

$$T_a + T_b + T_c \geq 3T + T_{|a-b|} + T_{|b-c|} + T_{|c-a|}. \tag{6}$$

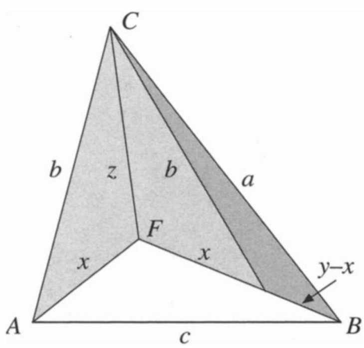


Figure 4

To prove that (3) implies (6) when all three angles measure less than 120°, we need only show that $|x - y| \geq |a - b|$, $|y - z| \geq |b - c|$, and $|z - x| \geq |c - a|$. Without loss of generality, assume that $a \geq b \geq c$. Within triangle ABC reflect the triangle with sides of length b, x, and z about the segment of length z as shown in FIGURE 4, to create two congruent light gray triangles (recall that each of the three angles at F measures 120°). Then in the dark gray triangle we have $b + y - x \geq a$, or equivalently, $y - x \geq a - b$. The other two inequalities are established similarly, and hence from (3) we have

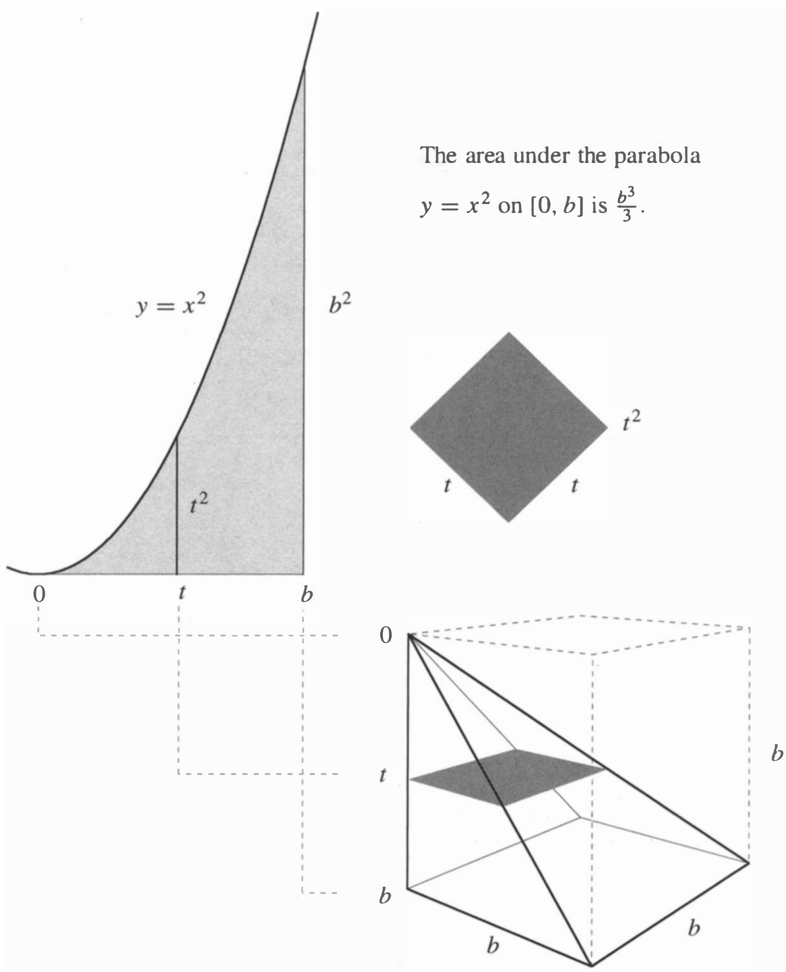
$$\begin{aligned} T_a + T_b + T_c &= 3T + T_{|x-y|} + T_{|y-z|} + T_{|z-x|} \\ &\geq 3T + T_{|a-b|} + T_{|b-c|} + T_{|c-a|}. \end{aligned}$$

In the case where one angle (say C) measures 120° or more, we have $z = 0$, $x = b$, and $y = a$. We refine the inequality in (4) to $T_a + T_b + T_c \geq 3T + T_{|a-b|} + T_a + T_b$ (see FIGURE 3(b)), and note that $a \geq |b - c|$, and $b \geq |c - a|$, from which (6) follows.

REFERENCES

1. O. Bottema, R. Ž. Djordjević, R. R. Janić, D. S. Mitrinović, and P. M. Vasić, *Geometric Inequalities*, Wolters-Noordhoff Publishing, 1969.
2. R. Courant and H. Robbins, *What Is Mathematics?* Oxford University Press, 1941.
3. D. Djukić, V. Janković, I. Matić, and N. Petrović, *The IMO Compendium*, Springer, 2006.
4. A. Engel, *Problem-Solving Strategies*, Springer, 1998.
5. S. L. Greitzer, *International Mathematical Olympiads 1959–1977*, MAA, 1978.
6. R. Honsberger, *Mathematical Gems*, MAA, 1978.
7. I. Niven, *Maxima and Minima Without Calculus*, MAA, 1981.
8. J. M. Steele, *The Cauchy-Schwarz Master Class*, Cambridge University Press and MAA, 2004.

Proof Without Words: Area of a Parabolic Segment



$$\int_0^b x^2 dx = \text{Volume of Pyramid} = \frac{1}{3} \cdot \text{height} \cdot \text{base} = \frac{1}{3} \cdot b \cdot b^2 = \frac{b^3}{3}$$

Carl R. Seaquist
Texas Tech University
Lubbock, TX 79409

PROBLEMS

PROPOSALS

To be considered for publication, solutions should be received by November 1, 2008.

1796. *Proposed by Matthew McMullen, Otterbein College, Westerville, OH.*

A point is selected at random from the region inside of a regular n-gon. What is the probability that the point is closer to the center of the n-gon than it is to the n-gon itself?

1797. *Proposed by Ovidiu Furdui, The University of Toledo, Toledo, OH.*

Let a, b, and c be nonnegative real numbers. Find the value of

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{\sqrt{n^2+kn+a}}{\sqrt{n^2+kn+b}\sqrt{n^2+kn+c}}.$$

1798. *Proposed by H. A. ShahAli, Tehran, Iran.*

Let x, y, and z be positive real numbers with $x+y+z=xyz$. Find the minimum value of

$$\sqrt{1+x^2}+\sqrt{1+y^2}+\sqrt{1+z^2},$$

and find all (x, y, z) for which the minimum occurs.

1799. *Proposed by Luz DeAlba, Drake University, Des Moines, IA.*

Let $s_1, s_2, \ldots, s_n$ be real numbers with $0 < s_1 < s_2 < \cdots < s_n$. For $1 \le i \le j \le n$ define $a_{ij} = a_{ji} = s_j$, and let A be the $n \times n$ matrix $A = [a_{ij}]_{1\le i,\, j\le n}$.

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, written or electronic, should include **on each page** the reader's name, full address, and an e-mail address and/or FAX number.

(a) Calculate $\det(A)$.

(b) Let $A^{-1} = [b_{ij}]_{1\le i,\, j\le n}$. Find the value of $\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}$.

1800. *Proposed by Michel Bataille, Rouen, France.*

Let ABC be a triangle, let E be a fixed point on the interior of side AC, and let F be a fixed point on the interior of side AB. For P on $\overline{EF}$, define

$$\rho(P) = \frac{[PBC]^2}{[PCA][PAB]}.$$

For which P does $\rho(P)$ take on its minimum value? What is this minimal value?

Quickies

Answers to the Quickies are on page 226.

Q981. *Proposed by Jan Mycielski, University of Colorado, Boulder, CO.*

Show that

$$\sum_p \left(\frac{1}{p}\prod_{q<p}\left(1-\frac{1}{q}\right)\right) = 1,$$

where p and q run over the primes.

Q982. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

It is well known that if $f : [0, 1] \to [0, 1]$ is a continuous function, then f must have at least one fixed point. However, the example $f(x) = x^2$, which only has 0 and 1 as fixed points shows that the set of fixed points need not be an interval.

Let $f : [0, 1] \to [0, 1]$ be a function with $|f(x) - f(y)| \le |x - y|$ for all $x, y \in [0, 1]$. Prove that the set of all fixed points of f is either a single point or an interval.

Solutions

An isosceles condition **June 2007**

1771. *Proposed by Mowaffaq Hajja, Yarmouk University, Irbid, Jordan.*

Let P be a point inside of triangle ABC, and let AA', BB', and CC' be the cevians through P. Prove that if $A'B' = A'C'$ and $BC' = CB'$, then triangle ABC is isosceles.

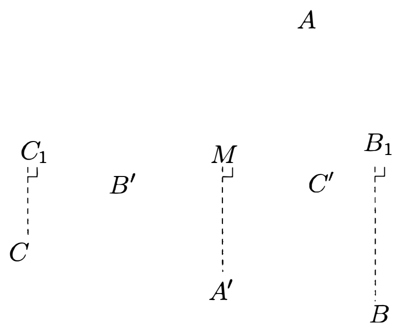
Solution by Elton Bojaxhiu, Albania, and Enkel Hyselaj, Australia.
It suffices to show that $AC' = AB'$, since then $AB = AC' + C'B = AB' + B'C = AC$. Assume that $AC' \ne AB'$. Then without loss of generality we may assume that $AB' > AC'$, and hence that $\angle AC'B' > \angle AB'C'$. Find points B_1, M, C_1 on line $B'C'$ so that each of BB_1, $A'M$, and CC_1 is perpendicular to line $B'C'$. Because $A'B' = A'C'$, it follows that M is the midpoint of segment $B'C'$, so $B'M = MC'$. Thus,

$$C_1M = (B'C)\cos(\angle AB'C') + B'M > (BC')\cos(AC'B') + MC' = B_1M.$$

It then follows from similarity that

$$\frac{BA'}{A'C} = \frac{B_1M}{MC_1} < 1, \qquad \text{and hence that} \qquad \frac{BA'}{A'C}\frac{CB'}{B'A}\frac{AC'}{C'B} = \frac{BA'}{A'C}\frac{AC'}{B'A} < 1.$$

But this last inequality contradicts Ceva's Theorem. Thus it follows that $AC' = AB'$ and $AB = AC$.



Also solved by Byoung Tae Bae (Korea), Herb Bailey, Michel Bataille (France), Robert Calcaterra, Adam Coffman, Chip Curtis, Fejéntaláltuka Szeged Problem Solving Group (Hungary), John Ferdinands, Dmitry Fleishman, Michael Goldenberg and Mark Kaplan, Peter Gressis and Dennis Gressis, Chris Hill, Geoffrey A. Kandall, Elias Lampakis (Greece), Charles McCracken, Mike Meehan, José H. Nieto (Venezuela), Volkhard Schindler (Germany), Raul A. Simon (Chile), Earl A. Smith, Albert Stadler (Switzerland), George Tsapakidis (Greece), Alexey Vorobyov and Michael Vorobyov, and the proposer.

Taylor minima — June 2007

1772. *Proposed by Rick Mabry, Louisiana State University at Shreveport, Shreveport, LA.*

Let n be an even positive integer and let a be a real number. Let $T_n(x; a)$ denote the degree n Taylor polynomial at the point $x = a$ for the exponential function e^x.

(a) Prove that for each real a, the polynomial $T_n(x; a)$ assumes its minimum at a unique point.

(b) Let t_a denote the value of x for which $T_n(x; a)$ assumes its minimum. Prove that the planar set

$$\{(t_a,\ T_n(t_a; a))\ :\ a \in \mathbb{R}\}$$

is itself the graph of an exponential function.

Solution by Michael Janas and Suzanne Dorée, Augsburg College, Minneapolis, MN.
Let n be an even positive integer.

(a) Note that $T_n(x; a) > 0$ whenever $x \geq a$. When $x < a$, Lagrange's form of the remainder gives

$$T_n(x; a) = e^x - \frac{e^c}{(n+1)!}(x-a)^{(n+1)} > e^x > 0,$$

where c is some real number in $[x, a]$. Thus, $T_n(x : a) > 0$ for all x. Next note that $T_n''(x; a) = T_{n-2}(x; a)$ with $n - 2$ even, so $T_n(x; a)$ is convex. To see that $T_n(x; a)$ has a unique minimum, note that $T_n'(x; a) = T_{n-1}(x; a)$ has odd degree and, because it has a positive derivative, is strictly increasing . Thus $T_n'(x; a)$ has a unique zero. This zero is the unique value at which $T_n(x; a)$ assumes its minimum.

(b) Note that $T_n(x; a) = e^a T_n(x - a; 0)$ achieves its minimum where $T_n(x - a; 0)$ does, namely at the unique x value with $x - a = t_0$. Thus $t_a = a + t_0$ and so

$$T_n(t_a; a) = e^a T_n(t_a - a; 0) = e^{t_a - t_0} T_n(t_0; 0) = \left[e^{-t_0} T_n(t_0; 0)\right] e^{t_a},$$

which is exponential in t_a, as desired.

Also solved by Michel Bataille (France), Elton Bojaxhiu (Albania) and Enkel Hysnelaj (Australia), Robert Calcaterra, John Christopher, Fejéntaláltuka Szeged Problem Solving Group (Hungary), John Ferdinands, Peter Gressis and Dennis Gressis, Elias Lampakis (Greece), David Lovit, Kim McInturff, Matthew McMullen, Jerry Metzger and Thomas Richards, José H. Nieto (Venezuela), Nicholas C. Singer, Albert Stadler (Switzerland), Richard Stephens, David Stone and John Hawkins, Marian Tetiva (Romania), Nora Thornber, Alexey Vorobyov and Michael Vorobyov, Michael Vowe (Switzerland), and the proposer.

Maximum of a product — June 2007

1773. *Proposed by H. A. ShahAli, Tehran, Iran.*

Let a, b, c, d be nonnegative real numbers with $a + b = c + d = 1$. Determine the maximum value of

$$(a^2 + c^2)(a^2 + d^2)(b^2 + c^2)(b^2 + d^2),$$

and determine conditions under which the maximum is attained.

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

We can rewrite the given expression as

$$\left((ac + bd)(ad + bc) - (ab - cd)^2\right)^2 + (ab - cd)^2 + \left((a + b)^2(c + d)^2 - 1\right)(ab - cd)^2.$$

Because $a + b = c + d = 1$, the last term is 0, so the expression reduced to

$$\left((ac + bd)(ad + bc) - (ab - cd)^2\right)^2 + (ab - cd)^2. \tag{1}$$

First note that $(ac + bd)(ad + bc) \geq (ab - cd)^2$. Indeed, we may assume, without loss of generality, that $a \geq c \geq d \geq b$. Then $cd - ab > 0$ and the result follows from the immediate inequalities $ac + bd \geq cd - ab$ and $ad + bc \geq cd - ab$.

Also, by the AM–GM inequality,

$$(ac + bd)(ad + bc) \leq \frac{(ac + ad + bc + bd)^2}{4} = \frac{(a + b)^2(c + d)^2}{4} = \frac{1}{4},$$

with equality if and only if $ac + bd = ad + bc$, that is, if and only if $(a - b)(c - d) = 0$. This holds when either $a = b = \frac{1}{2}$ or $c = d = \frac{1}{2}$.

Thus, expression (1) is less than or equal to

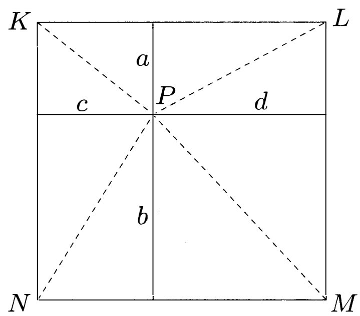
$$\left(\frac{1}{4} - (ab - cd)^2\right)^2 + (ab - cd)^2, \tag{2}$$

with equality if and only if $a = b = \frac{1}{2}$ or $c = d = \frac{1}{2}$. Expression (2) is an increasing function of $(ab - cd)^2$. With the given restrictions, $(ab - cd)^2$ reaches its maximum of $\frac{1}{16}$ at the points $(a, b, c, d) = \left(\frac{1}{2}, \frac{1}{2}, 1, 0\right)$, $\left(\frac{1}{2}, \frac{1}{2}, 0, 1\right)$, $\left(1, 0, \frac{1}{2}, \frac{1}{2}\right)$, and $\left(0, 1, \frac{1}{2}, \frac{1}{2}\right)$, yielding a maximum of $\frac{25}{256}$ for the given expression.

II. *Solution by Michel Bataille, Rouen, France.*
Let $KLMN$ be a square with side length 1, and let P be a point on or inside of the square. Let the distances from P to the sides be $a,\ b,\ c,\ d$ as in the diagram. Then

$$(a^2+c^2)(a^2+d^2)(b^2+c^2)(b^2+d^2) = (PK \cdot PL \cdot PM \cdot PN)^2 .$$

This reduces the problem to the problem of finding the maximum value of $PK \cdot PL \cdot PM \cdot PN$.



Let z_U be the complex affix of point U. We may assume that $z_N = 0$, $z_M = 1$, $z_K = i$, and $z_L = 1+i$, and set $z_P = z$. Then

$$PK \cdot PL \cdot PM \cdot PN = |f(z)|,$$

where $f(z) = z(z-1)(z-i)(z-1-i)$. Because f is entire, the maximum modulus theorem implies that the desired maximum is attained at a boundary point of the square. By symmetry, it is sufficient to investigate the maximum of $|f(z)|$ on the real interval $[0,\ 1]$, that is, the maximum on $[0,\ 1]$ of

$$\phi(x) = |f(x)| = x(1-x)\sqrt{x^2+1}\sqrt{(x-1)^2+1}.$$

By a straightforward calculation

$$\phi'(x) = (1-2x)\psi(x) \qquad \text{where} \qquad \psi(x) = \frac{3x^2-3x+2+2x^2(1-x)^2}{\sqrt{x^2+1}\sqrt{(x-1)^2+1}} > 0$$

on $[0,\ 1]$ It follows that ϕ attains its maximum on $[0,\ 1]$ only at the point $x = \frac{1}{2}$ with $\phi\left(\frac{1}{2}\right) = \frac{5}{16}$. Thus the desired maximum is $\frac{25}{256}$, and is attained if and only if $(a,\ b,\ c,\ d)$ is one of $\left(1,\ 0,\ \frac{1}{2},\ \frac{1}{2}\right)$, $\left(0,\ 1,\ \frac{1}{2},\ \frac{1}{2}\right)$, $\left(\frac{1}{2},\ \frac{1}{2},\ 1,\ 0\right)$, or $\left(\frac{1}{2},\ \frac{1}{2},\ 0,\ 1\right)$.

Also solved by Byoung Tae Bae (Korea), Herb Bailey, Elton Bojaxhiu (Albania) and Enkel Hysnelaj (Australia), Paul Budney, Robert Calcaterra, William Cross, Jim Delany, Gregory Dresden, Fall 2007 Calculus Class at Missouri State University—Maryville, Dmitry Fleischman, Elias Lampakis (Greece), Kee-Wai Lau (China), Jerry Metzger and Thomas Richards, José Nieto (Venezuela), Ángel Plaza and Sergio Falcón (Spain), Volkhard Schindler (Germany), Nicholas C. Singer, Albert Stadler (Switzerland), David Stone and John Hawkins, Marian Tetiva (Romania), Alexey Vorobyov and Michael Vorobyov, Michael Vowe (Switzerland), Stan Wagon, and the proposer. There were six incorrect submission..

Idempotent, hermetian, and equal **June 2007**

1774. *Proposed by Götz Trenkler, Department of Statistics, University of Dortmund, Dortmund, Germany.*

Let P and Q be idempotent, hermetian matrices of the same dimension and rank. Prove that if $PQP = P$, then $P = Q$.

Solution by Nicholas Singer, Annandale, VA.
Let P and Q be complex matrices of dimension n and rank r. Then

$$\mathbb{C}^n = \text{range}(P) \oplus \ker(P),$$

so P is the orthogonal projection onto range(P) and $I - P$ is the orthogonal projection onto $\ker(P) = \text{range}(P)^{\perp}$. Similar statements hold for Q. Because $PQP = P$ and P and Q both have rank r, it follows that $\text{rank}(PQ) = r = \text{rank}(QP)$. Furthermore, because $(PQ)(Px) = Px$ for all x in the range of P, we see that PQ is the identity on range(P). In addition, since $\text{rank}(PQ) = \dim(\text{range}(P))$, we have $PQ \equiv 0$ on $\ker(P)$. Hence, $PQ = P$ because the two agree on range(P) and $\ker(P)$. Also,

$$QP = Q^*P^* = (PQ)^* = P^* = P,$$

so Q is the identity on range(P) and, because $\text{rank}(Q) = r = \dim(\text{range}(P))$, $Q \equiv 0$ on $\ker(P)$. It follows that $P = Q$ because they agree on both range(P) and $\ker(P)$.

Also solved by Michael Andreoli, Michel Bataille (France), Elton Bojaxhiu (Albania) and Enkel Hysnelaj (Australia), Brian Bradie, Paul Budney, Luz M. DeAlba, Michael Goldenberg and Mark Kaplan, Eugene A. Herman, Joel Iiams, Eugen J. Ionascu, Jerry Metzger and Thomas Richards, José Nieto (Venezuela), John H. Smith, Jeffrey Stuart, Nora Thornber, Alexey Vorobyov and Michael Vorobyov, Xiaoshen Wang, and the proposer.

Vertex disjoint paths **June 2007**

1775. *Proposed by Christopher J. Hillar, Texas A&M University, College Station, TX.*

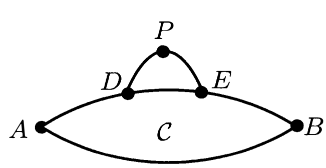
Characterize those graphs G that satisfy the following conditions: between each pair of vertices A and B in G,

(a) there exist two vertex disjoint paths.

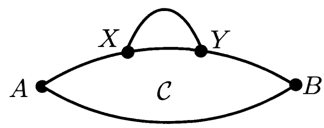
(b) any set of vertex disjoint paths between A and B has at most two elements.

Solution by Missouri State University Problem Solving Group, Missouri State University, Springfield, MO.
We assume that graph means *simple* graph so there are no loops and at most one edge joining any pair of distinct vertices. Conditions (a) and (b) the imply that G must have at least three vertices.

Let A and B be two of these vertices. Condition (a) implies that G contains a cycle $\mathcal{C}$ on which A and B are vertices. If P is a vertex of G not on $\mathcal{C}$, then because there are two vertex disjoint path between A and P and two between B and P, there must be two vertex disjoint paths from P to distinct points of $\mathcal{C}$. Let D and E be the endpoints of these paths on $\mathcal{C}$. (Note that it is not necessary to assume that D and E are distinct from A or B.) Thus we have a situation something like



This clearly yields at least three vertex disjoint paths between D and E, contradicting condition (b). If G has an edge not contained in $\mathcal{C}$, say an edge between the two distinct vertices X and Y on $\mathcal{C}$, then we have at least three vertex disjoint paths between X and Y. It follows that G must be the cycle $\mathcal{C}$.



Also solved by Elton Bojaxhiu (Albania) and Enkel Hysnelaj (Australia), Robert Calcaterra, Eddie Cheng and László Lipták, Dmitry Fleischman, Jerrold W. Grossman, Joel Iiams, S. C. Locke, José Nieto (Venezuela), Paul S. Peck and David I. Kennedy, Alexey Vorobyov and Michael Vorobyov, and the proposer. There was one solution with no name.

Answers

Solutions to the Quickies from page 221.

A981. For prime p let

$$S_p = \{n \,:\, n \text{ is a positive integer, } p|n, \text{ and } q \nmid n \text{ for any prime } q < p\},$$

and note that the density of S_p in the set of positive integers is

$$\frac{1}{p}\prod_{q<p}\left(1-\frac{1}{q}\right).$$

Because the sets S_p are pairwise disjoint and $\cup_p S_p = \{2, 3, 4, \ldots\}$, the result follows.

A982. Let $F = \{x \,:\, x \in [0, 1] \text{ and } f(x) = x\}$. Because f is continuous, it follows that F is compact. Let a be the smallest number in F and b the largest number in F. Clearly $F \subseteq [a, b]$. Now let x_0 be any point in $[a, b]$. Because a is a fixed point for f,

$$f(x_0) - a \leq |f(x_0) - a| = |f(x_0) - f(a)| \leq x_0 - a.$$

Therefore, $f(x_0) \leq x_0$. Similarly,

$$b - f(x_0) \leq |b - f(x_0)| = |f(b) - f(x_0)| \leq b - x_0,$$

from which $f(x_0) \geq x_0$. It follows that $f(x_0) = x_0$, so x_0 is a fixed point f. Therefore $F = [a, b]$.

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.

Julia Robinson and Hilbert's Tenth Problem. 2008. All-region video on NTSC DVD; running time 54 min. Produced and directed by George Csicsery, narrated by Danica McKellar. Available from Zala Films, `www.zalafilms.com` ; $149 for colleges/libraries (includes performance rights), $29.95 for home/personal use. Order form at `http://www.zalafilms.com/downloads/juliaflyer_111707.pdf` . ISBN 978–097245885–6.

I have a special affection for Hilbert's Tenth Problem, having taken a course devoted to it in graduate school just before the final step by Yuri Matiyasevich. And I have a special affection also for Constance Reid, sister of Julia Robinson, whose biographies of mathematicians—including *Julia*, about her sister, a major contributor to solving "Hilbert X"—have enriched us all. Will this video have a comparable effect on students and the public who know nothing about Hilbert X or Reid? I think so! Reid's personal recollections and reflections—delivered in a vivacious and memorable style over family photos—contribute to a vivid, entrancing, and beautiful portrait, enhanced by participation of Julia's theorem-partners and others currently delving yet deeper. The mathematics involved, touched on briefly in the film, is exposited in generous outline in the extra features included on the disc. The producer/director, George Csicsery, did also the award-winning *N is a Number: A Portrait of Paul Erdős* (1993) and the newly-released *Hard Problems: The Road to the World's Toughest Math Contest* (reviewed below). (Thanks to Phil Straffin.)

Flatland: The Movie. 2008. All-region video on NTSC DVD; running time 35 min. English, Spanish, and Italian subtitles. Directed by Jeffrey Travis. Available from Flat World Productions LLC, `www.flatlandthemovie.com`; $120 for educational edition (includes school site license for classroom and school showings, the book, and teacher worksheets), $29.95 for home/personal use. ISBN 978–1–60461–537–1. Abbott, Edwin A., with Thomas Banchoff and the filmmakers of *Flatland*, *Flatland: A Journey of Many Dimensions—The Movie Edition.* Princeton, NJ: Princeton University Press; xv + 168 pp, $15. ISBN 978–0–691–13657–8. *Flatland: The Film.* 2007. All-region video on NTSC DVD; running time 98 min. Directed by Ladd Ehlinger, Jr. Available from Flatland Productions, Inc., `http://www.flatlandthefilm.com/`; $24 (free license for educators to show the film to their students). ASIN: B000NJ60FM.

Not one, but two animated versions of Edwin Abbott's 1884 novel *Flatland*! The extraordinarily well-done *Flatland: The Movie* takes just the right contemporary liberties in updating a 125-year-old classic to a thoroughly enjoyable experience with geometric tones and moral overtones. One of my students called it "the perfect after-school special," which intrigues viewers with the potential of considering dimensionality. Extras on the disc include a PDF of the accompanying book and interviews with the actors portraying the voices (which feature Martin Sheen and his brother). The accompanying book includes the entire text of the original novel and of the screenplay, and an introduction by Thomas Banchoff (Brown University)—but most regrettably not Banchoff's "new introduction" from the Princeton University Press 1991 reprint of the novel, which admirably sets both the book and its author in an informative and inspiring context. I regret that I cannot report to you yet on the longer *Flatland: The Film*, since I have not seen it; but I felt that I should mention it here anyway.

Hard Problems: The Road to the World's Toughest Math Contest. 2008. Video on NTSC DVD; running time 82 min. Produced and directed by George Csicsery. Available from the MAA `http://www.hardproblemsmovie.com/`; $99 for institutions/libraries (includes performance rights), $24.95 for home/personal use (member price $19.95). ISBN 978–0–88385–902–5.

This inspiring and informative video profiles the drive, motivation, and humanity of high school students on the U.S. team in the 2006 International Mathematical Olympiad, together with their steps to reach it and their performance at it. Says a U.S. coach about the team, "They're Americans... they represent our educational system"; but what will strike viewers is that top U.S. contestants are overwhelmingly immigrants and sons of immigrants who go to elite private schools. [Why—as part of "coordination of problems"—are team coaches allowed multiple sessions to pressure judges for higher scores for their teams???] (Thanks to Rama Viswanathan.)

Hayes, Brian, *Group Theory in the Bedroom, and Other Mathematical Diversions*, Hill and Wang, 2008; xi + 269 pp, $25. ISBN 978–8080905219–6.

Get past the innuendo of the title (after all, it's about turning a mattress)—what you have here is a collection of a dozen pieces of the best scientific writing around ("Every essay in this book is a gem"—Martin Gardner). They are a small fraction of author Hayes's "Computing Science" column in *American Scientist* over the past 15 years (so we can look forward happily to future compilations!). Topics include the Strasbourg Cathedral brass clock ("Easter in 11842 falls on April 3"), demand for randomness, statistical physics in economics, the genetic code, frequency of wars, location of the Continental Divide, the Stern-Brocot tree, number partitions, namespace, counting in ternary, equality of numbers, and (not least) the bedroom piece.

Archimedes' Square. Puzzle, Kadon Enterprises, 2003; $29, `http://www.gamepuzzles.com/tiling3.htm#AS`. Rorres, Chris, Stomachion: an introduction, `https://www.cs.drexel.edu/~crorres/Archimedes/Stomachion/intro.html`.

The stomachion puzzle is a tangram-like puzzle consisting of 14 pieces; the Rorres Web page has links to mention of the stomachion by Roman authors and to geometric construction of the pieces from a 12×12 grid. The Archimedes Codex contains a mysterious fragment about the puzzle. The re-emergence of the codex in 1998 engaged Reviel Netz (Stanford University) as it main decoder. Mathematics-lover Joe Marasco sent Netz a replica of the puzzle but with the pieces assembled into a square in a different pattern than in the standard diagram. Marasco and Netz concluded that Archimedes' intention had been combinatorial: to calculate the number of ways that the pieces can be assembled to form a square. What total Archimedes may have arrived at, we don't know; the last part of the fragment is missing. Now you too can own a copy of the puzzle and consider its 536 fundamentally different solutions. (Thanks to Joe Marasco.)

Steen, Lynn Arthur, On being a mathematical citizen: The natural NExT step, `http://www.stolaf.edu/people/steen/Papers/leitzel.html`. Percival, Colin, Security is mathematics, `http://www.daemonology.net/blog/2008-03-21-security-is-mathematics.html`.

The message of Lynn Steen's James R.C. Leitzel Lecture at MathFest last August is, "you need to use your mathematics for more than mathematics itself." With thorough footnotes, he cites the value of mathematical thinking about measures of quantity (how should graduation rates be calculated?), quality (how should learning outcomes be assessed?), and readiness (how can admissions and placement exams be aligned with the mathematical skills expected in college?). He offers disturbing but convincing explanations for lack of progress in mathematics scores on the National Assessment of Educational Progress and for lack of decline in the need for remedial mathematics courses in college. He indicts the psychometric methodology behind "high stakes" state tests and laments educators' belated discovery of unintended consequences ("test only reading and mathematics, only reading and mathematics get taught"). Meanwhile, Percival's short blog asserts that by virtue of their training, all mathematicians have the "security mindset"—a drive to question hidden assumptions and to consider "edge" cases—that is essential for security professionals and fundamental to avoiding unanticipated dangers.

NEWS AND LETTERS

To the Editor:

With regard to "On Infinitely Nested Radicals" by Zimmerman and Ho (this MAGAZINE, February 2008, pp 3–15, hereafter referred to as "ZH"), the following additional references might be noted.

Lemma 1 of ZH appears in the solution to Problem 460, *Amer. Math. Monthly* **23** (1916) 209. ZH's Theorem 1 is essentially Problem E874 of the *Monthly*, whose solution appears in *Amer. Math. Monthly* **57** (1950) 186. An earlier appearance of the form $\sqrt{a - b\sqrt{a - b\sqrt{\cdots}}}$ occurs in T. S. E. Dixon, "Continued Roots," *The Analyst* **5** (1878) 20–21, while $\sqrt{a - \sqrt{a + \sqrt{a - \sqrt{a + \cdots}}}}$ is thoroughly treated in the solution to Problem 1174, this MAGAZINE **57** (1984) 299–300.

The nested radicals arising out of trigonometric identities have a long history of being independently rediscovered. Within the last 100 years we find M. Cipolla, "Intorno ad un radicale continuo," *Periodico di Mat.* series 3, **5** (1908) 179–185; G. Pólya and G. Szegö, *Aufgaben and Lehrsätz aus der Analysis, Volume I*, Springer, Berlin, 1925 (reprinted as *Problems and Theorems in Analysis, Volume I*, translated by D. Aeppli, Springer-Verlag, 1972), Problems 183–185 in Part I; P. J. Myrberg. "Iteratin von Quadratwurzeloperationen," *Ann. Acad. Sci. Fenn. Ser. A. I*, **259** (1958); and the ZH reference, L. D. Servi, "Nested square roots of 2," *Amer. Math. Monthly* **110** (2003) 326–330.

Dixon J. Jones
Fairbanks, AK

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW
Washington, DC 20036