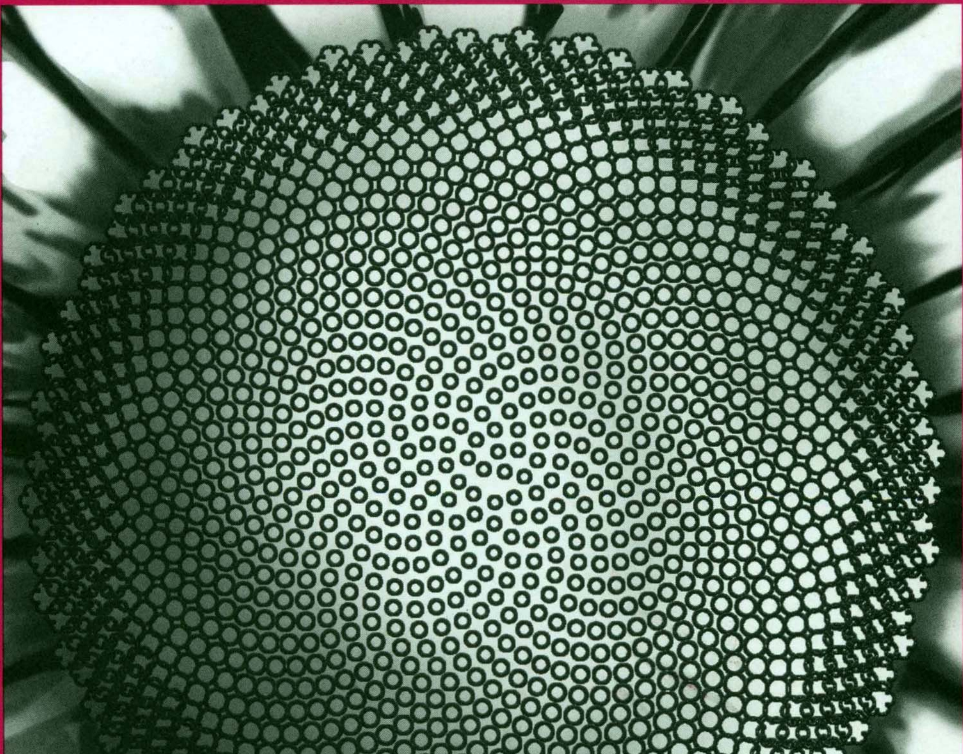
Vol. 75, No. 3, June 2002

MATHEMATICS MAGAZINE



- Golden, $\sqrt{2}$, and π Flowers: A Spiral Story
- Pillow Chess
- Doubly Recursive Multivariate Automatic Differentiation

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

EDITORIAL POLICY

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

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

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

Cover image: *A Clockwork Sunflower*, by Jason Challas and Frank Farris. The disk flowers of the sunflower are drawn using the algorithm from Naylor's article in this issue, with circles of varying size instead of dots; near the outer edge, the circles give way to curves with 6-fold symmetry to mimic the disk flowers that have opened. The petals of the ray flowers around the edge show the human touch. Jason Challas lectures on computer art at Santa Clara University.

AUTHORS

Michael Naylor is a former circus performer, who now juggles patterns in the Mathematics Department at Washington State University. His interests include geometry, mathematics history, elementary mathematics education, and music.

Grant Cairns studied electrical engineering at the University of Queensland, Australia, before doing a doctorate in differential geometry in Montpellier, France, under the direction of Pierre Molino. He benefited from two years as an assistant at the University of Geneva, and a one-year postdoctorate at the University of Waterloo, before coming to La Trobe University, Melbourne, where he is now an associate professor. When he is not being generally enthusiastic about all matters mathematical, his time is devoted to his sons Desmond and Maxwell, and his beautiful wife Romana.

Dan Kalman received his Ph.D. from the University of Wisconsin in 1980. He joined the mathematics faculty at American University in 1993, following an eight-year stint in the aerospace industry (working on automatic differentiation, among other things) and earlier teaching positions in Wisconsin and South Dakota. Kalman has won three MAA writing awards, is on the Editorial Board of *FOCUS*, and served a term as Associate Executive Director of the MAA. His interests include matrix algebra, curriculum development, and interactive computer environments for exploring mathematics, especially using *Mathwright* software.

Vol. 75, No. 3, June 2002

MATHEMATICS MAGAZINE

EDITOR
Frank A. Farris
Santa Clara University

ASSOCIATE EDITORS
Glenn D. Appleby
Santa Clara University
Arthur T. Benjamin
Harvey Mudd College
Paul J. Campbell
Beloit College
Annalisa Crannell
Franklin & Marshall College
David M. James
Howard University
Elgin H. Johnston
Iowa State University
Victor J. Katz
University of District of Columbia
Jennifer J. Quinn
Occidental College
David R. Scott
University of Puget Sound
Sanford L. Segal
University of Rochester
Harry Waldman
MAA, Washington, DC

EDITORIAL ASSISTANT
Martha L. Giannini

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

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

Advertising correspondence should be addressed to Dave Riska (*driska@maa.org*), Advertising Manager, the Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.

Copyright © by the Mathematical Association of America (Incorporated), 2002, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. Permission to make copies of individual articles, in paper or electronic form, including posting on personal and class web pages, for educational and scientific use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear the following copyright notice:

Copyright the Mathematical Association of America 2002. All rights reserved.

Abstracting with credit is permitted. To copy otherwise, or to republish, requires specific permission of the MAA's Director of Publication and possibly a fee.

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

Golden, $\sqrt{2}$, and π Flowers: A Spiral Story

MICHAEL NAYLOR
Western Washington University
Bellingham, WA 98226
mnaylor@cc.wwu.edu

Fibonacci numbers and the golden ratio are ubiquitous in nature. The number $(1+\sqrt{5})/2$ seems an unlikely candidate for what is arguably the most important ratio in the natural world, yet it possesses a subtle power that drives the arrangements of leaves, seeds, and spirals in many plants from vastly different origins. This story is something like these spirals, twisting and turning in one direction and then another, crisscrossing themes and ideas over and over again. We begin with a mathematical model for making these spirals. Many spirals in nature use the golden ratio, but something beautiful happens when we replace that ratio with some other famous irrational numbers. Another twist takes us to rational approximations and continued fractions. Let us follow these spirals into the beautiful world of irrational numbers.

Seed spirals

When a plant such as a sunflower grows, it produces seeds at the center of the flower and these push the other seeds outward. Each seed settles into a location that turns out to have a specific constant angle of rotation relative to the previous seed. It is this rotating seed placement that creates the spiraling patterns in the seed pod [**7**, p. 176].

These spirals can be very neatly simulated as follows: Let's say there are k seeds in the arrangement, and call the most recent seed 1, the previous seed 2, and so on, so that the farthest seed from the center is seed number k. As an approximation, if each seed has an area of 1, then the area of the circular face is k, and the radius is $\sqrt{k/\pi}$. The distance from the center of the flower to each seed, then, should vary proportionally to the square root of its seed number. If we call the angle α, since the angle between any two seeds is constant, the angle of seed k is simply $k\alpha$. We now have a simple way to describe the location of any seed with polar coordinates: $r=\sqrt{k}$, $\theta = k\alpha$.

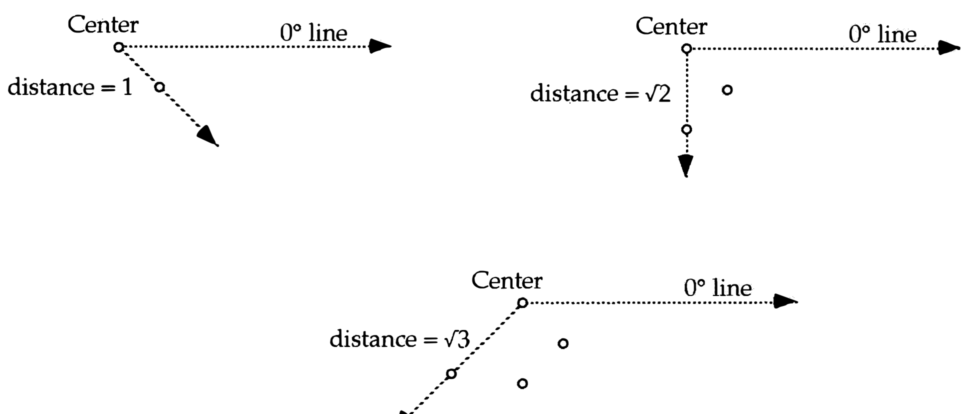


Figure 1 Growing a seed spiral

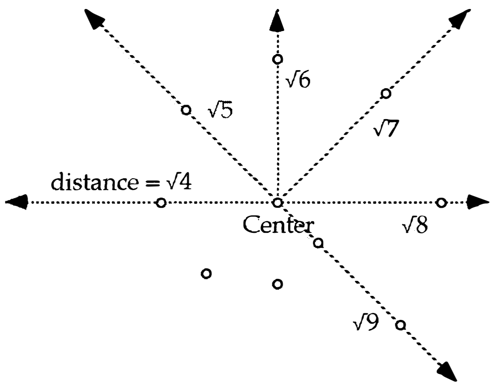


Figure 2 The first 9 seeds

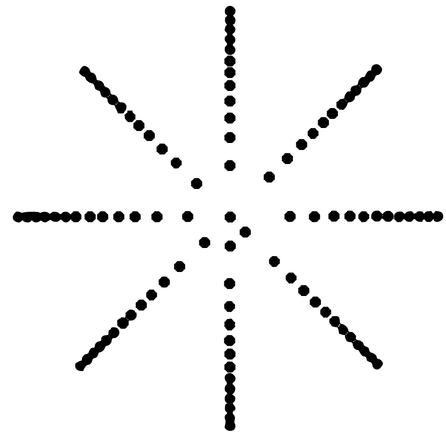


Figure 3 The first 100 seeds

Here's an example of a spiral formed with an angle $\alpha = 45°$, or $1/8$ of a complete rotation. Seed 1 is located at a distance of $\sqrt{1}$ and an angle of 45° (clockwise, in this example). The next seed is located 45° from this seed, or $2 * 45° = 90°$ from the zero line; its distance is $\sqrt{2}$. Seed 3 is located at $3 * 45°$ at a distance of $\sqrt{3}$.

Continuing in this manner, the eighth seed falls on the 0° line, the ninth seed is on the same line as seed 1, and so on (see FIGURE 2). FIGURE 3 shows what the spiral looks like with 100 seeds. It's easy to see the spiral near the center, but the pattern gets lost farther out as the eight radial arms become prominent. Notice how close together the seeds become, and how much space there is between rows of seeds; this is not a very even distribution of seeds. We can get a better distribution of seeds by choosing an angle that keeps the seeds from lining up so readily. If we try an angle of 0.15 revolutions (or 54°), the result is better, especially for the first few dozen seeds, but again we end up with radial arms, 20 this time (see FIGURE 4). Since $0.15 = 3/20$, the 20th seed will be rotated $20 * 3/20$ rotations, or 3 complete rotations to bring it to the 0° line. An angle of 0.48 results in 25 radial arms (see FIGURE 5), since the 25th seed will be positioned at an angle of $25 * 0.48$ rotations, or 12 complete rotations, and the cycle begins anew.

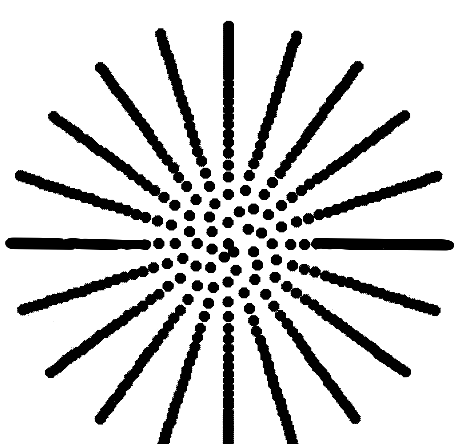


Figure 4 angle = 0.15

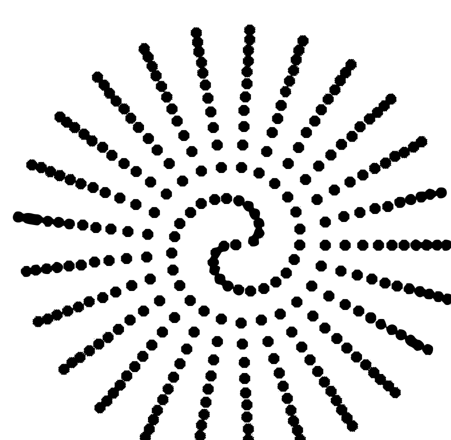


Figure 5 angle = 0.48

Clearly, if the angle is any rational fraction of one revolution, say a/b, seed b will fall on the 0° line, since the angle ab/b is an integral number of complete rotations. Therefore the pattern will repeat, radial arms will be formed, and the distribution be far from ideal. The best choice then, would be an irrational angle—we are then guaranteed that no seed will fall on the same line as any other seed.

Golden flowers

The irrational angle most often observed in plants is the golden ratio, $\phi = (1 + \sqrt{5})/2$, or approximately 1.618. This angle drives the placement of leaves, stalks, and seeds in pine cones, sunflowers, artichokes, celery, hawthorns, lilies, daisies, and many, many other plants [**5**, pp. 155–66; **2**, pp. 90–105; **1**, pp. 81–113]. With this angle of rotation, each seed is rotated approximately 1.618 revolutions from the previous seed—which is the same as 0.618 revolutions, or about 61.8% of a complete turn (approximately 222.5°). For our purposes, only the fractional part of the angle is significant and the whole number portion can be ignored. FIGURE 6 shows 1000 simulated seeds plotted with this angle of rotation, an arrangement we will call a *golden flower*. Notice how well distributed the seeds appear; there is no clumping of seeds and very little wasted space. Even though the pattern grows quite large, the distances between neighboring seeds appear to stay nearly constant. In the natural world, many plants grow their seeds (or stalks or leaves or thorns) simply where there is the most room [**5**, p. 161]. The resulting golden flower is the most even distribution possible [**1**, pp. 84–88; **6**, pp. 96–99]. (For an excellent discussion of the mechanics of the placement of seeds in a growing plant apex and the inevitability of these golden arrangements, see Mitchison [**3**, pp. 270–75].)

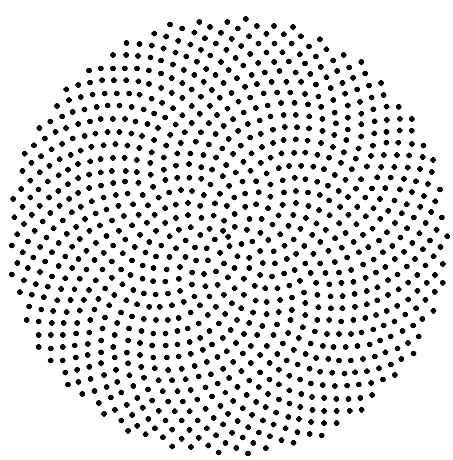


Figure 6 1000 seeds in a golden flower

Notice also the many different spiral arms. Spiral arms seem to fall into certain *families*. In this pattern above, you can see how a group of spirals twist in one direction, only to be taken over by another group of spirals twisting in the opposite direction. The interesting properties of spiral families form the heart of our discussion.

FIGURE 7 shows three families of spirals in the golden spiral. Each set of 300 seeds pictured is identical, but different spirals arms have been drawn on each set. The first set shown consists of 8 spiral arms, the second has 13, and the third 21—all Fibonacci numbers. You may be able to see other spirals not shown in these images, and the size of these groups are Fibonacci numbers as well.

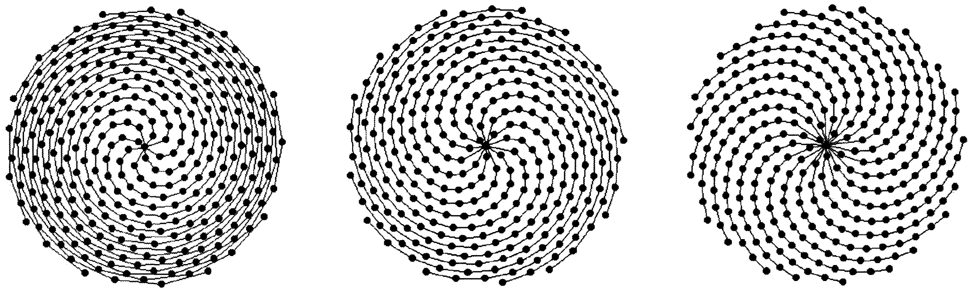


Figure 7 Spiral families 8, 13, and 21

To understand why spirals on a golden flower appear in groups whose size are Fibonacci numbers, it helps to consider placement of individually numbered seeds. In FIGURE 8, the first 144 seeds are numbered and the Fibonacci numbers are enclosed in rectangles. The baseline at 0° has also been added.

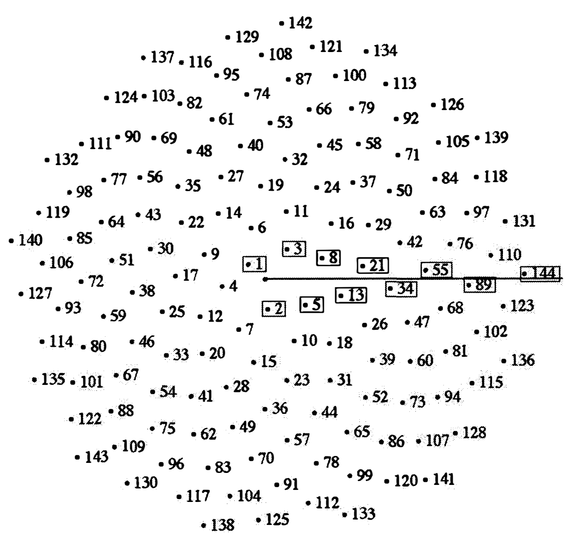


Figure 8 Fibonacci seeds

The Fibonacci numbered seeds converge on the 0° line, alternating above and below, just as the ratios of pairs of consecutive Fibonacci numbers converge to ϕ, alternately greater and less than ϕ. A seed that is numbered with a Fibonacci number will fall close to the zero degree line, since its angle (a Fibonacci number times ϕ) is approximately an integer. For example, since $55/34$ is approximately ϕ, seed 34 will be located at an angle of about $34 * (55/34)$, or very nearly 55 complete rotations (actually ~ 55.013 rotations, a slight over-rotation). The larger the Fibonacci numbers involved, the closer their ratio is to ϕ and therefore the closer the seeds lie to the zero degree line.

It is for this reason that seeds in each spiral arm in a golden flower differ by multiples of a Fibonacci number. For example, seed 34 is slightly over-rotated past the 0° line, seed 68 is rotated by the same angle from seed 34, as are seeds 102, 136, 170, and every other multiple of 34. These seeds form one spiral arm in family 34. Another arm in this family is 1, 35, 69, 103..., and another is 2, 36, 70, 104, ..., etc. Members of an arm in family 34 are seeds with numbers $34m + n$, where m and n are nonnegative integers and n is constant for that arm. Trace any spiral arm in the golden flower and you will find that its seed numbers are in arithmetic progression, since all share a common difference—a Fibonacci number.

π flowers

Why should the golden ratio be the preferred irrational number in nature? Shouldn't any irrational number work just as well? Let's take a look at a simulated seed pod generated with an angle of π rotations, or $\pi * 360°$. This angle is ~ 3.14159 revolutions, which is the same as ~ 0.14159 revolutions, or $\sim 50.97°$. FIGURE 9 shows the first 500 seeds—not a very even distribution at all! Seven spiral arms dominate the pattern with no new spirals apparent. With 10,000 seeds (FIGURE 10), a new set of spirals become visible, 113 arms in this family with so little curvature that the next set of spirals doesn't show until about a million seeds have been grown.

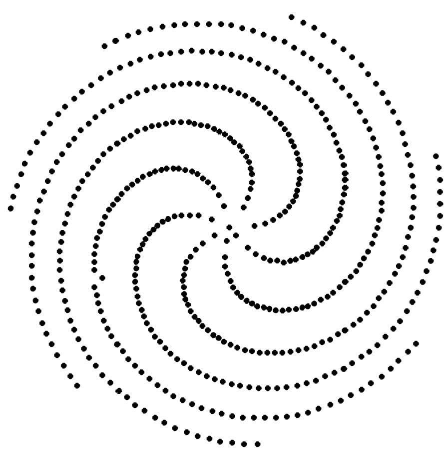


Figure 9 500 seeds, angle = π

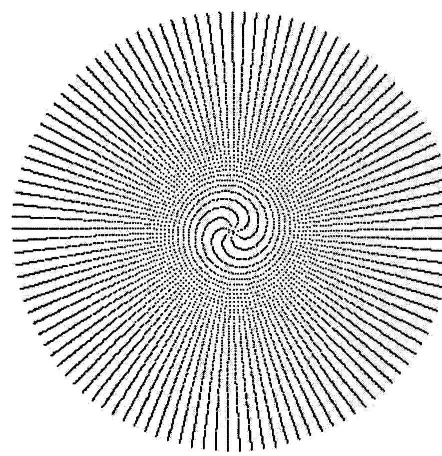


Figure 10 10,000 seeds, angle = π

Why should there be 7 spiral arms so prominently displayed in the center, and 113 arms in the next set of spirals? Perhaps you recognize these numbers as denominators in well-known rational approximations of π. An excellent approximation of π is $22/7$. The decimal expansion of $1/7$ is $0.142857\ldots$ and the angle of rotation in a π flower is $0.14159\ldots$—a close match! Another great approximation of π is $355/113$, accurate to 6 decimal places, and for this reason the next set of spirals has 113 arms.

The gap between these spiral families (7 and 113) in a π flower is huge compared to that of a golden flower. No other sets of spirals are apparent between family 7 and family 113—does this mean that there are no better rational approximations of π with denominators between 7 and 113? Plotting and numbering the seeds in a π flower suggests an answer. In FIGURE 11, seed 7 in the first spiral arm falls near the 0° line as expected, as does seed 113. Since seed 113 is part of the second spiral arm to cross this line, there is no seed less than 113 that lies closer to the 0° line than seed 7, and thus there is no better rational approximation of π than $22/7$ with a denominator less than 113. The approximation $355/113$ is so accurate that the spirals in family 113 have very little curvature and their members dominate the 0° line for generations. The nearest seed in the third arm to cross is seed 226—part of the same arm as seed 113. In fact, we need to check tens of thousands of seeds before we find one that falls closer to the 0° line than any multiple of 113—a topic we will visit again later.

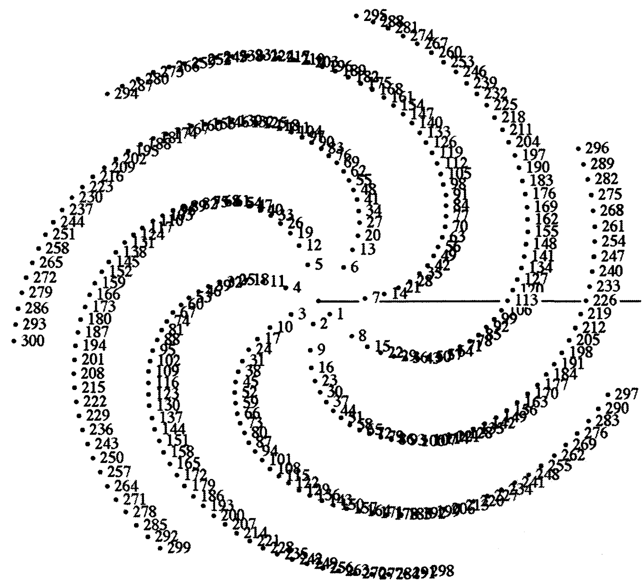


Figure 11 A numbered π flower

$\sqrt{2}$ flowers

An angle of rotation of $\sqrt{2}$ produces a very even distribution of seeds, rivaling that of the golden ratio. Five hundred seeds are shown in FIGURE 12; families of spirals are again readily apparent in this arrangement. A study of these $\sqrt{2}$ spirals is worthwhile, as their structure illuminates many properties of algebraic numbers and seed spirals in general.

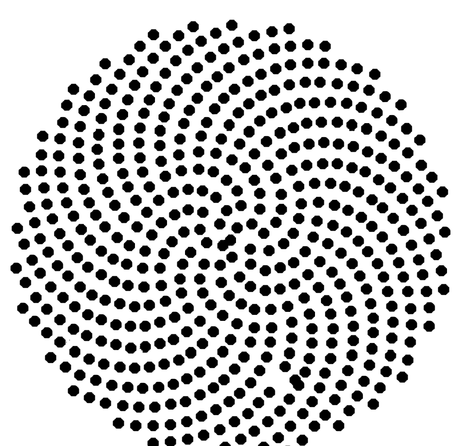


Figure 12 A root-two spiral

FIGURE 13 shows the results of a brute-force analysis—the first 12 families of spirals in a $\sqrt{2}$ spiral. Family 1 is made by connecting the seeds in order, family 2 is made by connecting seeds whose numbers differ by 2, family 3 by connecting seeds whose numbers differ by 3, etc. Study these spiral families for a moment. Notice that

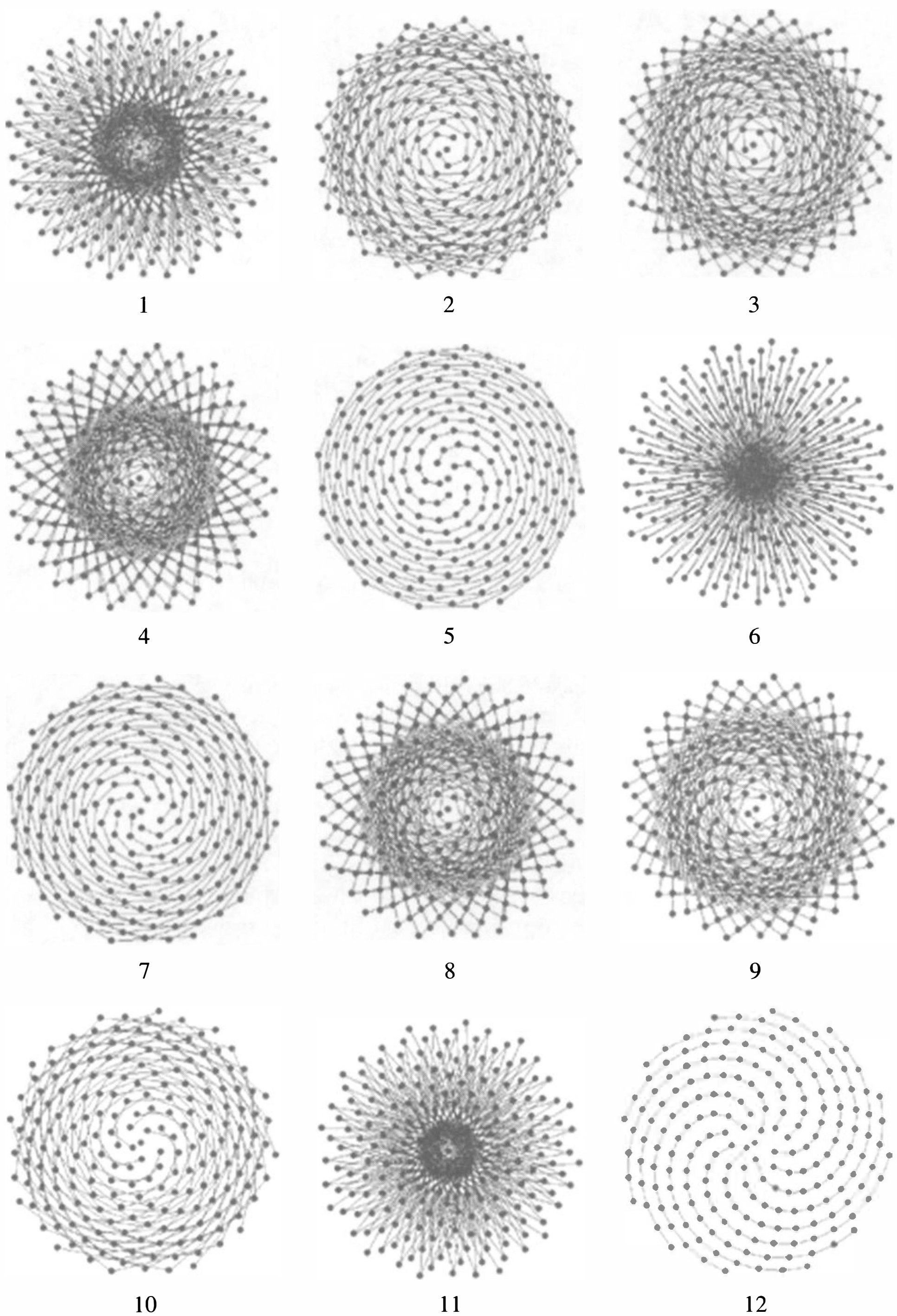


Figure 13 Spiral families 1–12 of a $\sqrt{2}$ spiral

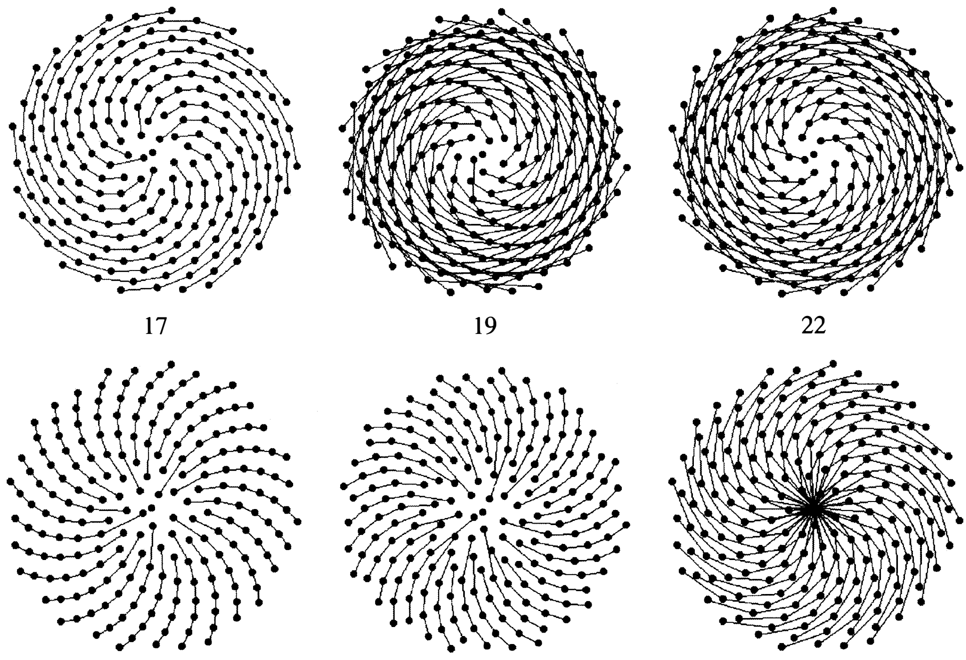


Figure 14 Selected families of the $\sqrt{2}$ spiral

some of the families produce very clean spiral arms, while others do not appear to be spirals at all, crossing themselves in a star-like or even scribble-like pattern.

Families 2 and 3 start well but quickly die, that is, they cross themselves after a small number of iterations. Families 5 and 7 are smooth. Family 10 looks like smooth spirals, but on closer examination it is seen that it crosses itself immediately—since 10 is a multiple of 5, this is the same as family 5 but with alternate seeds on the arms connected. Family 12 has the best looking spirals among the first 12.

The numbers of the families that produce nice spirals look suspicious: 2, 3, 5, 7, 12... could there be a Fibonacci-like relationship between spiral families in a $\sqrt{2}$ spiral as well? More spiral families are shown in FIGURE 14, but the next spiral family after 12 is not 19 as we might expect by adding 7 and 12, but rather 17, and the next family better than 17 is 29. The sequence is in fact: 1, 2, 3, 5, 7, 12, 17, 29, 41, 70, 99, ... Before reading further, can you find the pattern in this sequence and extend it? The numbers in this sequence are the numbers in the Columns of Pythagoras. The Columns of Pythagoras are a pair of columns of integers. The top entry in each column is 1. Given a row with numbers A and B in that order, the next row is generated by summing A and B and writing this number, C, in the first column underneath A, then summing A and C and writing it in the second column underneath B. This process generates all of the spiral families of the $\sqrt{2}$ flower (see FIGURE 15). Further, the ratio of the numbers in each row converges to $\sqrt{2}$: $1/1 = 1$, $3/2 = 1.5$, $7/5 = 1.4$, $17/12 = 1.41666\ldots$, $41/29 = 1.41379\ldots$, etc.

Continued fractions

Let us follow one more twist on this spiraling journey. The golden ratio may be written as the following continued fraction:

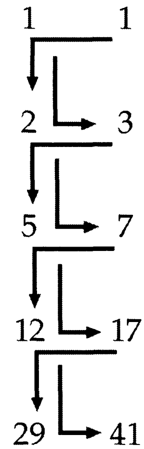


Figure 15 The Columns of Pythagoras

$$\frac{1+\sqrt{5}}{2} = 1 + \cfrac{1}{1 + \cfrac{1}{1+\cfrac{1}{1+\cdots}}}$$

This result is easily verified by setting the continued fraction equal to some variable, say x, and then recognizing that x is repeated in the denominator of the fractional part, that is, $x = 1 + 1/x$. This expresses the continued fraction perfectly as one root of an easily evaluated quadratic.

Partial evaluations of this continued fraction, called *convergents*, result in ratios of Fibonacci numbers, that is,

$$1+\frac{1}{1}=\frac{3}{2}, \quad 1+\cfrac{1}{1+\frac{1}{1}}=\frac{5}{3}, \quad 1+\cfrac{1}{1+\cfrac{1}{1+\frac{1}{1}}}=\frac{8}{5}, \quad 1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\frac{1}{1}}}}=\frac{13}{8}.$$

The reader may enjoy checking that the following continued fraction gives an expression for $\sqrt{2}$:

$$\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{2+\cfrac{1}{2+\cdots}}}.$$

Partial evaluations of this continued fraction yield the following ratios:

$$1+\frac{1}{2}=\frac{3}{2}1+\cfrac{1}{2+\frac{1}{2}}=\frac{7}{5}, \quad 1+\cfrac{1}{2+\cfrac{1}{2+\frac{1}{2}}}=\frac{17}{12}, \quad 1+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\frac{1}{2}}}}=\frac{41}{29}.$$

These are the same numbers in the Columns of Pythagoras and the same ratios found in the $\sqrt{2}$ flower!

Let us examine the continued fraction for the other irrational number we have used to build flowers, π. The continued fraction begins $3 + 1/7\ldots$ and the values of numbers leading the expressions under the denominators at each level, starting with the 7, are: $7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, \ldots$.

Looking at the partial evaluations yields rational approximations to π that reflect the number of spiral arms in the π flower:

$$\pi = 3 + \frac{1}{7} = \frac{22}{7},$$

$$\pi = 3 + \frac{1}{7 + \frac{1}{15}} = 3 + \frac{1}{\frac{106}{15}} = 3 + \frac{15}{106} = \frac{333}{106},$$

$$\pi = 3 + \frac{1}{7 + \frac{1}{15 + \frac{1}{1}}} = 3 + \frac{1}{7 + \frac{1}{16}} 3 + \frac{1}{\frac{113}{16}} = 3 + \frac{16}{113} = \frac{355}{113}.$$

Remember the 113 arms in the π spiral? It would take a lot of seeds to begin to find the next series of spirals—the next partial evaluation explains why:

$$\pi = 3 + \frac{1}{7 + \frac{1}{15 + \frac{1}{1 + \frac{1}{292}}}} = 3 + \frac{1}{7 + \frac{1}{15 + \frac{292}{293}}} = 3 + \frac{1}{7 + \frac{293}{4687}} = 3 + \frac{4687}{33102} = \frac{103993}{33102}.$$

The next family of spirals past family 113 is family 33102. We would need 33,102 seeds just to get one seed in each spiral arm! If we plot about a million seeds we may be able to starting seeing these spirals; however, there would be nearly 92 spirals packed into each degree arc of the circular face. An illustration 10 cm in diameter would have over 1000 spiral arms in each cm of the circumference—the illustration would appear to be nothing other than a black circle! (Note that 333/106 is also an approximation of π. However, due to the closeness of a better approximation, 355/113, the set of 106 spiral arms is immediately obscured by the set of 113 spiral arms.)

The continued fraction expansion for the golden ratio uses the smallest possible numbers in the expansion, namely 1s. Therefore, it converges to a rational number the least quickly. In this sense, the golden ratio is the most irrational number and therefore gives the best possible distribution [**6**, pp. 96–99].

More... Given that seed spirals are easily plotted using polar coordinates, you may wish to create your own irrational flowers using mathematics software. Software (Mac OS) used to create many of the images seen here is also available to download for free from the author's web page at http://www.wwu.edu/~mnaylor.

REFERENCES

1. T. Cook, *The Curves of Life*, 1914. Reprint 1979, Dover Press.
2. S. Colman, *Nature's Harmonic Unity*, Benjamin Blom, Inc., New York, 1912.
3. G. J. Mitchison, Phyllotaxis and the Fibonacci series, *Science* **196** (April 1977), 270–275.
4. K. Niklas, *Plant Biomechanics*, University of Chicago Press, Chicago, 1992.
5. P. Stevens, *Patterns in Nature*, Little, Brown, and Co., New York, 1974.
6. I. Stewart, Daisy, Daisy, give me your answer, do, *Scientific American* (January 1995), 96–99.
7. D. Thompson, *On Growth and Form*, abridged version, Cambridge University Press, London, 1971.

Pillow Chess

GRANT CAIRNS
La Trobe University
Melbourne, Australia 3086

In cylindrical chess, one plays on a regular chessboard with the pieces in their standard positions, but imagines that the left and right edges of the board are identified. So, for example, when a rook travels horizontally off the right-hand edge, it reappears in the same row on the left-hand side of the board. It is easy to make a cylindrical chessboard out of paper, by taping together the left- and right-hand edges. However, to avoid the pieces falling off, the simplest way to play cylindrical chess is to use the regular (flat) chessboard and just remember the edge identification. Like 3-dimensional chess, cylindrical chess is well over 100 years old. According to Pritchard [**59**] it dates to the early eighteenth century. *Byzantine* or *Round* chess is played on a cylindrical 4×16 board, and is possibly a thousand years old [**59, 77, 53**].

In chess on a torus, one identifies the left and right edges and also the top and bottom edges, as in FIGURE 1a. These identifications really do produce the torus; the left and right edge identification gives a cylinder, and the top and bottom edge identification amounts to connecting up the ends of the cylinder (see Barr [**3**] or Stillwell [**70**] for drawings of this construction). Once again, for playing toroidal chess, it is to best to use the regular chessboard and just remember the edge identifications. The origins of toroidal chess are not clear, but it goes back at least as far as Pólya's 1918 paper [**57**].

One can also play chess on the Klein bottle; the situation is similar to the torus, but the horizontal edge identifications involve a reflection (FIGURE 1b): as a rook travels up from h8, it reappears in a1. For games on the Klein bottle, there are even stronger reasons to use the regular chessboard; the Klein bottle can't be constructed in 3-dimensional Euclidean space without self-intersections [**3**].

The projective plane is another surface that is commonly represented by a square with edge identifications, as in FIGURE 1c. The projective plane is the space of straight lines through the origin in 3-dimensional Euclidean space. It can be obtained from the sphere by identifying antipodal points. Like the Klein bottle, it is not orientable and can't be realized in 3-dimensional Euclidean space without self-intersections. See Barr [**3**], Stillwell [**70**] or Prasolov [**58**] for more information about representing surfaces. An alternative *projective chess* is obtained from the traditional board by adding squares at infinity [**56**, Chapter 6.16]. This is also called *projected chess* [**59**].

The "fairy chess" games of FIGURE 1 are easy to understand; conceptually they are really only a small variation on traditional chess since after all, the usual board can be

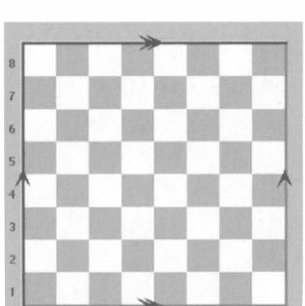


Figure 1a Torus

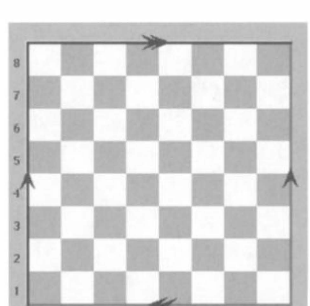


Figure 1b Klein bottle

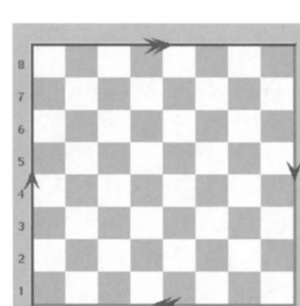


Figure 1c Projective plane

thought of as being composed of 64 squares, with rules that tell you how the squares are connected up. The game is quite different in practice; for starters, one wouldn't want to commence one of these chess games with the pieces in their traditional positions, since the kings would begin on adjacent squares! Pritchard [**59**] gives various possible starting positions on the torus. Also, on the Klein bottle and the projective plane, one doesn't have the usual black-white pattern; the black square a1 is immediately above the black square h8. We could fix this if we were willing to use a 9×8 board, for example.

The pillow board

In this paper, we introduce *pillow chess*, a way to play chess on a surface equivalent to the sphere where one can play with the standard pieces in their traditional positions. The edge identifications are shown in FIGURE 2; the left and right edges are identified with each other, while the top and bottom edges are identified with themselves. So, for example, when a rook travels up the "b" column and off the top edge of b8, it reappears in g8. Notice that unlike the torus chessboard, the board depicted in FIGURE 2 isn't homogeneous; like the projective plane chessboard, it has *corners*. The midpoint of the bottom edge is such a corner: when you attempt to circle that point, the total angle traversed is only π. This also occurs at the midpoint of the top edge, at the extreme top left point (which is identified with extreme top right point), and the extreme bottom left point (which is identified with extreme bottom right point).

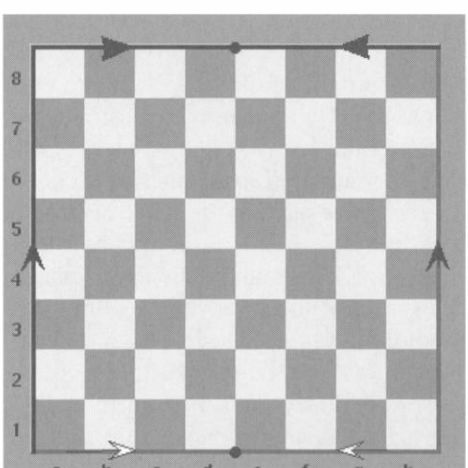


Figure 2 The pillow chessboard

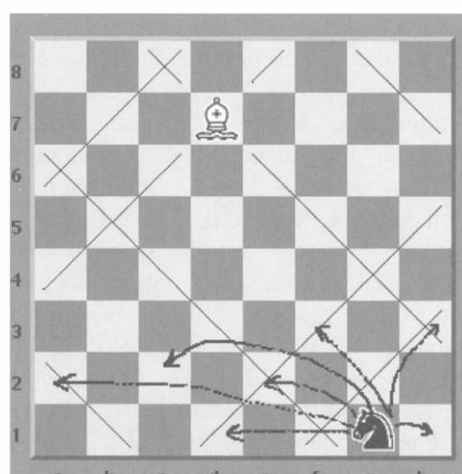


Figure 3 Example of moves

A little work with the Euler characteristic of surfaces such as these will show that the existence of 4 corners is virtually forced. Indeed, for any choice of edge identifications, the squares of the chessboard form a cell decomposition of the resulting surface S. For each integer $i \geq 1$, let n_i denote the number of zero-cells (that is, points) that are adjacent to i two-cells (squares). Then there are $\sum_{i>1} n_i$ zero-cells, $\sum_i \frac{i}{2} n_i$ one-cells, and $\sum_i \frac{i}{4} n_i$ two-cells and the Euler characteristic [**48**] is

$$\chi(S) = \sum_i \frac{i}{4} n_i - \sum_i \frac{i}{2} n_i + \sum_i n_i = \sum_i \left(1 - \frac{i}{4}\right) n_i .$$

If we agree only to form boards with $n_i = 0$ for i other than 2 and 4, we get $\chi(S) = \frac{n_2}{2}$ and because $\chi(S) \leq 2$, this gives $n_2 = 4, 2$ or 0. The case $n_2 = 2$ is the projective plane. In the case $n_2 = 4$, we get the sphere with 4 corners, which we call a *pillow*, for obvious reasons. One can form *hyperbolic boards* by allowing $n_i \neq 0$ for $i \neq 2, 4$, but then one loses the concepts of horizontal and vertical, as long as one persists with squares that are really square. Boards paved by triangles or hexagons have been studied [**59, 33, 7**], but we will stay with squares.

As in the other fairy chessboards described above, the pieces in pillow chess have a greater command of the board than in traditional chess; see FIGURE 3. For example, the knight at g1 can reach h1 by moving two squares to the right and one square down. The bishops travel along parallel lines: if the bishop moves from d7 through e8, it reappears in c8, travelling along the parallel line through b7. (The bishops don't *bounce* off the top edge as in the *reflecting queens* introduced by Klarner [**37**]; for more on reflecting queens, see articles by Guy [**27**, Section C18] or Gardner [**24**, Chapter 15], and the Klarner's *Maths Review* [**38**] of Huff's paper.) To get used to pillow chess, we suggest considering the pillow chess problem posed in FIGURE 4; the solution is given at the end of the paper. When considering this problem, it may assist the reader to imagine what the game looks like "across the edges" of the board; such an expanded view is shown in FIGURE 5.

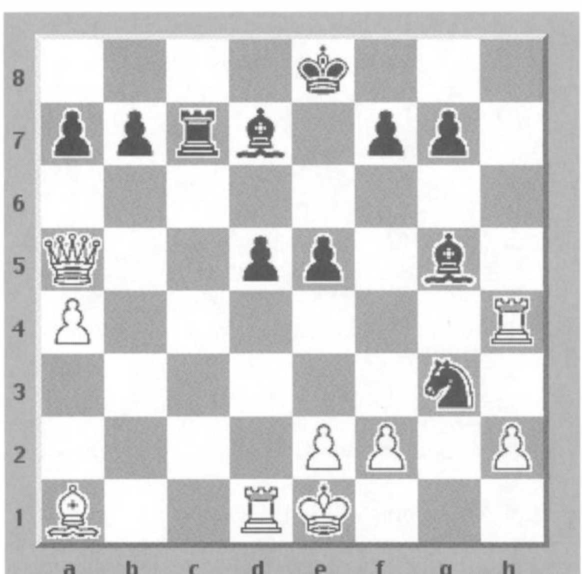


Figure 4 White to play. Black to mate in one!

The aim of this paper is to revisit, on the pillow board, two classical chess problems of a mathematical nature: the *knight's tour* problem, and the *n-queens* problem. As is the tradition, we do not restrict ourselves to 8×8 boards. However, in order to retain the usual black-white pattern, we will restrict our attention to $n \times 2m$ boards.

First, we need to understand better the nature of the pillow board. The pillow is an example of an *orbifold*. In the same way that a manifold is a space that looks locally like Euclidean space, an orbifold is a space that looks locally like the quotient of Euclidean space by a finite group action. In fact, the pillow was one of the first examples treated in Thurston's book [**72**], where the term orbifold first appeared. Orbifolds were

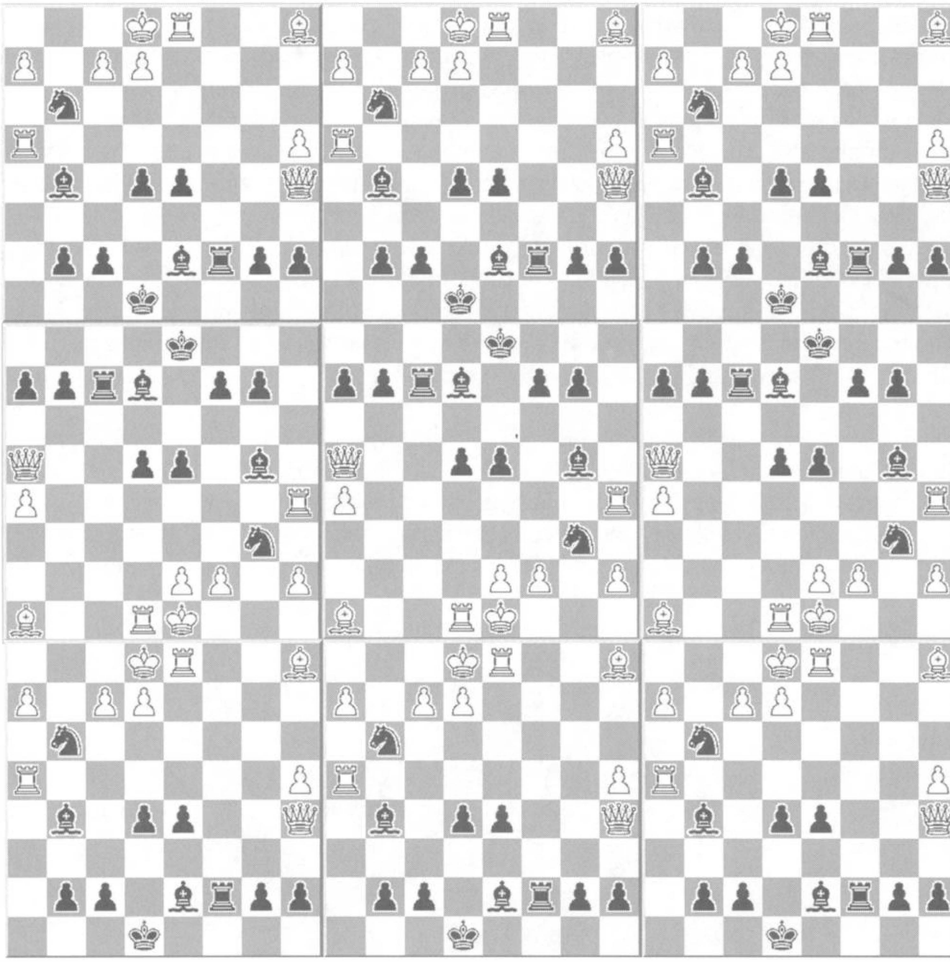


Figure 5 Expanded view of the problem

previously defined by Satake [**63**], who called them V-manifolds. For a nice introductory treatment of orbifolds in a somewhat more restrictive sense, see Stillwell [**70**, Chapter 8].

The simplest example of an orbifold is the quotient of the plane by a finite group of rotations about the origin; the quotient is a cone, which is smooth everywhere except at a single conical singularity (the image of the origin). The pillow is an orbifold which is smooth everywhere except at its 4 conical singularities. But the concept of an orbifold is quite subtle. Indeed, just as the pillow is homeomorphic to the sphere S^2, every 2-dimensional orbifold is homeomorphic to a 2-dimensional manifold; what distinguishes the orbifold from the manifold is not just what it looks like topologically, but what it is *geometrically*. In our case, the pillow has the flat Euclidean geometry, and each of the 4 conical singularities has angle π.

The pillow is a special kind of orbifold. It is not just locally a quotient by a finite group action; we will show that it is *globally* the quotient of the torus by an action of the 2-element group Z_2. First recall that the torus can be regarded as the quotient space of the plane $\mathbb{R}^2$ under the action of the group $\mathbb{Z}^2$ by translation: $(i, j) : (x, y) \to (x + i, y + j)$[**54**]. In this way, one can play chess on the torus by playing on the infinite plane and identifying appropriate squares; if the plane is tiled by unit-square chessboards, centered on the vertices of the integer lattice $\mathbb{Z}^2$, then one identifies the square at $(x, y) \in \mathbb{R}^2$ with the squares at $(x+i, y+j)$, for all $(i, j) \in \mathbb{Z}^2$. The board centered at (0, 0) is a fundamental domain [**54**] for the $\mathbb{Z}^2$ action. Every chess piece on the torus board is thus represented by infinitely many replicas, each in the corresponding position on its tile/board. (FIGURE 6 depicts the torus with a queen and a knight.)

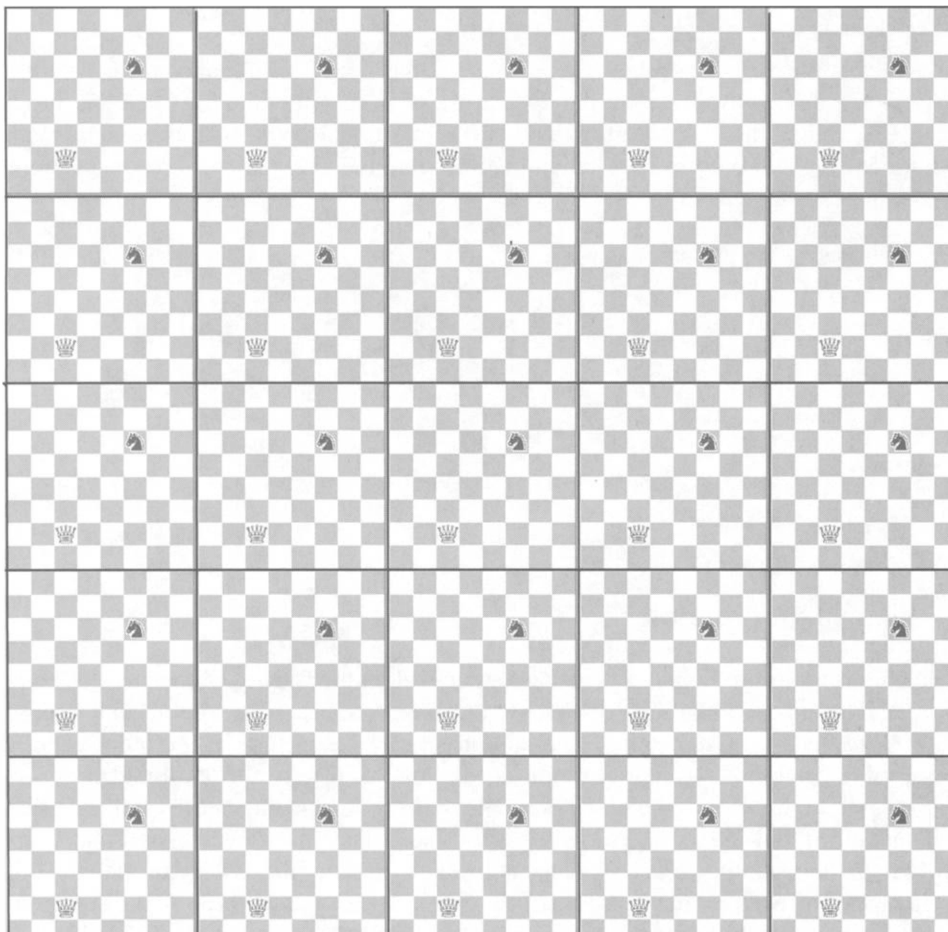


Figure 6 The torus as a quotient of the plane

In a manner similar to the representation of the torus as a quotient of the plane, the pillow can be realized as a quotient of the torus. One represents the torus $\mathbb{T}^2$ as the unit square, centered at the origin in $\mathbb{R}^2$ with opposite edges identified as in FIGURE 1a, and one considers the rotation $\sigma : \mathbb{T}^2 \to \mathbb{T}^2$ defined by $\sigma(z) = -z$. The set $\{\mathrm{id}, \sigma\}$ defines a 2-element group isomorphic to $\mathbb{Z}_2$. The upper half of the unit square is a fundamental domain for this action of this group, and you can convince yourself that the identifications on the boundary of the fundamental domain are precisely those of the pillow board. Thus, one can play chess on the pillow by playing on the torus centered at the origin and identifying diametrically opposite squares.

The realization of the pillow board as a quotient of the torus is not just a curiosity; it is a useful tool. To fix ideas, let us say that a *board* is a set of squares, with marked edges, together with a rule for connecting edges in pairs. We say that two boards are *equivalent* if there is a bijection between their sets of squares that respects the edge connections. Now notice that, whereas in the above description of the pillow board as a quotient of the torus we took the upper half of the torus as the fundamental domain, we could equally as well have taken the left-hand half as the fundamental domain. This simple observation immediately gives:

LEMMA. *The $n \times 2m$ pillow chessboard and the $m \times 2n$ pillow chessboard are equivalent.*

For example, the 8×8 pillow is equivalent to the 4×16 pillow, and one can readily experience this equivalence by playing on the 4×16 pillow; see FIGURE 7. From this perspective, the 4×16 pillow is just the traditional Byzantine chessboard with additional top and bottom edge identifications, and the starting chess piece positions are also the same.

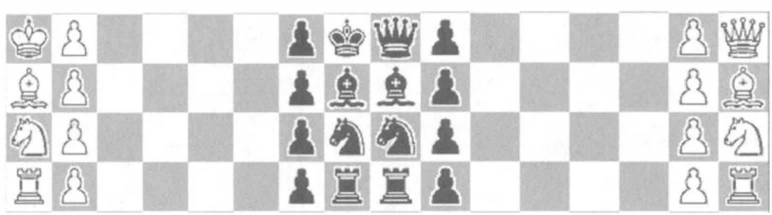


Figure 7 The 4×16 pillow chessboard

The knight's tour problem

The knight's tour problem is this: can a knight visit all the squares of the board exactly once and return to its starting position? In this paper we use the word tour in the sense of a *closed* or *re-entrant* tour. Some authors use the word tour in the more general sense of an *open* tour—one not requiring the knight to return to its initial position, and some authors refer to closed tours as *circuits*. As documented by Murray [**52**] (see also [**53**] and [**68**]), the knight's tour problem dates back over a thousand years to Indian chess and has numerous appearances throughout the history of the game of chess (but not back as far as 200 BCE, as some have claimed [**73, 74**]). The problem was investigated by mathematicians such as Euler [**2**] and Vandermonde [**71**]; in modern terminology, the tour is an example of a Hamiltonian circuit [**80**], [**5**, Chapter 11]. There is a vast literature on the problem. As Kraitchik remarked [**42**] (in a paper first published in 1941), "Many generalizations of the knight's problem have been proposed. Many alterations of the size and shape of the board have already been considered."

One recurring topic is the (open) knight's tour on the half board [**52, 53**]. There are cute proofs of the impossibility of a closed tour on $4 \times n$ boards; Honsberger gives Pósa's proof [**32**, p. 145], and Gardner [**22**] gives a proof that he attributes to Golomb. The connection between knight tours on the 4×4 board (minus one square) and the 15 puzzle was investigated in this MAGAZINE in 1993 [**34**]. The number of knight's tours on the 8×8 board [**46, 49**] and the $n \times n$ board [**43**] have both been studied. Heuristics for generating tours are given by Shufelt and Berliner [**67**]. Boards of other shapes have also been studied [**42**; **47**, Vol. 4]. Tours on the cylinder, Möbius band and Klein bottle appear in the "fanciful" account of Stewart [**69**, Chapter 7] and are studied in Watkins [**78**]. See Eggleton and Eid [**17**] for tours on infinite boards. For tours on boards with hexagonal tiles, see [**44**].

Constructing a knight's tour on the (traditional) $k \times l$ rectangular chessboard is classical. The following theorem is stated without proof in Kraitchik's *Mathematical Recreations* [**42**, Chapter 11]; independent proofs were given by Cull and De Curtins [**16**] (except for the $k = 3$ case) and Schwenk [**65**].

RECTANGULAR BOARD THEOREM. *For $k \leq l$, there is a knight's tour on the $k \times l$ rectangular chessboard unless one or more of the following three conditions hold:*
(a) k and l are both odd,
(b) $k = 1, 2$ or 4,
(c) $k = 3$ and $l = 4, 6$ or 8.

Another proof of the rectangular board theorem for square boards is available [**15**]; this also treats open tours that commence and terminate on specified squares. Open tours on $n \times m$ boards with $\min(n, m) \geq 5$ are examined in [**16**]. Results for open tours on $4 \times m$ boards are known [**41, 73**], and the $3 \times n$ case is treated elsewhere [**41, 73, 75**]. (We alert the reader to the typographical error in the statement of [**75**, Theorem 2], in which *tour* should read *circuit* in the language of that paper.)
On the torus, Watkins and Hoenigman proved [**79**]:

TORAL BOARD THEOREM. *For all k, l, there is a knight's tour on the $k \times l$ torus chessboard.*

We now turn to the pillow board. Note that since the $n \times 2m$ pillow board is the quotient of the $2n \times 2m$ toral board, it follows immediately from the previous theorem that there is a closed knight's path on the $n \times 2m$ pillow board that visits each square exactly *twice*. In fact, one has:

THEOREM 1. *For all n, m, there is a knight's tour on the $n \times 2m$ pillow chessboard.*

Proof. By our Lemma, we may assume that $n \leq m$. By the rectangular board theorem, it suffices to consider the cases $n \leq 4$ and in the case $n = 3$ we need only consider $m = 3$ and $m = 4$. Moreover, by the Lemma, the $4 \times 2m$ board is equivalent to the $m \times 8$ board, so by the rectangular board theorem again, for $n = 4$ and $n \leq m$, we need only consider $m = 4$. So this leaves us with 5 cases to consider: (1) $n = 1$, (2) $n = 2$, (3) $n = m = 3$, (4) $n = 3$, $m = 4$, and (5) $n = m = 4$. Tours on the 3×6 and 3×8 boards are shown in FIGURE 8; the 3×6 example was taken from Watkins and Hoenigman's study of tours on the torus [**79**] (this particular tour uses only the side edge identifications). A tour on the 4×8 board is shown in FIGURE 9. It remains to deal with cases (1) and (2).

1	16	13	10	7	4
8	5	2	17	14	11
15	12	9	6	3	18

1	16	7	22	13	4	19	10
20	11	2	17	8	23	14	5
15	6	21	12	3	18	9	24

Figure 8 Tours on the 3×6 and 3×8 pillow boards

Each move of a knight can be represented by a pair (i, j), with $i, j \in \{\pm 1, \pm 2\}$, where for example $(1, 2)$ means move 1 square to the right and 2 squares up (in the obvious sense). For convenience, we adopt a notation similar to that of Monsky [**50**]; we

1	20	11	26	17	8	23	14
24	15	2	21	12	27	18	9
19	10	25	16	7	22	13	32
30	29	6	3	28	31	4	5

Figure 9 Tour on the 4×8 pillow board

let $A = (1,2)$, $B = (1,-2)$, $C = (-1,2)$, $D = (-1,-2)$, $E = (2,1)$, $F = (2,-1)$, $G = (-2,1)$, $H = (-2,-1)$, and we interpret the word $A(BC)^2$, for example, to mean the sequence of moves A, B, C, B, C. In Case (1), a simple tour is given by the sequence A^{2m}; FIGURE 10 shows the example of the 1×8 board. In Case (2), we start at the top left corner and make the following sequence of moves:

for m even, $m = 2k$, $B(CB)^{m-1}H(EF)^{k-1}EB(GH)^{k-1}G^2$.
for m odd, $m = 2k-1$, $B(CB)^{m-1}H(EF)^{k-1}A(HG)^{k-1}G$.

FIGURE 11 shows the example of the 2×6 and 2×8 boards. ■

1	2	3	4	5	6	7	8

Figure 10 Tour on the 1×8 pillow board

1	12	3	8	5	10
6	7	4	11	2	9

1	16	3	10	5	14	7	12
8	9	6	15	4	11	2	13

Figure 11 Tours on the 2×6 and 2×8 pillow boards

The n-queens problem

The n-queens problem is this: can one place n queens on an $n \times n$ board such that no pair is attacking each other? Such queens are said to be *nonattacking* or *invulnerable.*

This problem dates back to the middle of the 19th century [**10**]. The traditional board has solutions for $n \neq 2, 3$, while the $n \times n$ torus board has solutions when n is not divisible by 2 or 3. The problem was solved for the torus by Pólya [**57**] and the traditional board by Ahrens [**1**] and essentially the same underlying idea has been used by several authors. The key construction for n relatively prime to 6 appeared in Theorem II at the end of the 4th section of Lucas' book [**47**, Vol. 1], which was published originally in 1883. For more recent presentations and formulations of this idea, see [**13, 25, 40**], for the torus, and [**4, 20, 31, 60, 81**], for the traditional board.

(Incidentally, the n-queen problem for the torus is obviously equivalent to the n-queen problem for the cylinder.)

The basic construction can be described as follows: start on any given square on the torus, and make repeated knight moves $(2, 1)$, placing a queen on each square that one visits. Regarding the initial square as the origin in $\mathbb{Z}_n \times \mathbb{Z}_n$, this places queens on the squares $k(2, 1)$, where the entries are reduced modulo n. More generally, solutions of the form $k(d, 1)$, for some d, are said to be *linear* or *regular*. Most authors are content in giving a particular linear solution. A systematic approach is adopted by Erbas, Tanik, and Aliyazicioglu **[19]** (see also **[18, 45, 66]**). Nonlinear solutions are given by Bruen and Dixon **[9]** and by Chandra **[11]**. When n is not divisible by 2 or 3, the $d = 2$ case gives a solution on the torus, and one can easily deduce a solution on the $n \times n$ traditional board for $n \not\equiv 2, 3 \pmod 6$. The cases $n \equiv 2, 3 \pmod 6$ can be treated by combining two linear parts; a succinct description is given in a *Monthly* article on the subject **[61]**. These solutions were already considered by Lucas, who called them *semi-regular* [**47**, Vol. 1, p. 64].

Many other aspects of the n-queen problem have been studied. In addition to the question of existence, one may also investigate the number of solutions. This latter problem is still largely open **[62, 61]**. When no solutions exist, one is interested in the maximum number of nonattacking queens that can be placed on the board. This has been completely solved for the torus **[50, 28, 12, 29, 51]**. Like the knight's tour problem, the n-queens problem has a graph theoretic interpretation: the existence of a maximal stable set [**5**, Chapter 4; **6**, Chapter 13]. The problem has been examined in higher dimensions **[39, 55]**, and on infinite boards **[14]**. The n-queens problem has found many echoes in computer science; according to Vardi **[76]**, it is "a canonical homework assignment in introductory programming classes." See the references at the end of Section 6.4 of **[30]**. The queens problem has also been turned into a game **[64]**.

Variations on the n-queens problem have been studied by changing the possible moves of the queen: a *semiqueen* is a queen that can't move on the negative diagonals [**21**; **36**; **47**, Vol. 1, p. 84, Theorem I; **76**]; an *amazon* is a piece that can move like a queen and a knight **[26, 11]**. Incidentally, the term *amazon* is traditional [**8**, Section 30; **33**]; the term *superqueen* is sometimes used [**23**, Chapter 16; **26**], but *superqueen* is also used to mean a queen on a torus **[11, 50, 35]**. Amazons are called *nite-queens* by Chandra **[11]**.

Now we turn to the pillow board. By our Lemma, the $2m \times 2m$ pillow board is equivalent to the $m \times 4m$ pillow board, which has m rows. So one can place at most m nonattacking queens on the $2m \times 2m$ pillow board. Thus a solution requires the placing of half as many queens as on the traditional board, but each queen commands more squares, in fact, twice as many squares, for most positions. Investigations on small boards show that solutions are more abundant than on the traditional board. According to my calculations, for $m = 2, 3, 4, 5, 6$, the $2m \times 2m$ pillow chessboard has 24, 64, 768, 6464, 54656 solutions respectively. In general, one has:

THEOREM 2. *For all m, the* $2m \times 2m$ *pillow chessboard admits m nonattacking queens.*

Before proving this result, let us look more closely at the notion of nonattacking queens. We give the squares coordinate labels (x, y), $x, y \in \{1, \dots, 2m\}$ in the obvious way, starting at $(1, 1)$ in the bottom left-hand corner. Notice that queens on the traditional board at distinct positions (x_1, y_1), (x_2, y_2) are nonattacking if and only if the following 4 conditions hold: $x_1 \neq x_2$, $y_1 \neq y_2$, $(y_1 - x_1) \neq (y_2 - x_2)$, $(y_1 + x_1) \neq (y_2 + x_2)$. On the torus one has similar conditions, with equality replaced by equivalence modulo $2m$. On the pillow, the conditions are: $x_1 \neq x_2$, $y_1 \neq y_2$,

$\mu(y_1 - x_1) \neq \mu(y_2 - x_2)$, $\mu(y_1 + x_1 - 1) \neq \mu(y_2 + x_2 - 1)$, where the function μ is defined as follows. Let $[i]$ denote the remainder of i modulo $2m$. Then

$$\mu(i) = \begin{cases} [i] & \text{if } [i] \leq m, \\ [-i] & \text{otherwise.} \end{cases}$$

Proof. There is an obvious solution for $m = 1$. For $m > 1$, let $A = \{(i, 2i) : i < \frac{2}{3}m\} \cup \{(i, 2i+1) : \frac{2}{3}m \leq i < m\}$ and $B = \{(i, 2i+1) : i < m\}$ and place a queen at each point in the following set:

$$C = \begin{cases} A \cup \{(m+1, m+1)\} & \text{if } m \equiv 0 \text{ (Mod 6)}, \\ B \cup \{(m, m+1)\} & \text{if } m \equiv 1 \text{ or } 5 \text{(Mod 6)}, \\ B \cup \{(m+1, m)\} & \text{if } m \equiv 2 \text{ or } 4 \text{(Mod 6)}, \\ A \cup \{(m, m)\} & \text{if } m \equiv 3 \text{(Mod 6)}. \end{cases}$$

These positions are just a variation on the classical idea; we begin near the bottom left corner and proceed up the board using knight moves, placing a queen on each square visited. When m is divisible by 3, there is a little skip near the major diagonal. Finally there is a solitary rogue queen near the middle of the board. The positions for $m = 6, \ldots, 9$ are shown in FIGURES 12 to 15.

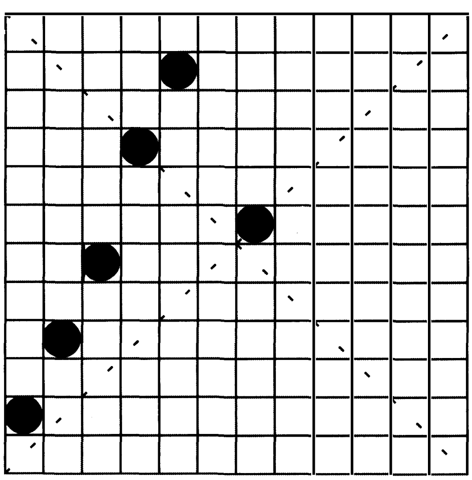


Figure 12

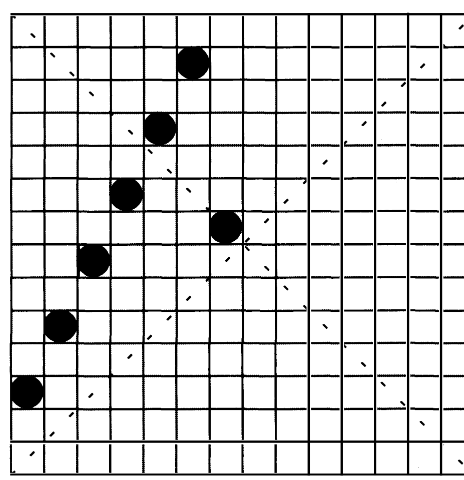


Figure 13

The proof that the m positions in C are nonattacking is perhaps best done by considering three separate cases: (a) m divisible by 6, (b) m divisible by 3 but not by 6, and (c) m not divisible by 3. We briefly describe case (a). From the construction of C, no two positions are in the same column. So we must show that the positions are also in distinct rows, distinct negative diagonals, and distinct positive diagonals. This gives three things to check:

Row Check

(1) $2i = m + 1$ is impossible for m even.

(2) $2i + 1 = m + 1 \Longrightarrow i = m/2 \Longrightarrow \frac{2}{3}m \not\leq i$.

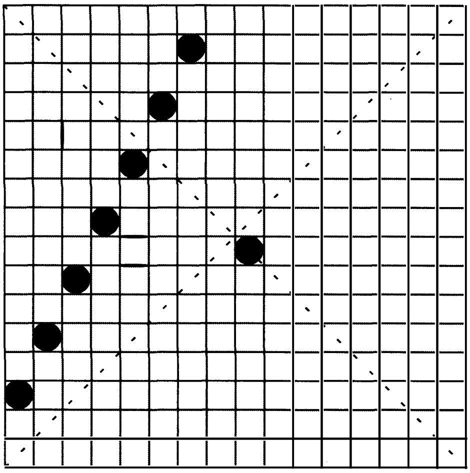


Figure 14

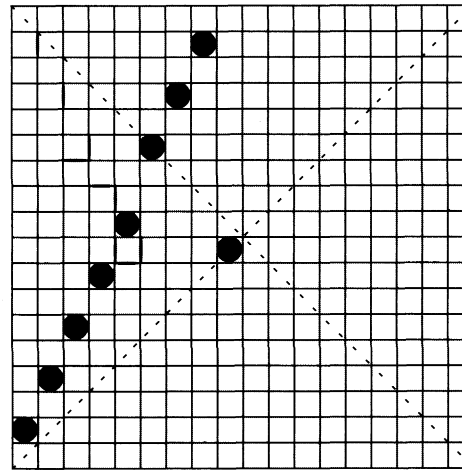


Figure 15

Negative Diagonals Check

(1) For $i, j < \frac{2}{3}m$, $\mu(i) = \mu(j) \Longrightarrow i = j$.

(2) For $i < \frac{2}{3}m$, $\frac{2}{3}m \leq j < m$, $\mu(i) = \mu(j+1) \Longrightarrow i = j+1$, which is impossible as $i < j$.

(3) For $i < \frac{2}{3}m$, $j = m+1$, $\mu(i) = \mu(0) \Longrightarrow i = 0$, which is impossible.

(4) For $\frac{2}{3}m \leq i, j < m$, $\mu(i+1) = \mu(j+1) \Longrightarrow i = j$.

(5) For $\frac{2}{3}m \leq i < m$, $j = m+1$, $\mu(i+1) = \mu(0) \Longrightarrow i+1 = 0$, which is impossible.

Positive Diagonals Check

(1) For $i, j \leq \frac{m+1}{3}$, $\mu(3i-1) = \mu(3j-1) \Longrightarrow i = j$.

(2) For $i \leq \frac{m+1}{3}$, $\frac{m+1}{3} < j < \frac{2}{3}m$, $\mu(3i-1) = \mu(3j-1) \Longrightarrow 3i - 1 = 2m - 3j + 1$, which is impossible for m divisible by 3.

(3) For $i \leq \frac{m+1}{3}$, $\frac{2}{3}m \leq j < m$, $\mu(3i-1) = \mu(3j) \Longrightarrow 3i - 1 = 3j - 2m$, which is impossible for m divisible by 3.

(4) For $i \leq \frac{m+1}{3}$, $j = m+1$, $\mu(3i-1) = \mu(2m+1) \Longrightarrow 3i - 1 = 1$, which is impossible.

(5) For $\frac{m+1}{3} < i, j < \frac{2}{3}m$, $\mu(3i-1) = \mu(3j-1) \Longrightarrow i = j$.

(6) For $\frac{m+1}{3} < i < \frac{2}{3}m$, $\frac{2}{3}m \leq j < m$, $\mu(3i-1) = \mu(3j) \Longrightarrow 2m - 3i + 1 = 3j - 2m \Longrightarrow 4m + 1 = 3(i+j)$, which is impossible for m divisible by 3.

(7) For $\frac{m+1}{3} < i < \frac{2}{3}m$, $j = m+1$, $\mu(3i-1) = \mu(2m+1) \Longrightarrow 2m - 3i + 1 = 1$, which is impossible for $i < \frac{2}{3}m$.

(8) For $\frac{2}{3}m \leq i, j < m$, $\mu(3i) = \mu(3j) \Longrightarrow i = j$.

(9) For $\frac{2}{3}m \leq i < m$, $j = m+1$, $\mu(3i) = \mu(2m+1) \Longrightarrow 3i - 2m = 1$, which is impossible for m divisible by 3.

This completes case (a). Cases (b) and (c) can be treated in an analogous manner; we leave their verification to the reader. ■

Final comments

We finish with two comments. First, it would be interesting to extend Theorems 1 and 2 to pillow chessboards that do not have the usual black-white pattern, as is already the case for the results on the torus. This means considering pillow boards that are obtained from the $n \times m$ torus, with m and/or n odd, by identifying diametrically opposite squares, as discussed above. This entails some unusual features; when n and m are both odd, the fundamental domain no longer consists of an integer number of squares, so the "board" looks rather odd. Also, if m is odd and n is even, one of the squares has an edge which is identified with itself!

Second, we remark that pillow chess is by no means the first chess game to be played on the sphere. There is a great number of chess variants (see Pritchard [**59**] and the web site www.chessvariants.com). In particular, we mention *Global Chess*, which is a commercial variant played on two revolving disks representing the hemispheres (here the "squares" at the poles are triangles), and Andrea Mori's *small spherical chess*, which is obtained from a cylindrical board by adding two additional (pole) squares, one at each end. Don Miller's *spherical chess* [**59**] as modified by Leo Nadvorney, is very close in appearance to pillow chess (the identifications are shown in FIGURE 16), but in fact this board is not a sphere but a Klein bottle!

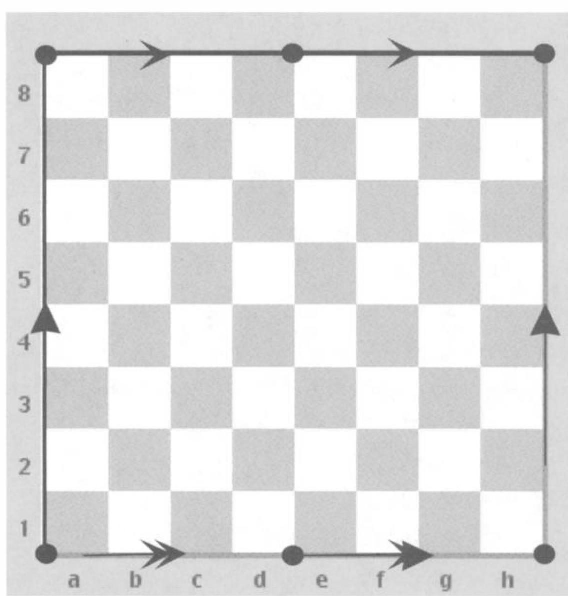


Figure 16

Solution to the pillow chess problem. First notice that white can't move its rook at h4, since this would disclose check by black's bishop at g5! As well, e2 and f1 are covered by black's bishop at d7. Thus black is threatening mate in three ways: Rf1, Nf1 and Nxe2. White can avoid all these attacks only by playing Rc1, in which case black mates with RxR.

Acknowledgments. My thanks go to John Bamberg for battling me in several games of pillow chess, to Andrej Panjkov for interesting suggestions for further generalizing pillow chess, and to Andrew Gove and Danny Sleator of Chessclub.com, whose software *ChessViewer* produced the chessboard graphics in this paper.

REFERENCES

1. W. Ahrens, *Mathematische Unterhaltungen und Spiele*, B.G. Teubner, Leipzig, 1921.
2. W. W. R. Ball, *Mathematical Recreations and Essays*, Revised by H. S. M. Coxeter, Macmillan and Co., London, 1947.
3. S. Barr, *Experiments in Topology*, Dover Publications, Inc., Mineola, NY, 1989.
4. J. D. Beasley, *The Mathematics of Games*, Oxford Uni. Press, Oxford, 1989.
5. C. Berge, *The Theory of Graphs*, Methuen and Co., London, 1962.
6. ———, *Graphs and Hypergraphs*, North Holland Publ., Amsterdam, 1973.
7. J-P. Bode and H. Harborth, Independent chess pieces on Euclidean boards, *J. Combin. Math. Combin. Comput.* **33** (2000), 209–223.
8. J. Boyer, *Nouveaux Jeux d'Echecs Non Orthodoxes*, J. Boyer, Paris, 1954.
9. A. Bruen and R. Dixon, The n-queens problem, *Discrete Math.* **12**:4 (1975), 393–395.
10. P. J. Campbell, Gauss and the eight queens problem: a study in miniature of the propagation of historical error, *Historia Math.* **4**:4 (1977), 397–404.
11. A. K. Chandra, Independent permutations, as related to a problem of Moser and a theorem of Pólya, *J. Combinatorial Theory* Ser. A **16** (1974), 111–120.
12. M.R. Chen, R. G. Sun, and J. C. Zhu, Partial n-solution to the modular n-queens problem II, *Combinatorics and graph theory*, 1–4, World Sci. Publishing, River Edge, NJ, 1993.
13. D. S. Clark, A combinatorial theorem on circulant matrices, *Amer. Math. Monthly* **92**:10 (1985), 725–729.
14. D. S. Clark and O. Shisha, Invulnerable queens on an infinite chessboard, *Ann. New York Acad. Sci.* **555** (1989), 133–139.
15. A. Conrad, T. Hindrichs, H. Morsy, and I. Wegener, Solution of the knight's Hamiltonian path problem on chessboards, *Discrete Appl. Math.* **50** (1994), no. 2, 125–134.
16. P. Cull and J. De Curtins, Knight's tour revisited, *Fibonacci Quarterly* **16**:3 (1978), 276–286.
17. R. B. Eggleton and A. Eid, Knight's circuits and tours, *Ars Combin.* **17A** (1984), 145–167.
18. C. Erbas and M. M. Tanik, Generating solutions to the N-queens problem using 2-circulants, this MAGAZINE **68**:5 (1995), 343–356.
19. C. Erbas, M. M. Tanik, and Z. Aliyazicioglu, Linear congruence equations for the solutions of the N-queens problem, *Inform. Process. Lett.* **41**:6 (1992), 301–306.
20. B-J. Falkowski and L. Schmitz, A note on the Queens' problem, *Inform. Process. Lett.* **23**:1 (1986), 39–46.
21. N. J. Fine, Solution of problem E1699, *Amer. Math. Monthly* **72** (1965), 552–553.
22. M. Gardner, Problems that are built on the knight's move in chess, *Scientific American* **217** (1967), no. 4, 128–132.
23. ———, *Further Mathematical Diversions*, Penguin Books, New York, 1977.
24. ———, *Fractal music, hypercards and more.... Mathematical recreations from Scientific American magazine*, W. H. Freeman and Company, New York, 1992.
25. Z. Goldstein, Solution of problem E2698, *Amer. Math. Monthly* **86** (1979), 309–310.
26. S. W. Golomb, Sphere packing, coding metrics, and chess puzzles, *Proc. Second Chapel Hill Conf. on Combinatorial Mathematics and its Applications*, Univ. North Carolina, Chapel Hill, NC, 1970, pp. 176–189.
27. R. K. Guy, *Unsolved Problems in Number Theory*, Springer-Verlag, New York, 1994.
28. O. Heden, On the modular n-queen problem, *Discrete Math.* **102**:2 (1992), 155–161.
29. ———, Maximal partial spreads and the modular n-queen problem, *Discrete Math.* **120**:1–3 (1993), 75–91; II, *ibid.* **142**:1–3 (1995), 97–106.
30. S. M. Hedetniemi, S. T. Hedetniemi, and R. Reynolds, Combinatorial Problems on chessboards II, *Domination in graphs, advanced topics* (T. W. Haynes, S. T. Hedetniemi, P. J. Slater eds.), Marcel Dekker, Inc., New York, 1998, pp. 133–162.
31. E. J. Hoffman, J. C. Loessi, and R. C. Moore, Constructions for the solution of the m queens problem, this MAGAZINE **42** (1969), 66–72.
32. R. Honsberger, *Mathematical Gems from Elementary Combinatorics, Number Theory, and Geometry*, MAA, Buffalo, 1973.
33. D. Hooper and K. Whyld, *The Oxford Companion to Chess*, Oxford Uni. Press, Oxford, New York, 1984.
34. S. P. Hurd and D. A. Trautman, The knight's tour on the 15-puzzle, this MAGAZINE **66**:3 (1993) 159–166.
35. F. K. Hwang and K. W. Lih, Latin squares and superqueens, *J. Combin. Theory* Ser. A **34**:1 (1983), 110–114.
36. E. Just, Solution of problem E2302, *Amer. Math. Monthly* **79**:5 (1972), 522–523.
37. D. A. Klarner, The problem of reflecting queens, *Amer. Math. Monthly* **74** (1967), 953–955.
38. ———, *Mathematical Review* 48 #10849 of "On pairings of the first $2n$ natural numbers", by G. B. Huff, *Acta Arith.* **23** (1973), 117–126.
39. ———, Queen squares, *J. Recreational Math.* **12**:3 (1979/80), 177–178.
40. T. Kløve, The modular n-queen problem, *Discrete Math.* **19**:3 (1977), 289–291; II, *ibid.* **36**:1 (1981), 33–48.
41. M. Kraitchik, *Le Problème du Cavalier*, Gauthiers-Villars, Paris, 1927.
42. ———, *Mathematical Recreations*, Dover Publ., New York, 1953.

43. O. Kyek, I. Parberry, and I. Wegener, Bounds on the number of knight's tours, *Discrete Appl. Math.* **74**:2 (1997), 171–181.
44. P. C. B. Lam, W. C. Shiu, and H. L. Cheng, Knight's tour on hexagonal nets, *Congr. Numer.* **141** (1999), 73–82.
45. L. C. Larson, A theorem about primes proved on a chessboard, this MAGAZINE **50**:2 (1977), 69–74.
46. M. Loebbing and I. Wegener, The number of knight's tours equals 33, 439, 123, 484, 294—counting with binary decision diagrams, *Electron. J. Combin.* **3**:1 (1996), #R5 and #R5 Comment 1.
47. E. Lucas, *Récréations Mathématiques*, Blanchard, Paris, 1960.
48. W. S. Massey, *A Basic Course in Algebraic Topology*, Springer-Verlag, New York, 1991.
49. B. McKay, Comments on "The number of knight's tours equals 33, 439, 123, 484, 294—counting with binary decision diagrams" [*Electron. J. Combin.* **3**:1 (1996), #R5] by M. Loebbing and I. Wegener, *Electron. J. Combin.* **3**:1 (1996), #R5 Comment 2.
50. P. Monsky, Solution of problem E3162, *Amer. Math. Monthly* **96** (1989), 258–259.
51. ———, Addendum: "On the modular n-queen problem," by O. Heden, *Discrete Math.* **118**:1–3 (1993), 293.
52. H. J. R. Murray, The knight's tour: ancient and oriental, *British Chess Mag.* (1903):Jan. 1–7.
53. ———, *A History of Chess*, Oxford Univ. Press, London, 1913.
54. V. V. Nikulin and I. R. Shafarevich, *Geometries and Groups*, Springer-Verlag, Berlin-New York, 1987.
55. S. P. Nudelman, The modular n-queens problem in higher dimensions, *Discrete Math.* **146**:1–3 (1995), 159–167.
56. M. Petković, *Mathematics and Chess*, Dover Publications, Inc., Mineola, NY, 1997.
57. G. Pólya, Über die "doppelt-periodischen" Lösungen des n-Damen-Problems, *George Pólya: Collected papers* Vol. IV (G-C. Rota, ed.), MIT Press, Cambridge, London, 1984, pp. 237–247.
58. V. V. Prasolov, *Intuitive Topology*, American Mathematical Society, Providence, RI, 1995.
59. D. B. Pritchard, *The Encyclopedia of Chess Variants*, Games and Puzzles Publ., Surrey, 1994.
60. M. Reichling, A simplified solution of the n queens problem, *Inform. Process. Lett.* **25**:4 (1987), 253–255.
61. I. Rivin, I. Vardi, and P. Zimmermann, The n-queens problem, *Amer. Math. Monthly* **101**:7 (1994), 629–639.
62. I. Rivin and R. Zabih, A dynamic programming solution to the n-queens problem, *Inform. Process. Lett.* **41**:5 (1992), 253–256.
63. I. Satake, On a generalization of the notion of manifold, *Proc. Nat. Acad. Sci. U.S.A.* **42** (1956), 359–363.
64. G. Schrage, The eight queens problem as a strategy game, *Internat. J. Math. Ed. Sci. Tech.* **17**:2 (1986), 143–148.
65. A. J. Schwenk, Which rectangular chessboards have a knight's tour?, this MAGAZINE **64**:5 (1991), 325–332.
66. H. D. Shapiro, Generalized Latin squares on the torus, *Discrete Math.* **24**:1 (1978), 63–77.
67. J. A. Shufelt and H. J. Berliner, Generating Hamiltonian circuits without backtracking from errors, *Theoret. Comput. Sci.* **132**:1–2 (1994), 347–375.
68. D. Singmaster, Some early sources in recreational mathematics, *Mathematics from manuscript to print, 1300–1600*, (C. Hay ed.), 195–208, Oxford Univ. Press, New York, 1988.
69. I. Stewart, *Another Fine Math You've Got Me Into...*, W. H. Freeman and Company, New York, 1992.
70. J. Stillwell, *Geometry of Surfaces*, Springer-Verlag, New York, 1992.
71. J. J. Tattersall, Vandermonde's contributions to the early history of combinatorial theory, *Ars Combin.* **25C** (1988), 195–203.
72. W. Thurston, *The Geometry and Topology of 3-Manifolds*, mimeographic notes, Princeton Uni., 1978–1979.
73. M. Valtorta and M. I. Zahid, Warnsdorff's tours of a knight, *J. Recreational Math.* **25**:4 (1993), 263–275.
74. ———, Tie-breaking rules for $4 \times n$ Warnsdorff's tours, *Congr. Numer.* **95** (1993), 75–86.
75. G. H. J. van Rees, Knight's tours and circuits on the $3 \times n$ chessboard, *Bull. Inst. Combin. Appl.* **16** (1996), 81–86.
76. I. Vardi, *Computational Recreations in Mathematica*, Addison-Wesley, Redwood City, 1991.
77. G. H. Verney, *Chess Eccentricities*, Longmans, Green and Co., London, 1885.
78. J. Watkins, Knight's tours on cylinders and other surfaces, *Congr. Numer.* **143** (2000), 117–127.
79. J. Watkins and R. L. Hoenigman, Knight's tours on a torus, this MAGAZINE **70**:3 (1997), 175–184.
80. R. J. Wilson, A brief history of Hamiltonian graphs, *Graph Theory in Memory of G.A. Dirac*, (Sandbjerg, 1985), *Ann. Discrete Math.*, **41**, North-Holland, Amsterdam-New York, 1989, pp. 487–496.
81. A. M. Yaglom and I. M. Yaglom, *Challenging Mathematical Problems with Elementary Solutions Vol 1*, Holden-Day, San Francisco, 1964.

Doubly Recursive Multivariate Automatic Differentiation

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

Automatic differentiation is a way to find the derivative of an expression without finding an expression for the derivative. More specifically, in a computing environment with automatic differentiation, you can obtain a numerical value for $f'(x)$ by entering an expression for $f(x)$. The resulting computation is accurate to the precision of the computer system—it does not depend on the approximation of derivatives by difference quotients. Indeed, the computation is equivalent to evaluating a symbolic expression for $f'(x)$, but no one has to find that expression—not even the computer system.

That's right. The automatic differentiation system never formulates a symbolic expression for the derivative. Automatically calling on something like *Mathematica* to produce a symbolic derivative, and then plugging in a value for x is the wrong image entirely. Automatic differentiation is something completely different.

Well OK, but so what? Symbolic algebra systems are so prevalent and powerful today, why should we be concerned with avoiding symbolic methods? There are two answers. The first is practical. Symbolic generation of derivatives can lead to exponential growth in the length of expressions. That causes computational problems in real applications. Accordingly, there is a practical applied side to the subject of automatic differentiation, as witnessed by the serious attention of computer scientists and numerical analysts [**3, 4**].

The second answer is more mathematical. It is a relatively easy task to create a single variable automatic differentiation system capable of evaluating first derivatives. In fact, writing in this MAGAZINE in 1986, Rall [**10**] gives a beautiful presentation of just such a system. What is mathematically interesting is an amazingly elegant extension of the one-variable/one-derivative system that handles essentially any number of variables and derivatives. The extension is recursively defined, employing an induction on both the number of variables and the number of derivatives, and using fundamental definitions that are virtually identical to the ones used in Rall's system.

The purpose of this paper is to present the recursive automatic differentiation system. To set the stage, we will begin with a brief review of Rall's one-variable/one-derivative system, followed by an example of the recursive system in action. Then the mathematical formulation of the recursive system will be presented. The paper will end with a brief discussion of practical issues related to the recursive system.

Rall's system

Because automatic differentiation is a computational technique, it is best understood in the context of a computer language. In particular, recall that in a scientific computer language such as Basic, or FORTRAN, variables correspond to memory locations. For example, consider the statements

$$x = 3$$
$$f = x^2 - 5.$$

The first causes a value of 3 to be stored in the memory location for x, while the second reads the value of x, squares it, subtracts 5, and stores the result in the memory location for f. We can think of this as a procedure for evaluating the function $f(x) = x^2 - 3$.

In Rall's system, the idea is to simultaneously evaluate both $f(x)$ and $f'(x)$. In this system, each variable corresponds to an ordered pair of memory locations, one for the value of a function, and one for the value of the derivative. Now the goal is for the statements above to produce the pair (4, 6), incorporating the values of both $f(3)$ and $f'(3)$.

This is accomplished as follows. First, when a variable is assigned a value in a statement such as $x = 3$ the automatic differentiation system stores in the memory for x the pair (3, 1). This corresponds to the value of the identity function $I(x) = x$, and its derivative, at $x = 3$. Second, any numerical constant that appears in an expression is represented by a pair corresponding to the value and derivative of a constant function. For the example above, the constant 5 is represented by (5, 0)—the value of the constant function $C(x) \equiv 5$, and its derivative. Finally, each operation appearing in the expression is carried out in an extended sense, operating on pairs. The rule for pair addition or subtraction is just the usual componentwise operation. The rule for pair multiplication is

$$(a_1, a_2) \times (b_1, b_2) = (a_1 b_1, a_2 b_1 + a_1 b_2). \tag{1}$$

Using these definitions, we can anticipate what the automatic differentiation system will do in response to the pair of statements

$$\begin{aligned} x &= 3 \\ f &= x^2 - 5. \end{aligned}$$

The first statement leads to the creation of the pair (3,1). The second statement translates into a sequence of operations on pairs:

$$\begin{aligned} f &= (3, 1) \times (3, 1) - (5, 0) \\ &= (3 \cdot 3, 1 \cdot 3 + 3 \cdot 1) - (5, 0) \\ &= (9, 6) - (5, 0) \\ &= (4, 6). \end{aligned}$$

As easily verified, this result correctly represents the value of both $x^2 - 5$ and its derivative at $x = 3$. Notice that there is no symbolic computation here. However, the equivalent of symbolic differentiation rules are built into the definitions of pair addition and multiplication. Thus, the expression for f is evaluated to produce both the value of the expression and of its derivative.

It should be stressed that the operations on pairs can be formulated without any reference to functions and derivatives. We adopt an abstract framework with objects (ordered pairs) and operations. As defined above, ordered pairs can be added, subtracted, and multiplied. In fact, extended operations for pairs can be defined for all the usual elementary functions. For example, the sine of a pair is defined according to

$$\sin(a_1, a_2) = (\sin a_1, a_2 \cos a_1). \tag{2}$$

Of course, these abstract definitions are inspired by the idea that each ordered pair will contain values of a function and its derivative. To make the connection explicit, we will use the notation $f^{[1,1]}(x) = (f(x), f'(x))$, where the [1, 1] indicates the presence of one variable, and the inclusion of one derivative. Thus, in the original computation, we found $f^{[1,1]}(3) = (4, 6)$. Similarly, using the sine operation for pairs, the statements

$$x = 3$$
$$g = \sin(x^2 - 5)$$

result in the computation of $\sin(4, 6) = (\sin 4, 6\cos 4)$. The elements of this ordered pair are the correct values of $\sin(x^2 - 5)$ and its derivative at $x = 3$. That is, with $g(x)$ defined as $\sin(x^2 - 5)$, the lines above compute $g^{[1,1]}(3)$.

What makes the system work is that each operation correctly propagates derivative values. For the arithmetic operations, that means

$$\begin{aligned} f^{[1,1]}(3) + g^{[1,1]}(3) &= (f + g)^{[1,1]}(3) \\ f^{[1,1]}(3) - g^{[1,1]}(3) &= (f - g)^{[1,1]}(3) \\ f^{[1,1]}(3) \times g^{[1,1]}(3) &= (fg)^{[1,1]}(3). \end{aligned} \tag{3}$$

Observe that the rules for addition, subtraction, and products of pairs are based on the sum and product rules for derivatives. Similarly, (2) is really nothing more than the chain rule, since the derivative of $\sin(f(x))$ is given by $\cos(f(x))f'(x)$. With a_1 in place of $f(x)$ and a_2 in place of $f'(x)$, this becomes $\cos(a_1)a_2$. That shows that in (2), if $(a_1, a_2) = f^{[1,1]}(3)$, then $\sin(a_1, a_2) = \sin(f^{[1,1]}(3)) = (\sin \circ f)^{[1,1]}(3)$. In a similar way, any differentiable function ϕ can be extended to pairs by the formula

$$\phi(a_1, a_2) = \big(\phi(a_1), \phi'(a_1)a_2\big). \tag{4}$$

With this definition, we have

$$\phi(f^{[1,1]}) = (\phi \circ f)^{[1,1]}. \tag{5}$$

Although these examples pertain to a function of a single variable, and involve only a single derivative, it is easy to envision extensions involving several variables and partial derivatives of various orders. Throughout, we will restrict our attention to functions sufficiently smooth so that order of differentiation does not matter.

In the recursive system that we will present below, the idea is to compute all of the partial derivatives up to some specified order. In this system, evaluating a function f at a point in its domain means determining an object $f^{[n,m]}$ that contains the function value as well as the values of all partial derivatives through order m with respect to n variables. These objects are referred to as *derivative structures*. Since m defines the maximum number of derivatives, it is called the derivative index. Similarly, n is the variable index. As in the discussion above, we can proceed abstractly by defining derivative structures and appropriate operations without any mention of functions and derivatives. However, given a function f, we do need some way to construct $f^{[n,m]}$ as one of our abstract derivative structures, and equations analogous to (3) and (5) must hold.

The recursive system in action

Before describing the abstract system, let's take a look at how the system operates. Consider the function

$$f(x, y, z) = \frac{\sqrt{x + y}}{\sqrt{z - y}},$$

and suppose we wish to evaluate f and all partial derivatives through second order at the point $(4, 5, 14)$. The recursive automatic differentiation system can be given this problem with the following commands (with slightly modified syntax for readability):

```
x = DS-Make-Var(3,2,1,4)
y = DS-Make-Var(3,2,2,5)
z = DS-Make-Var(3,2,3,14)
u = DS-Sqrt(DS-Add(x,y))
v = DS-Sqrt(DS-Sub(z,y))
Print DS-Divide(u,v)
```

These commands involve applications of several different functions within the automatic differentiation system. First, there are three invocations of `DS-Make-Var`. This function creates the derivative structures corresponding to the independent variables x, y, and z. For example, `x = DS-Make-Var(3,2,1,4)` creates a derivative structure for **3** variables, and for partial derivatives through order **2**, corresponding to variable number **1** (x), and assigning that variable a value of **4**. This command is the equivalent of $x = 4$ in the one-variable/one-derivative system. Similarly, the next two statements create the derivative structures corresponding to variables y and z, assigning values of 5 and 14, respectively. The other commands are the derivative structure versions of standard operations; `DS-Add` is addition of derivative structures, `DS-Sqrt` applies the square root for derivative structures, and so on. So the fourth statement adds the derivatives structures for `x` and `y` and takes the square root of the result. That defines a new derivative structure, `u`. Similarly, the next line defines `v` by subtracting `y` from `z`, and applying the derivative structure for square roots. The final command applies derivative structure division to `u` and `v`, and prints the result.

As in Rall's system, the computations above are completely numerical. For example, the derivative structure for the variable x stores the value of x, 4, as well as all the partial derivatives through second order with respect to x, y, and z. These values are, of course, trivially determined. The partial derivative with respect to x is 1, and all the other partial derivatives are 0. But the point is that the derivative structure called `x` is just some sort of array with entries of 4, 1, and many zeroes. In the same way, `y` and `z` are arrays of numbers as well. When these are combined according to the commands listed above, the final result is printed out as

0.01235					
0.11111	0.00000		−0.01235		
1.00000	0.05556	−0.00309	−0.05556	−0.00309	0.00926.

These are the values of f and its derivatives, in the following arrangement:

$$\begin{array}{llllll} f_{yy} & & & & & \\ f_y & f_{xy} & & f_{yz} & & \\ f & f_x & f_{xx} & f_z & f_{xz} & f_{zz}. \end{array}$$

The subscripts indicate partial differentiation: f_x for $\frac{\partial f}{\partial x}$, f_{xy} for $\frac{\partial^2 f}{\partial y \partial x}$, and so on. The rationale for laying out the derivatives in this way will become clear when the general system is defined. For this example, it is enough to see how the system operates, and to observe that all the desired partial derivatives are correctly computed.

At this point, I hope that the basic idea of the automatic differentiation system is clear. Numerical values for a function and its derivatives are arranged in some sort of data structure, and operations on these structures are defined according to the rules of differentiation so that derivatives are correctly propagated. The structures for the simplest functions, namely the constant functions (like $c(x, y, z) \equiv 5$) and variables (like

$I_1(x, y, z) \equiv x$) are easy to specify directly. By operating on these simple derivative structures, we can formulate derivative structures for essentially arbitrary expressions involving the variables and elementary functions.

Although these ideas are feasible in principle, I also hope the reader has some sense of the difficulty of handling all the details in practice. At first glance, the idea of defining appropriate structures to contain all the partial derivatives through second order relative to three variables, and then specifying the proper operations of arithmetic, as well as proper definitions for functions like sine and cosine, should seem fairly intimidating, or at least unpleasantly tedious. Happily, and surprisingly, there is a remarkably simple recursive formulation that is no more complicated than Rall's one-variable/one-derivative system. Indeed, considered formally, the operations within this recursive formulation are virtually identical to the operations in Rall's system. With that in mind, let us turn now to the recursive development of an automatic differentiation system.

The objects

The first step in constructing the recursive system is to define the objects, or derivative structures, on which we will operate. Let us consider a few motivating examples. First, for functions of a single variable, automatic calculation of m derivatives can be provided by operating on $(m + 1)$-tuples. A typical object in the system, $a = (a_0, a_1, \cdots, a_m)$, includes the value of a function and its first m derivatives. For example, with $m = 3$, we can write

$$a = f^{[1,3]} = (f, f_x, f_{xx}, f_{xxx}).$$

For a function of two variables, assuming equality of mixed partials, the partial derivatives through order m are conveniently arranged in a triangular array. This is illustrated in FIGURE 1 for $m = 3$. It is important to note that the entry in the lower left-hand corner has a special significance. In the derivative structure $f^{[2,m]}$, the lower left-hand corner is the value of the original function f.

$$\begin{array}{llll} f_{yyy} & & & \\ f_{yy} & f_{yyx} & & \\ f_y & f_{yx} & f_{yxx} & \\ f & f_x & f_{xx} & f_{xxx} \end{array}$$

Figure 1 Layout of $f^{[2,3]}$

Observe that the array in FIGURE 1 can be decomposed into two parts. The bottom row is a vector of derivatives with respect to a single variable, as described in the preceding paragraph. That is, the bottom row is just $f^{[1,3]}$. The second part, all of the triangle *except* the bottom row, is also a derivative structure, namely ${f_y}^{[2,2]}$; it contains the value of f_y, and all of its first and second order partial derivatives with respect to x and y. This gives $f^{[2,3]}$ as a combination of $f^{[1,3]}$ and ${f_y}^{[2,2]}$.

In a similar way, we can lay out the entries of $f^{[3,3]}$, that is, the partial derivatives through third order with respect to three variables (see FIGURE 2). The partial derivatives are arranged in a pyramid composed of several triangular layers. Each layer has the same form as the triangular array in FIGURE 1. As before, there is a distinguished entry identifying the function f, at the lower left-hand corner of the lowest level. Also, as before, there is a natural decomposition into two parts. The first part is the bottom triangular array, which is recognizable as $f^{[2,3]}$. It contains all partial derivatives through order $m = 3$ with respect to x and y. The complementary part is the sub-pyramid made up of levels 2, 3, and 4. This can be recognized as $f_z^{[3,2]}$. It contains all partial derivatives relative to the three variables x, y, and z, through order 2 of the function f_z. The decomposition gives $f^{[3,3]}$ as a combination of $f^{[2,3]}$ and $f_z^{[3,2]}$.

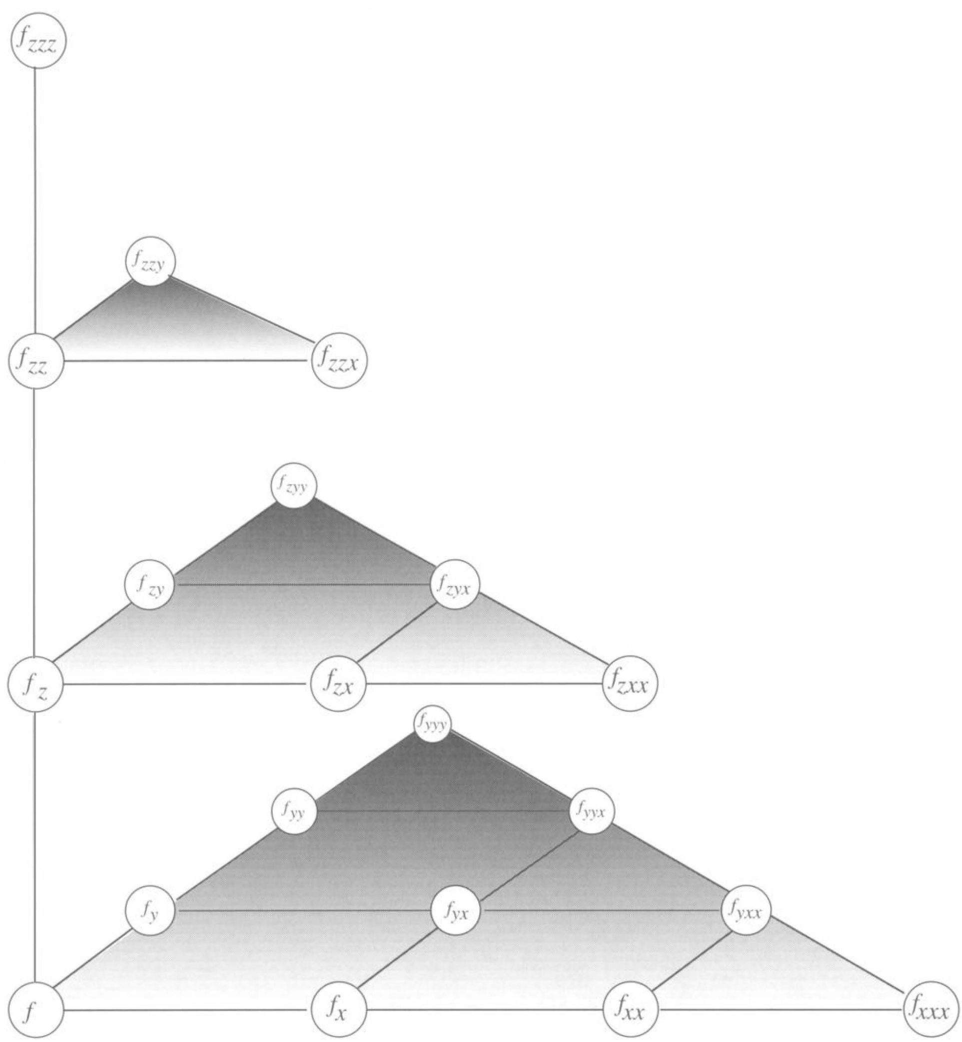


Figure 2 Layout of $f^{[3,3]}$

These examples suggest a hierarchy of automatic differentiation objects. For any n and m, we can imagine a set of objects that contain all partial derivatives through order m with respect to n variables. These will be our *derivative structures*. Thus, for a single variable we have derivative vectors; for two variables, derivative triangles; for three variables, derivative pyramids; and in general, derivative structures.

The decomposition discussed in the examples above can be described in general using the terminology of derivative structures. For each example we considered, a derivative structure of partial derivatives through order m with respect to n variables was partitioned into two smaller derivative structures. The first part had the same number of derivatives (m) and one fewer variables ($n - 1$) than the original structure, while the second part had one fewer derivatives ($m - 1$) and the same number of variables as the original. These observations inspire the following recursive definition of derivative structures.

DEFINITION 1. *For $m, n \geq 0$, we define $DS(n, m)$, the set of derivative structures with derivative index m and variable index n, as follows. If $m = 0$ or $n = 0$, $DS(n, m)$ is just $\mathbb{R}$, the real numbers. Otherwise*

$$DS(n, m) = DS(n-1, m) \times DS(n, m-1)$$

(where $\times$ denotes the Cartesian product).

It should be emphasized here that this definition makes no mention of functions or derivatives. It abstractly defines a class of objects, built up recursively, and reducing to real numbers at the lowest level of the recursion. In this context, a derivative structure is understood most simply as a binary tree, with real numbers as the leaves. An element $a \in DS(4, 7)$, for example, has two components, one in $DS(3, 7)$ and the other in $DS(4, 6)$. Each of these components likewise has two components, as shown in FIGURE 3. Each branch of the tree ends when one of the two indices reaches zero, indicating that the corresponding component is a real number. For $a = f^{[n,m]}$, the real numbers at the leaves are simply the values of partial derivatives of f. However, this visualization turns out to be of limited value. Instead, the best approach is to retain the recursive image of an element of $DS(n, m)$ as an ordered pair, each of whose components is a lower order derivative structure.

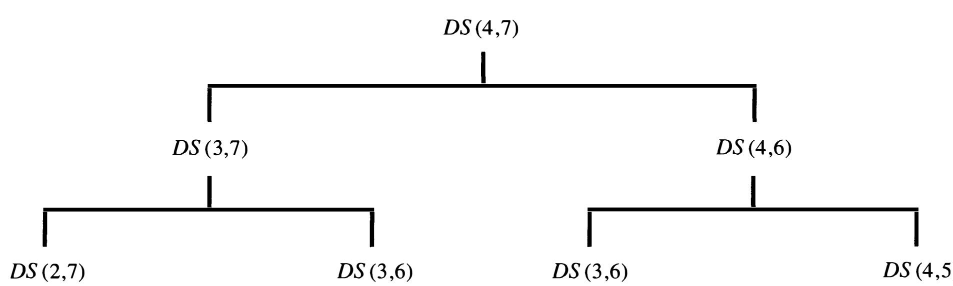


Figure 3 Partial Tree for $a \in DS(4, 7)$.

The idea of a derivative structure as an ordered pair hints at the connection to Rall's automatic differentiation system. Shortly we will see that the definitions for operations on derivative structures make this connection into a perfect analogy. But there is one final prerequisite needed. In terms of the triangular arrays and pyramids considered earlier, the two components of a derivative structure are particular substructures. For example, if $a = (a_1, a_2)$ is a derivative pyramid, then a_1 is a derivative triangle, and a_2 is a smaller derivative pyramid. We also need a third substructure, denoted a_1^*. Later an abstract recursive definition of a_1^* will be provided. But conceptually, think of a_1^* as follows: If the derivative structure $a = f^{[n,m]}$, then it contains within it $f^{[n,m-1]}$, the derivatives up to order $m - 1$. That substructure is a_1^*. Thus, in FIGURE 1, a_1 is the bottom row, a_2 is the sub-triangle consisting of everything but the bottom row, and a_1^* is the triangle that contains everything except the third order derivatives lying along the hypotenuse. Notice that a_2 and a_1^* have the same size and shape, but are derivative structures for different functions. Similarly, in FIGURE 2, the triangle on the lowest level is a_1, the remaining levels form the sub-pyramid a_2, and a_1^* is the sub-pyramid consisting of everything except the highest order derivatives lying on the slanting outer face of the pyramid.

This completes the background we need to define derivative structure operations. We know that a derivative structure a is an ordered pair (a_1, a_2), that the components are derivative substructures of lower order, and that a_1^* is another sub-structure with the same size and shape as a_2. The operations on derivative structures are defined in terms of these substructures.

Operations on derivative structures

To build expressions out of derivative structures, we need to be able to apply arithmetic operations and elementary functions. By considering the reciprocal function $r(x) = 1/x$ as one of our elementary functions, we eliminate the need to define derivative structure division. To divide a/b we simply multiply $a \times r(b)$. Accordingly, the only arithmetic operations that we need are addition, subtraction, and multiplication. As a convenience we will also include scalar multiplication.

The definitions of all the arithmetic operations are recursive. The case of addition, subtraction, and scalar multiplication will make this clear.

DEFINITION 2. *For $DS(0, m)$ and $DS(n, 0)$, the elements are real numbers and addition, subtraction, and multiplication are the usual real number operations. For $n, m > 0$, let $a = (a_1, a_2)$ and $b = (b_1, b_2)$ be elements of $DS(n, m)$, and let r be a real number. Then addition, subtraction, and scalar multiplication are defined by*

$$\begin{aligned} a + b &= (a_1 + b_1, a_2 + b_2) \\ a - b &= (a_1 - b_1, a_2 - b_2) \\ ra &= (ra_1, ra_2). \end{aligned}$$

Formally, these are identical to the componentwise definitions in Rall's system. But they have a slightly different meaning in the present context. To add a and b we must add their components, which are themselves derivative structures. The computer implementation of the addition is thus recursive. To add two elements of $DS(3, 4)$, for example, we recall the addition operation for components in $DS(3, 3)$ and in $DS(2, 4)$. Those additions, in turn, spawn additions of more derivative structures. At each recursion, though, one of the two indices is reduced. Eventually, an index becomes zero, and the recursion terminates with an addition of real numbers. Subtraction and scalar multiplication operate similarly.

The definition of multiplication is again an analog of what we saw in Rall's system.

DEFINITION 3. *For $DS(0, m)$ and $DS(n, 0)$ multiplication is defined to be the usual real number operation. For $n, m > 0$, if $a = (a_1, a_2)$ and $b = (b_1, b_2)$ are elements of $DS(n, m)$, define*

$$a \times b = (a_1 \times b_1, a_2 \times b_1^* + a_1^* \times b_2).$$

Formally, this is virtually identical to the one-variable/one-derivative multiplication rule defined by (1). The only difference is that there are no asterisks in (1). Indeed, the ordered pairs in Rall's systems are elements of $DS(1, 1)$, and in that setting, a_1 and a_1^* are identical. However, while there are clear formal similarities between multiplication in Rall's system and in $DS(n, m)$, it must be remembered that in the latter system the definition is recursive. As for the operations of addition, subtraction, and scalar multiplication, the multiplication of derivative structures requires multiplying their components, and hence a recursive use of multiplication. And as we saw earlier, the recursive process keeps generating more and more multiplications, finally reaching a point at which the derivative objects reduce to real numbers. So, while the multiplication definition seems to have the same simplicity as in Rall's system, under the surface there is a complex sequence of operations implicitly defined.

Finally we come to the elementary functions. Given a derivative structure a and an elementary function ϕ, we wish to define $\phi(a)$. Once again, the definition is almost identical to what appeared in Rall's system.

DEFINITION 4. *Let ϕ be an m-times differentiable function of a real variable. If $m = 0$ or $n = 0$, $DS(n, m)$ is just $\mathbb{R}$ and ϕ is applied to the elements in the usual way. For $n, m > 0$, if $a = (a_1, a_2) \in DS(n, m)$, define*

$$\phi(a) = \left(\phi(a_1), \phi'(a_1^*) \times a_2\right).$$

This definition is a direct analog of (4), to which it reduces in the case that $n = m = 1$. As we saw with multiplication, the only formal difference is the appearance of an asterisk in the general derivative structure definition. Here again, the actual computation of $\phi(a)$ is recursive, and the recursion terminates when ϕ or one of its derivatives is finally called upon to operate on a real number.

That's it. That is all you need to construct arbitrary elementary function expressions involving general derivative structures. As promised, the definitions are virtually the same as those in Rall's system, and yet they provide for the automatic generation of partial derivatives to essentially arbitrary order with respect to an essentially arbitrary number of variables. But the presentation is not quite complete. We still have to see how to create the fundamental derivative structures that correspond to constants and variables. And at some point we need to see why the definitions just given really work.

Fundamental derivative structures

So far, we have defined derivative structures and their operations abstractly, without mention of functions and partial derivatives. To make the connection with automatic differentiation clear, we must have a definition of $f^{[n,m]}$ as an element of $DS(n, m)$.

DEFINITION 5. *Let f be a function of at least n variables with continuous partial derivatives through order m, and let x be an element of the domain of f. Then the derivative structure for f with derivatives through order m with respect to the first n variables is given at x by*

$$f^{[n,m]}(x) = \begin{cases} f(x) & \text{if } n = 0 \text{ or } m = 0 \\ \left(f^{[n-1,m]}(x), (\partial_n f)^{[n,m-1]}(x)\right) & \text{otherwise} \end{cases}$$

where ∂_n denotes partial differentiation with respect to the nth variable of f.

This definition is a formalization of the pattern we saw in special cases, but some caution is needed. How do we know that $f^{[n,m]}$, as defined here, really does contain all the partial derivatives it is supposed to? For now the reader is asked to accept the validity of the definition. We will return to the justification in the next section.

Given the preceding definition, we can construct derivative structures for constants recursively. For example, to create the derivative structure for the constant 5, we consider the constant function $f(x, y, z, \ldots) \equiv 5$. Now $f^{[n,m]}$ has two components. The first is $f^{[n-1,m]}$, and that can be constructed recursively. The second is $\partial_n f^{[n-1,m]}$, and since f is constant, the partial derivative is 0. But that is again a constant function. Thus, a recursive construction algorithm can operate similarly to the operation algorithms. To construct a constant in $DS(n, m)$, we must first construct constants in

$DS(n-1, m)$ and $DS(n, m-1)$. The recursion proceeds until one index becomes 0, and at that point the value of the constant is returned. That constant is 5 just once, corresponding to tracing the left branch all the way down the tree to a leaf. In any path that involves a right branch, the function will be differentiated at least once, and it will be a zero function that is finally evaluated. On some level, however, this image is irrelevant. All that really matters is that a simple recursive construction algorithm for constants exists in the automatic differentiation system.

To illustrate the situation for the independent variables, let's consider the function $I_2(x, y, z) = y$. How do we construct ${I_2}^{[3,2]}$ at $y = 8$, for example? At the top level, ${I_2}^{[3,2]}$ is an ordered pair. The first component is ${I_2}^{[2,2]}$, which will be constructed recursively. The second component is $\partial_z {I_2}^{[3,1]}$, and since $\partial_z y = 0$ that is just the derivative structure of the constant 0. It can be constructed using the algorithm for a constant. At the next level, ${I_2}^{[2,2]}$ is decomposed into ${I_2}^{[1,2]}$ and $\partial_y {I_2}^{[2,1]}$. For the first of these, notice that the first index is 1. This is a derivative structure that does not involve any derivatives with respect to y, and for its construction we can treat y as the constant 8. For the second component, $\partial_y I_2 = \partial_y y = 1$. Again we need only construct a derivative structure for a constant. In a similar way, the derivative structure for any of the independent variables can be constructed recursively. Indeed, ${x_j}^{[n,m]} = (a_1, a_2)$ is defined as follows: If $j < m$, then a_1 is defined by a recursive construction of ${x_j}^{[n,m-1]}$ and a_2 is a derivative structure for the constant 0. If $j = m$, then a_1 is constructed as a constant derivative structure, with whatever value was assigned to x_j, and a_2 is the derivative structure for the constant 1. And if $j > m$, a_1 is again a constant derivative structure with the value of x_j, but a_2 is the derivative structure of the constant 0.

This is the construction used to define `DS-Make-Var` in the sample computation presented earlier. In fact, if you review that computation, you will see that we have now defined every operation that appears there. The automatic differentiation system is complete. With algorithms for constructing derivative structures for independent variables and constants, and definitions of derivative structure operations and elementary functions, nothing more is needed. However, we have yet to see any verification that the system actually works. How do we know, for example, that the arithmetic definitions propagate derivatives correctly? How do we know that applying an elementary function to a derivative structure as in Definition 4 produces the desired derivative information at the end? For that matter, how do we even know that the recursive definition for $f^{[n,m]}$ is correct? The next section will address these questions.

Validation of the system

There are two aspects of the system that require validation. First, we have to verify that the recursive definition of $f^{[n,m]}$ properly represents the intuition suggested by the triangle and pyramid examples. Second, it must be established that the definitions of derivative structure operations correctly propagate derivative information. That is, we must see that

$$\begin{aligned} f^{[n,m]} + g^{[n,m]} &= (f+g)^{[n,m]} \\ f^{[n,m]} - g^{[n,m]} &= (f-g)^{[n,m]} \\ f^{[n,m]} \times g^{[n,m]} &= (fg)^{[n,m]} \end{aligned} \tag{6}$$

and

$$\phi(f^{[n,m]}) = (\phi \circ f)^{[n,m]}. \tag{7}$$

For both of these ends, expressing a derivative structure a as an ordered pair (a_1, a_2) and referring to the components and to a_1^* will be of central importance. It simplifies the presentation to express these substructures using an operator notation. Thus, if $a = (a_1, a_2)$ is a derivative structure, we define $V(a) = a_1$ and $D(a) = a_2$. The names of these operators reflect the meaning of the components in the one-variable/one-derivative system, where a_1 is the *value* of the function, and a_2 is the *derivative*. Recall that a_1^* is obtained from a by removing all the highest order derivatives, so that a_1^* is a *lower order* version of a. Accordingly, we use the notation $L(a) = a_1^*$.

Although the conceptual meaning of the operators is clear, formal definitions will be given for completeness. For L, this is particularly important as there has not yet been given an abstract definition in terms of derivative structures.

DEFINITION 6. *Let $a \in DS(n, m)$. If $n = 0$ or $m = 0$, a is a real number and $V(a)$, $D(a)$, and $L(a)$ are all defined to equal a. Otherwise, $a = (a_1, a_2)$. In this case, we define $V(a) = a_1$, $D(a) = a_2$, and $L(a)$ according to*

$$L(a) = \begin{cases} L(a_1) & \text{if } m = 1 \\ (L(a_1), L(a_2)) & \text{if } m > 1. \end{cases}$$

It may not be immediately apparent that this definition of L is consistent with the earlier explanation of a_1^*. The reader may wish to verify that the definition works correctly for triangles and pyramids. However, for the arguments that will follow, it is not logically necessary to connect the definition of L with the conceptual image of $f^{[n,m]}$. Instead, we will be content to take $L(a)$ as the *definition* of a_1^*, and show that this definition has the properties we need for automatic differentiation.

The three operators provide the means to connect the abstract definition of $DS(n, m)$ to the ideas illustrated by the derivative vectors, triangles, and pyramids. As a first instance of this, we have the following result.

THEOREM 1. *Derivative structures for functions are related to the operations V, D, and L as follows:*

$$V(f^{[n,m]}) = f^{[n-1,m]}$$
$$D(f^{[n,m]}) = (\partial_n f)^{[n,m-1]}$$
$$L(f^{[n,m]}) = f^{[n,m-1]}.$$

If the derivatives in $f^{[n,m]}$ are laid out as in the examples of triangles and pyramids, these identities are obvious. However, it is possible to prove the identities using only the abstract definitions of the operators and of $f^{[n,m]}$. In fact, the first two identities are immediate consequences of the abstract definition of $f^{[n,m]}$. The third identity can be proved by a straightforward induction argument that exploits the recursive definitions of both $f^{[n,m]}$ and L. This same style of proof is effective for a number of the results to follow, and while a detailed proof for the third identity above will not be given, a sample proof will be given for a later theorem. In any case, it is important to note that the induction proof uses only the abstract definitions of L and $f^{[n,m]}$, and so makes no direct use of the full image of how partial derivatives are laid out in $f^{[n,m]}$. Thus, the fact that the third identity can be established by an abstract proof confirms that, at least in this regard, L and $f^{[n,m]}$ operate according to expectation.

Theorem 1 lends itself to a simple formal algorithm for applying V, D, or L to $f^{[n,m]}$: V decrements the variable index by 1; L decrements the derivative index by 1; and D both decrements the derivative index and differentiates f once with respect to the nth variable. Using just the first two of these rules we can prove the next result.

THEOREM 2. *Suppose $f^{[n,m]}$ is defined at x. Let e_j be a nonnegative integer for $1 \le j \le n$ with $\sum e_j \le m$. Then the partial derivative $\partial_1^{e_1} \cdots \partial_n^{e_n} f(x)$ can be obtained from $f^{[n,m]}(x)$ as follows: If $\sum e_j = m$ then*

$$\partial_1^{e_1} \cdots \partial_n^{e_n} f(x) = D^{e_1} V D^{e_2} V \cdots V D^{e_n} f^{[n,m]}(x);$$

otherwise

$$\partial_1^{e_1} \cdots \partial_n^{e_n} f(x) = V D^{e_1} V D^{e_2} V \cdots V D^{e_n} f^{[n,m]}(x).$$

The proof is simply a matter of applying the identities in Theorem 1. Rather than present the details in a formal way, it will be more illuminating to work through an example. Consider the derivative structure $f^{[3,6]}$ and suppose we want to obtain $\partial_1^2 \partial_2 \partial_3^2 f(x)$. Since this is a fifth derivative and $m = 6$, the theorem says to compute $V D^2 V D V D^2 f^{[3,6]}$. We can verify that the desired result is obtained by applying the identities in Theorem 1 as follows:

$$\begin{aligned}
V D^2 V D V D^2 f^{[3,6]}(x) &= V D^2 V D V (\partial_3^2 f)^{[3,4]}(x) \\
&= V D^2 V D (\partial_3^2 f)^{[2,4]}(x) \\
&= V D^2 V (\partial_2 \partial_3^2 f)^{[2,3]}(x) \\
&= V D^2 (\partial_2 \partial_3^2 f)^{[1,3]}(x) \\
&= V (\partial_1^2 \partial_2 \partial_3^2 f)^{[1,1]}(x) \\
&= (\partial_1^2 \partial_2 \partial_3^2 f)^{[0,1]}(x) \\
&= \partial_1^2 \partial_2 \partial_3^2 f(x).
\end{aligned}$$

This example reveals the general nature of the algorithm for extracting a particular derivative from $f^{[n,m]}$. Notice that the D operator only performs differentiation of $f^{[n,m]}$ with respect to x_n. But each time we apply V, we reduce the value of n, and hence change the variable that D differentiates. If we want a certain number of derivatives with respect to x_n, we apply D that many times. Then we apply V, in effect, shifting the focus to x_{n-1}. If we want one or more derivatives with respect to x_{n-1}, we apply D that many times again. So we continue, alternately applying D to differentiate and V to shift to a new variable, until all the desired derivatives have been applied. For an mth derivative, there will be m applications of D, reducing the derivative index to 0, and so reducing the derivative structure to a real number. Otherwise, there will be exactly n applications of V. This will reduce the variable index to 0, and so again result in a real number.

It should be stressed again that the operators V and D were defined completely abstractly, with no reference to derivatives. In a computational system, a particular derivative structure is simply an organized network of memory locations which store real values. The algorithm above navigates through such a network to a particular entry. Theorems 1 and 2 show that when a derivative structure is constructed according to the abstract definition of $f^{[n,m]}$, the desired derivative values can all be located and extracted. More specifically, visualizing the network as a binary tree, each application of V selects a left branch from a node, each application of D selects a right branch, and after either m applications of D or n applications of V a terminal node is reached. Thus, Theorem 2 can be understood as a prescription for finding the appropriate terminal node for a particular partial derivative.

To complete the validation of the system, we must see that derivative structure operations really do succeed in constructing $f^{[n,m]}$. That is, we must verify (6) and (7). The formal statement is given in the following theorem.

THEOREM 3. *Let f and g be real valued functions of n or more variables, with continuous partial derivatives through order m, let x be in the domain of f and g, let r be a real number, and let ϕ be a real function m times differentiable at $f(x)$. Then the following identities hold:*

$$\begin{aligned}
f^{[n,m]}(x) + g^{[n,m]}(x) &= (f+g)^{[n,m]}(x) \\
f^{[n,m]}(x) - g^{[n,m]}(x) &= (f-g)^{[n,m]}(x) \\
rf^{[n,m]}(x) &= (rf)^{[n,m]}(x) \\
f^{[n,m]}(x) \times g^{[n,m]}(x) &= (fg)^{[n,m]}(x) \\
\phi(f^{[n,m]}(x)) &= (\phi \circ f)^{[n,m]}(x).
\end{aligned}$$

As mentioned earlier, the recursive nature of the definitions makes induction a natural approach to proving results like these. To illustrate, here is a proof of the final identity above. It assumes that the preceding identities have already been established.

Proof. The proof is by induction on $n+m$. If either n or m is zero, the conclusion holds trivially. So assume that both n and m are positive, and that the conclusion holds for $f^{[n',m']}$ whenever $n'+m' < n+m$. From the definition of ϕ for derivative structures, if $f^{[n,m]}(x)$ is expressed as the pair (a_1, a_2), then

$$\phi\left(f^{[n,m]}(x)\right) = \left(\phi(a_1), \phi'(a_1^*) \times a_2\right).$$

In terms of the V, D, and L operators, this becomes

$$\phi\left(f^{[n,m]}(x)\right) = \left(\phi\left(Vf^{[n,m]}(x)\right), \phi'\left(Lf^{[n,m]}(x)\right) \times Df^{[n,m]}(x)\right).$$

Applying Theorem 1 we obtain

$$\phi\left(f^{[n,m]}(x)\right) = \left(\phi\left(f^{[n-1,m]}(x)\right), \phi'\left(f^{[n,m-1]}(x)\right) \times (\partial_n f)^{[n,m-1]}(x)\right).$$

Now we are ready to use the induction hypothesis. On the right side of the preceding equation, the real functions ϕ and ϕ' are applied to derivative structures with lower order than $f^{[n,m]}(x)$. By induction, we can *bring ϕ and ϕ' inside their respective parentheses*, leading to

$$\phi\left(f^{[n,m]}(x)\right) = \Big((\phi \circ f)^{[n-1,m]}(x), (\phi' \circ f)^{[n,m-1]}(x) \times (\partial_n f)^{[n,m-1]}(x)\Big).$$

Similarly, the identity for derivative structure multiplication allows us to bring the product on the right side of the equation inside the parentheses. Performing that reduction and recognizing the normal real function chain rule then produces

$$\begin{aligned}
\phi(f^{[n,m]}(x)) &= \Big((\phi \circ f)^{[n-1,m]}(x), (\phi' \circ f \cdot \partial_n f)^{[n,m-1]}(x)\Big) \\
&= \left((\phi \circ f)^{[n-1,m]}(x), [\partial_n(\phi \circ f)]^{[n,m-1]}(x)\right) \\
&= (\phi \circ f)^{[n,m]}(x).
\end{aligned}$$

This shows that the identity holds for $f^{[n,m]}$, completing the induction argument. ∎

This concludes the validation of the recursively defined automatic differentiation system. It has been demonstrated that the simple recursive definitions for derivative structure operations properly propagate partial derivatives. To put it more simply, we have seen that the recursive automatic differentiation system *works*. In a final section, we discuss a few ideas connected with implementation and computational efficiency.

Implementation and efficiency

The recursive automatic differentiation system presented here can be implemented in any computer programming language that supports recursion. A working version is described in [**6**]. There, the interested reader will find LISP code for the entire system, amounting to about 150 lines. Although the presentation in [**6**] is from a different point of view than the double recursion described here, the LISP code can be considered an implementation of either point of view. In fact, the double recursion described here was discovered as a direct result of studying the implementation in [**6**]. It should also be mentioned that the original idea for treating the number of derivatives recursively is due to Neidinger [**8**]. His work provided a critical inspiration for both the approach of [**6**] and the double recursion presented here.

It is beyond the scope of this paper to discuss the LISP implementation in detail. However, there is one aspect that is worth considering. The programming for the automatic differentiation system must include derivative structure formulations for all the familiar elementary functions: exponential, sine, cosine, etc. Each of these is programmed according to Definition 4. Interestingly, this definition can be implemented quite generally, and then used to create the procedures for all the desired elementary functions. The basic idea is to define a procedure that will combine the original function ϕ, the derivative ϕ', and the derivative structure a to compute $\phi(a)$. For the sake of discussion, let us call the procedure `Compose`. It will take as arguments procedures `phi` and `phi-prime`, and a derivative structure `a`. If `a` is actually just a real value, `Compose` applies `phi` to `a` and returns the result. Otherwise, `Compose` uses the V, D, and L operators to compute `a1`, `a2`, and `a1*`, respectively. Then it applies `phi` to `a1`, `phi-prime` to `a1*`, and returns the ordered pair `(phi(a1), phi-prime(a1*) * a2)`.

All of the elementary functions are defined in terms of the procedure `Compose`. For example, here is what the definition of the derivative structure exponential function might look like:

```
Function DS-Exp(a)
  if a is real
     return exp(a)
  else
     return Compose(DS-Exp, DS-Exp, a)
  end
```

Note that `DS-Exp` plays the role of both `phi` and `phi-prime` in the call to `Compose`. Thus, the computation of `DS-Exp(a)` requires evaluations of `DS-Exp(a1)` and `DS-Exp(a1*)`. This is simply a direct implementation of the recursive nature of Definition 4. In a similar way, the reciprocal function is defined as follows:

```
Function DS-Recip(a)
  if a is real
     return 1/a
  else
```

```
    return Compose(DS-Recip, DS-DRecip, a)
  end
```

Here, DS-DRecip is a derivative structure function that plays the role of the derivative of the reciprocal function. That is, with $\phi(x) = 1/x$, the derivative is $\phi'(x) = -1/x^2$. This can be defined by

```
Function DS-DRecip(a)
   recip-a = DS-Recip(a)
   return -1 * recip-a * recip-a
```

And now that we have defined the reciprocal function, it is no problem to add the natural logarithm.

```
Function DS-Ln(a)
  if a is real
     return ln(a)
  else
     return Compose(DS-Ln, DS-Recip, a)
  end
```

As these examples suggest, the development of a complete automatic differentiation system requires very little programming, once the derivative structure operations are in place. For each elementary function that is included, the developer does have to explicitly specify the derivative. However, that is a small price to pay for the automatic generation of derivatives to essentially arbitrary order. And in any case, one cannot reasonably hope to avoid defining derivatives altogether in a system that is supposed to compute derivatives automatically. In comparison to other approaches to automatic differentiation for higher derivatives [**2, 7**], the development presented here is remarkably simple.

This simplicity streamlines the task of implementing an automatic differentiation system. How the system performs is quite another issue, and it turns out that the elegance of the recursive approach is accompanied by some significant sources of inefficiency. While we will not take up this issue in any significant way here, a few brief comments are in order.

A little reflection reveals that a naive implementation of the doubly recursive approach involves widespread recomputation of previously obtained results. To illustrate this idea, consider the third derivative of the product fg. We know by Leibniz' rule that

$$(fg)''' = f''' + 3f''g' + 3f'g'' + g'''.$$

This can be derived by repeatedly applying the product rule, and then algebraically simplifying the result. In particular, three different terms, each equal to $f''g'$, would appear, giving rise to the single term $3f''g'$ in Leibniz' rule. The recursive automatic differentiation system is similar to repeatedly applying the product rule without algebraic simplification. That would entail three separate evaluations of $f''g'$.

In contrast, Neidinger [**9**] has developed a multivariate automatic differentiation system that uses explicit looping and subscripting. This system avoids the recomputation that can arise in the recursion process, and should be expected to outperform a direct implementation of the design presented here.

Inspired by Neidinger's approach, there are obvious strategies for reducing some of the recursive approach's inefficiency. In particular, a carefully optimized multiplication procedure, based on Leibniz' rule rather than simple recursion, might make a significant impact. Another attractive idea is to identify and exploit redundant calculations in the recursion process. Yet another improvement would be to take advantage of sparseness, eliminating computations that ultimately lead to multiplication by zero. Whether a modified version of the recursive system would be competitive with Neidinger's system is a question for further study.

However, no matter what formulation is used, direct computation of all partial derivatives of an expression is simply not the fastest approach. A more efficient alternative is to use systems of univariate automatic differentiation computations and an interpolation scheme [**1**]. Although this does increase memory requirements, it is easily shown to produce huge reductions in execution for large scale systems. Thus, for example, in a system with several hundred variables and a need for third order partial derivatives, any direct computation of all partial derivatives would be much slower than the alternative using interpolation.

On the other hand, computational speed is not always an issue. An automatic differentiation system of the type described here has been used successfully in an interactive application for analyzing systems of constraints arising in the design of satellite systems. In that context, automatic differentiation was used to perform sensitivity analyses among dozens of variables. For this application, computation was limited by the speed of user input, not by the speed with which the automatic differentiation system operated. In that situation, the speed of the automatic differentiation system was of no concern at all.

More generally, as computational speed continues to increase, the importance of execution efficiency will continue to decline, particularly for problems with small numbers of variables. In these cases, the directness and simplicity of the current development offers an attractive paradigm for implementing an automatic differentiation system.

Acknowledgment. This paper is based on an invited address at the January 1997 AMS/MAA meeting [**5**].

REFERENCES

1. C. Bischof, G. Corliss, and A. Griewank, Structured second- and higher-order derivatives through univariate Taylor series, preprint MCS-P296-0392, Argonne National Laboratory, Argonne, Illinois, May 1992.
2. H. Flanders, Automatic differentiation of composite functions, in *Automatic Differentiation of Algorithms: Theory, Implementation, and Application*, A. Griewank and G. F. Corliss, eds., SIAM, Philadelphia, 1991, pp. 95–99.
3. A. Griewank, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, SIAM, Philadelphia, 2000.
4. A. Griewank and G. F. Corliss, eds., *Automatic Differentiation of Algorithms: Theory, Implementation, and Application*, SIAM, Philadelphia, 1991.
5. D. Kalman, Automatic differentiation: computing derivative values without derivative formulas, Invited address, Joint Meetings of the American Mathematical Society and the Mathematical Association of America, San Diego, January 1997.
6. D. Kalman and R. Lindell, Recursive multivariate automatic differentiation, *Optimization Methods and Software* **6** (1995), 161–192.
7. C. L. Lawson, Automatic differentiation of inverse functions, in *Automatic Differentiation of Algorithms: Theory, Implementation, and Application*, A. Griewank and G. F. Corliss, eds., SIAM, Philadelphia, 1991, pp. 87–94.
8. R. D. Neidinger, Automatic differentiation and APL, *College Math. J.* **20** (1989), 238–251.
9. R. D. Neidinger, An efficient method for the numerical evaluation of partial derivatives of arbitrary order, *ACM Transactions on Mathematical Software* **18** (1992), 159–173.
10. L. B. Rall, The arithmetic of differentiation, this MAGAZINE **59** (1986), 275–282.

NOTES

Avoiding Your Spouse at a Party Leads to War

MARC A. BRODIE
The College of St. Benedict
St. Joseph, MN 56374
mbrodie@csbsju.edu

In her note in the February 2001 issue of this MAGAZINE, "Avoiding Your Spouse at a Bridge Party," Barbara H. Margolius [**4**] addresses the Bridge Couples Problem, a question originally considered in terms of dancing by James Brawner a year earlier ("Dinner, Dancing and Tennis, Anyone?" this MAGAZINE, February 2000 [**2**]):

THE BRIDGE COUPLES PROBLEM. Suppose n married couples ($2n$ people) are invited to a bridge party. Bridge partners are chosen at random, without regard to gender. What is the probability that no one will be paired with his or her spouse?

I see no reason to avoid only my spouse. In addition to a wife, I have two children, and we would all like to avoid each other. Presumably, there are other families of four in this situation. Thus we will consider the following problem:

THE BRIDGE FAMILIES PROBLEM. Suppose n families of four ($4n$ people) are invited to a bridge party. Bridge partners are chosen at random, without regard to gender or generation. What is the probability that no one will be paired with a member of his or her family?

Note that the Bridge Families Problem is different from the generalization of the Bridge Couples Problem that Margolius offered as an exercise at the end of her article.

The Bridge Families Problem can also be interpreted as a problem arising in the playing of the card game War. War is played by thoroughly shuffling a standard deck of 52 playing cards and dealing 26 cards to each of two players. The players then compare the top cards in their hands, with both of those cards going to the player with the higher ranking card. Play continues in this fashion until each player has played all of his or her 26 cards. In the case of a match (sometimes called a "battle"), that is both players turning over cards of the same rank, additional rules exist concerning how to proceed. However, we will not be concerned with that issue here. Neither will we consider plays beyond the first 26, though in the traditional game of War, play generally extends well beyond going through the deck once. In terms of War, the Bridge Families Problem becomes:

THE PROBLEM OF WAR WITHOUT BATTLES. Suppose a well-shuffled deck consisting of $4n$ cards (4 distinguishable cards of each of n linearly ordered ranks) is dealt to two players so that each player has $2n$ cards. What is the probability that when the players play through their decks and compare the cards, there are no matches?

We choose to use the language of cards over that of families because cards in a deck come with an obvious ordering by rank. This ordering allows for many additional interesting questions, some of which have been investigated by two of my students in their undergraduate research projects and honors theses. Here we will answer the following question, which is closely related to the Problem of War Without Battles:

THE ANNIHILATION PROBLEM. What is the probability that in a game of War with a deck of $4n$ cards, player 1 annihilates player 2 (that is, has a higher ranking card in all $2n$ plays)?

The Annihilation Problem is of particular interest to me since I was once almost annihilated by one of my children (he won all but the last card)! As we shall later see, the probability of being annihilated when using a standard deck of cards, and the probability of being almost annihilated in this fashion, are both less than $\frac{1}{300,000,000}$. Thus, I am left wondering just who shuffled the deck before that ill-fated game.

Computing the probability of k matches We envision the deal of the cards as a $2 \times 2n$ rectangular array, with the cards of player 1 in the first row, and those of player 2 in the second (see FIGURE 1). A match then corresponds to two cards of the same rank in the same column. Thus in FIGURE 1 there are four matches shown, J$\diamondsuit$ and J$\clubsuit$, 3$\heartsuit$ and 3$\spadesuit$, J$\spadesuit$ and J$\heartsuit$, and 8$\clubsuit$ and 8$\spadesuit$. Note that there are two matches involving jacks; we will call such a situation a *double match.* The presence of the (unmatched) 3$\clubsuit$ and 8$\diamondsuit$, indicates that the 3s and 8s are *single matches.*

3♣	5♠	J♢	7♠	3♡	Q♡	K♡	4♣	J♠	5♢	6♣	2♡	A♠	8♣	⋯	4♠
7♢	8♢	J♣	Q♣	3♠	6♠	5♡	2♣	J♡	7♣	2♢	9♣	9♡	8♠	⋯	K♢

Figure 1 A possible deal of the cards for a standard deck ($n = 13$)

It is the potential for double matches that makes computing the probability of k matches an interesting generalization of the Bridge Couples Problem. Since there are $4n$ distinguishable cards being placed into $4n$ positions in the array, there are $(4n)!$ possible deals. Let k be a fixed integer, $1 \leq k \leq 2n$. We will compute the probability of a deal with at least k matches, then use the inclusion-exclusion principle to compute the probability of a deal with exactly k matches.

If we have k matches, some may be doubles. We let m be the number of double matches, and r be the number of single matches. Thus $2m + r = k$, and we must consider all *possible* values of m from $m = 0$ to $m = \lfloor k/2 \rfloor$. Note, for example, that in a standard deck of 52 cards (where $n = 13$), if we let $k = 20$, we will need at least seven double matches. In general, the smallest possible value of m will be the maximum of the set $\{0, k - n\}$. To count the number of deals with at least k matches, we first observe that there are $\binom{2n}{k}$ ways to specify which positions contain the matches. We then select the r ranks for single matches, choose two of each of those four cards for the match, and choose which of those cards goes to each player. This stage results in

$$\binom{n}{r}\binom{4}{2}^r \cdot 2^r = \binom{n}{r} \cdot 12^r \tag{1}$$

possibilities. Next, from the remaining $n - r$ ranks we choose the m ranks involved in the double matches, split the four cards of each of those ranks into an unordered pair of unordered pairs, and choose which card goes to which player in each of the resulting $2m$ pairs. This next stage of our process results in

$$\binom{n-r}{m}\left(\binom{4}{2}/2\right)^m \cdot 2^{2m} = \binom{n-r}{m} \cdot 12^m \tag{2}$$

possibilities. Finally, the k pairs we have chosen can be put into the k specified positions in $k!$ ways, and the remaining $4n-2k$ cards can be put into the array in $(4n-2k)!$ ways. Thus, for a fixed m, the number of deals that have at least k matches *in specified positions* is the product of $k!\,(4n-2k)!$ and formulas (1) and (2):

$$\binom{n}{r}\binom{n-r}{m}\cdot 12^{m+r}\cdot k!\,(4n-2k)! = \frac{12^{k-m}\cdot k!\,n!\,(4n-2k)!}{m!\,(k-2m)!\,(n-k+m)!}. \tag{3}$$

Thus the *probability* of a deal with at least k matches in specified positions is given by

$$\sum_{m=\max\{0,k-n\}}^{\lfloor\frac{k}{2}\rfloor}\frac{12^{k-m}\cdot k!\,n!\,(4n-2k)!}{m!\,(k-2m)!\,(n-k+m)!\,(4n)!}$$

$$= \frac{12^{k}\cdot k!\,n!\,(4n-2k)!}{(4n)!}\cdot\sum_{m=\max\{0,k-n\}}^{\lfloor\frac{k}{2}\rfloor}\frac{1}{12^{m}\cdot m!\,(k-2m)!\,(n-k+m)!}.$$

We are now in a position to solve the Problem of War Without Battles. We will find the probability of no matches, $p_{0,n}$, by computing the probability of the complement, at least one match. Keep in mind that, for example, the event of a match in the first position, and that of a match in the second position are not disjoint. Thus we cannot find the probability of at least one match by simply evaluating our formula above when $k=1$. Instead, we use the inclusion-exclusion principle; we include all the nondisjoint possibilities with one match, then subtract off the overlapping possibilities with two matches, and so on. Since there are $\binom{2n}{k}$ ways to specify the k positions in which there are matches, inclusion-exclusion tells us that the probability of at least one match is $p_{1^*,n}=$

$$\sum_{k=1}^{2n}\left((-1)^{k+1}\binom{2n}{k}\frac{12^{k}\cdot k!\,n!\,(4n-2k)!}{(4n!)}\cdot\sum_{m=\max\{0,k-n\}}^{\lfloor\frac{k}{2}\rfloor}\frac{1}{12^{m}\cdot m!\,(k-2m)!\,(n-k+m)!}\right). \tag{4}$$

(The asterisk in $p_{1^*,n}$ signifies that this is the probability of *at least* one match.) Hence, the solution to the Problem of War Without Battles is $p_{0,n}=1-p_{1^*,n}$.

At this point, we pause to answer the Annihilation Problem. Since the Annihilation Problem does not consider situations involving battles, we need only determine the probability that player 1 has a higher ranking card than player 2 does in each of the $2n$ plays, *given that there are no matches.* Since all deals are assumed to be equally likely, this probability is quickly seen to be $1/2^{2n}$, so that the probability of annihilation is $p_{0,n}/2^{2n}$. When $n=13$, we compute that for a standard deck of cards, the probability of no battles is approximately 0.210214, and the probability of my annihilation is $0.210214/2^{26}\approx 3.13243\times 10^{-9}$. Note that under the same assumptions, the probability that player 1 has a higher ranking card in all but the last of the $2n$ plays is also $1/2^{2n}$. Thus the probability of my almost being annihilated in this fashion is also approximately 3.13243×10^{-9}.

Using more general Inclusion-Exclusion formulas (see [**3**, Chapter IV]), we now compute the probability of exactly j matches, for any j: $p_{j,n}=$

$$\left|\sum_{k=j}^{2n}\left((-1)^{k+1}\binom{k}{j}\frac{12^{k}\cdot n!\,(2n)!\,(4n-2k)!}{(2n-k)!\,(4n)!}\cdot\sum_{m=\max\{0,k-n\}}^{\lfloor\frac{k}{2}\rfloor}\frac{1}{12^{m}\cdot m!\,(k-2m)!\,(n-k+m)!}\right)\right|. \tag{5}$$

As a reality check, we had *Mathematica* use (5) to compute the probability distribution for a standard deck of cards, some of which is shown in TABLE 1.

TABLE 1: The probability distribution for the number of matches using a standard deck of cards.

j	$p_{j,13}$	j	$p_{j,13}$
0	0.210214	7	0.000595043
1	0.334183	8	0.0000946984
2	0.259336	9	0.0000128083
3	0.130713	$\vdots$	$\vdots$
4	0.048028	24	3.33819×10^{-25}
5	0.0136839	25	0
6	0.00314026	26	5.34967×10^{-28}

Given the relative complexity of the formula for $p_{j,13}$, it was particularly pleasing to find that $p_{25,13} = 0$, which of course must be the case since it is impossible to have 25 matches without the two remaining cards matching also.

The asymptotic behavior of the probability of war without battles In Margolius' article [**4**], it was shown that the probability that no one will be paired with his or her spouse converges to $e^{-1/2}$. Here we investigate the convergence (as $n \to \infty$) of the probability, $p_{0,n}$, of no battles in a game of War with a deck of $4n$ cards. In TABLE 2, values of $p_{0,n}$, as computed by *Mathematica* for increasing n, are displayed.

TABLE 2: The probability of zero matches in a game of War with different sized decks of cards.

n	$p_{0,n}$	n	$p_{0,n}$
13	0.210214	50	0.219779
20	0.214742	100	0.221456
30	0.217542	500	0.222795
40	0.218941	1000	0.222963

Note that although convergence is slow, appearances certainly indicate that the sequence is increasing, and in this case, since the sequence is bounded above by 1, a limit must exist. The interested reader might wish to make a conjecture as to the limit before reading further. (Hint: the limit involves e.)

THEOREM. *The limit of the probability of no battles in a game of War with cards of n different ranks is*

$$\lim_{n\to\infty} p_{0,n} = e^{-3/2} \approx 0.223130.$$

Proof. The presence of the double summation and the floor function in formula (4) makes finding the limit of the $p_{0,n}$ directly from that formula somewhat difficult, so we take a different approach. We will first compute the probability, $q_{1^*,n}$, of having at least one *double* match. Using (3), with $k = 2$ and $m = 1$, choosing two positions to contain the double match, and noting that we're overcounting, we find that

$$q_{1^*,n} \le \binom{2n}{2} \frac{12 \cdot 2 \cdot n!\,(4n-4)!}{(n-1)!\,(4n)!} = \frac{6n(2n-1)}{(4n-1)(4n-2)(4n-3)}.$$

Therefore,

$$\lim_{n\to\infty} q_{1*,n} = 0.$$

Since the probability of a double match goes to 0, we may evaluate the limit of the probability of no battles in a game of War by using (4) with the inner sum evaluated only at $m = 0$. Doing so yields

$$p_{1*,n} = \sum_{k=1}^{2n} \frac{(-1)^{k+1}12^k}{k!} \cdot \frac{n!\,(2n)!\,(4n-2k)!}{(n-k)!\,(2n-k)!\,(4n)!}.$$

A formal argument using epsilonics (similar to that used by Margolius [**4**]) now shows that

$$\lim_{n\to\infty} p_{1*,n} = 1 - e^{-3/2}.$$

However, we choose to note that one can see this result more informally by observing that

$$\frac{n!\,(2n)!\,(4n-2k)!}{(n-k)!\,(2n-k)!\,(4n)!} = \frac{n(n-1)\cdots(n-k+1)\cdot 2n(2n-1)\cdots(2n-k+1)}{4n(4n-1)\cdots(4n-2k+1)},$$

so that for fixed k,

$$\lim_{n\to\infty} \frac{n!\,(2n)!\,(4n-2k)!}{(n-k)!\,(2n-k)!\,(4n)!} = \frac{n^k(2n)^k}{(4n)^{2k}} = \frac{1}{8^k}.$$

Thus,

$$\lim_{n\to\infty} p_{1*,n} = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}12^k}{k!} \cdot \frac{1}{8^k} = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}(3/2)^k}{k!} = 1 - e^{-3/2}.$$

Suggestions for further investigation In [**1**], Blom, Holst, and Sandell consider (among other problems) the "matching Sing-Sing problem." This problem is equivalent to the Bridge Couples Problem. Blom, Holst, and Sandell prove that as $n \to \infty$, the probability distribution of the number of matches approaches a Poisson distribution with parameter $\lambda = 1/2$. Given the result of our theorem, the obvious conjecture is that in the Problem of War Without Battles, the probability distribution of the number of matches approaches a Poisson distribution with parameter $\lambda = 3/2$. Is this the case? Numerical evidence at least indicates this is plausible. For example, in TABLE 3,

TABLE 3: A comparison of the probability distributions of the number of matches for $n = 13$ with the Poisson distribution with parameter $\lambda = 3/2$.

j	$p_{j,13}$	$P(j;3/2)$	j	$p_{j,13}$	$P(j;3/2)$
0	0.210214	0.223130	7	0.000595043	0.000756426
1	0.334183	0.334695	8	0.0000946984	0.000141830
2	0.259336	0.251021	9	0.0000128083	0.0000236383
3	0.130713	0.125511	$\vdots$	$\vdots$	$\vdots$
4	0.048028	0.0470665	24	3.33819×10^{-25}	6.05401×10^{-21}
5	0.0136839	0.0141200	25	0	3.63240×10^{-22}
6	0.00314026	0.00352999	26	5.34967×10^{-28}	2.09562×10^{-23}

we compare the probability distribution for the number of matches playing with a standard deck of cards (from TABLE 2) with the probabilities from the Poisson distribution $P(j; 3/2)$.

In the annihilation problem, we considered only the case that player 1 had a higher ranking card than player 2 in each play. What if we allow matches, and adopt the convention that if there is a match, the winner of the next play takes all four cards? What if we adopt the standard convention that in the event of a match, each players' next three cards are placed face down, and their fourth cards are compared, with the winner taking all ten cards involved? What are the probabilities of annihilation in these cases? Here, "annihilation" means player 2 doesn't win any cards during the course of play. Numerous generalizations can be investigated. What if we play with a deck consisting of six cards per rank? What about m cards per rank? What if m is odd? Finally, what if, as often seems to be the case in my family, we're not playing with a full deck?

Acknowledgment. The author would like to thank the two anonymous referees for their comments. This paper was substantially improved by their suggestions.

REFERENCES

1. Gunnar Blom, Lars Holst, and Dennis Sandell, Three Sing-Sing problems, *Amer. Math. Monthly* **102** (1995), 887.
2. J. N. Brawner, Dinner, dancing and tennis, anyone?, this MAGAZINE **73** (2000), 29–35.
3. W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed., revised printing, John Wiley and Sons, New York, NY, 1970.
4. Barbara H. Margolius, Avoiding your spouse at a bridge party, this MAGAZINE **74** (2001), 33–41.

Bernoulli on Arc Length

VICTOR MOLL
Tulane University
New Orleans, Louisiana 70148
vhm@math.tulane.edu

JUDITH NOWALSKY
University of New Orleans
New Orleans, Louisiana 70148
jnowalsk@math.uno.edu

GINED ROA
LEONARDO SOLANILLA
Universidad del Tolima
Ibagué, Colombia
gined@latinmail.com
solanila@bunde.tolinet.com.co

The academic life of the Bernoulli family was always surrounded by controversy. The disputes between Johann (John) and his older brother and former teacher Jacob and with his son Daniel are famous and well documented. An interesting discussion of this remarkable family is found in Section 12.6 of [**3**]. After the death of L'Hôpital, John claimed the authorship of his classical analysis book. In the controversy between Leibniz and Newton about the creation of calculus, he stood on Leibniz' side. His controversial positions were not restricted to mathematics: he was even accused of denying the possibility of the resurrection of Christ.

In the course of our study of the history of elliptic integrals, we found a paper by Johann Bernoulli [**1**] which, in our opinion, both illuminates the calculation of arc lengths of smooth curves, a topic covered in most undergraduate calculus programs around the world, and provides an additional tool for producing new and interesting examples of *rectifiable curves*. According to Bernoulli, these are curves whose arc length can be expressed as elementary functions of their end points. The paper contains a main theorem that is perfectly valid even today, and admits a nice interpretation in terms of the notion of radius of curvature. Furthermore, we discovered in it a colorful antecedent of Landen integral transformations [**2**].

Let $y = y(x)$ be a differentiable function defined on $[a, b]$. Then its *arc length* is defined by

$$g(x) = \int_a^x \sqrt{1 + \left(\frac{dy}{d\xi}\right)^2}\, d\xi.$$

In general, this integral is not trivial. The examples and exercises provided in most textbooks look *unnatural*: for instance, the first example given in Thomas [**4**], page 395, deals with the arc length of the curve

$$y = \frac{4\sqrt{2}}{3}x^{3/2} - 1$$

for $0 \le x \le 1$. This is an easy example in the sense that it is computable:

$$g(1) = \int_0^1 \sqrt{1 + 8\xi}\, d\xi = \frac{13}{6}.$$

The reader can easily verify that the integral corresponding to the length of a circle can be evaluated. However, the calculation of the arc length of an ellipse leads to the integral

$$L(x) = a\int_0^x \sqrt{\frac{1-e^2\xi^2}{1-\xi^2}}\,d\xi,$$

where a is the semimajor axis of the ellipse, and e its eccentricity. This last integral is one of the fundamental *elliptic integrals* and is not an elementary function. It was the starting point of our research on Bernoulli's work.

Bernoulli's *universal theorem* The main goal of this section is to present Bernoulli's result on how to produce rectifiable curves, which in this sense might also be called *rectifiable by straight lines*; their arc length can be expressed as elementary functions of their end points.

THEOREM 1. *Let $y = y(x)$ be a twice differentiable function satisfying*

$$\left(\frac{dy}{dx}\right)^2 + 3x\frac{dy}{dx}\frac{d^2y}{dx^2} \geq 0 \quad (\leq 0)$$

in its interval of definition $[a, b]$. Define a new curve with coordinates

$$X = x\left(\frac{dy}{dx}\right)^3, \quad Y = \frac{3x}{2}\left(\frac{dy}{dx}\right)^2 - \frac{1}{2}\int_a^x \left(\frac{dy}{d\xi}\right)^2 d\xi.$$

Now let $g(x)$ and $G(x)$ be the arc lengths of y and the parametric curve $(X(x), Y(x))$ starting at $x = a$. Then

$$g(x) + (-)\, G(x) = \xi\left(\frac{dg}{d\xi}\right)^3\Bigg]_{\xi=a}^{\xi=x}$$

for all $x \in [a, b]$.

Proof. First observe that

$$\left(\frac{dg}{dx}\right)^3 = \left[1+\left(\frac{dy}{dx}\right)^2\right]^{3/2}.$$

Following Bernoulli's recommendation, we compute

$$\frac{d}{dx}x\left(\frac{dg}{dx}\right)^3 = \left[1+\left(\frac{dy}{dx}\right)^2\right]^{3/2} + 3x\left[1+\left(\frac{dy}{dx}\right)^2\right]^{1/2}\frac{dy}{dx}\frac{d^2y}{dx^2}.$$

On the other hand, careful differentiation shows that

$$\begin{aligned} G(x) &= \int_a^x \sqrt{\left(\frac{dX}{d\xi}\right)^2 + \left(\frac{dY}{d\xi}\right)^2}\,d\xi \\ &= \int_a^x \sqrt{1+\left(\frac{dy}{d\xi}\right)^2}\left|\left(\frac{dy}{d\xi}\right)^2 + 3x\frac{dy}{d\xi}\frac{d^2y}{d\xi^2}\right| d\xi. \end{aligned}$$

To conclude the proof, note that the integrand of $g(x) + (-)\, G(x)$ is $\frac{d}{dx}x(\frac{dg}{dx})^3$, so the result follows from the Fundamental Theorem of Calculus. ∎

For example, the function $y = \ln x$ yields

$$X = \frac{1}{x^2}, \quad Y = \frac{2}{x^2} - \frac{1}{2} = 2\sqrt{X} - \frac{1}{2}.$$

After struggling to get the correct constants in some assertions in Bernoulli's article, we discovered a nice interpretation of Theorem 1. This formulation eluded Bernoulli as he did not relate the result to the curvature of the graph $y = y(x)$. Recall that the radius of curvature of the graph $y = y(x)$ at a point x is

$$R(x) = \left[1 + \left(\frac{dy}{dx}\right)^2\right]^{3/2} \times \left(\frac{d^2y}{dx^2}\right)^{-1}.$$

This is the radius of a circle whose curvature matches that of the curve at the given point. Let us restate the previous theorem in terms of curvature.

THEOREM 2. *Under the assumptions of Theorem 1,*

$$g(x) \pm G(x) = \xi \frac{d^2y}{d\xi^2} R(\xi)\Big]_{\xi=a}^{\xi=x}.$$

Geometrically, denote respectively by $C(a), C(x)$ the centers of the osculating circles at the points $A(a) = (a, y(a)), A(x) = (x, y(x))$ on the curve. Also, let $\alpha(a) = \angle BAC(a), \alpha(x) = \angle BAC(x)$ be the corresponding angles between the radii of curvature $R(a) = CA(a)$, $R(x) = CA(x)$ at these points and the hypotenuses of some right triangles $ABC(a)$, $ABC(x)$ as sketched as in the figure below. The positions of points $B(a)$ and $B(x)$ along the rays shown are determined by the angles $\alpha(a)$ and $\alpha(x)$, respectively. Then

$$CB(a) = R(a) \tan \alpha(a), \quad CB(x) = R(x) \tan \alpha(x).$$

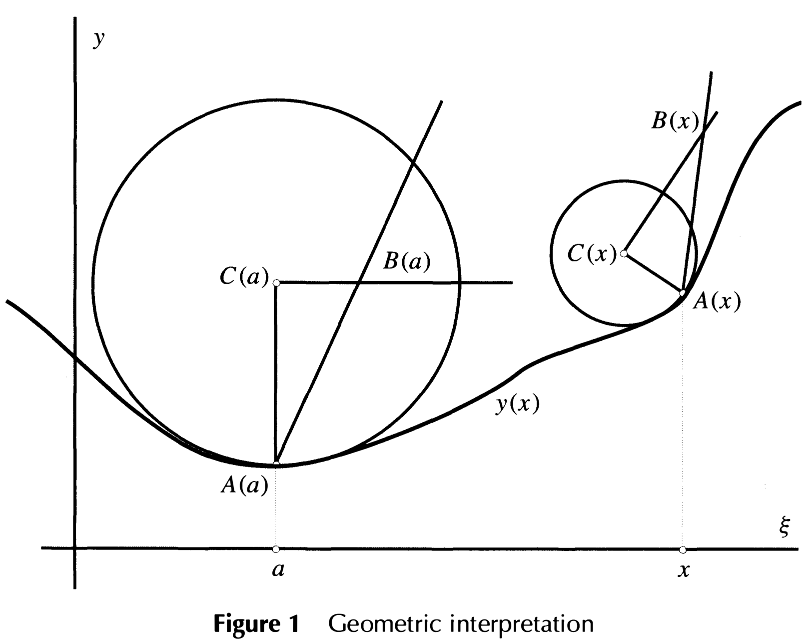


Figure 1 Geometric interpretation

In order to get the rectification, set

$$\tan\alpha(a) = a\frac{d^2y}{d\xi^2}\bigg]_{\xi=a}, \qquad \tan\alpha(x) = x\frac{d^2y}{d\xi^2}\bigg]_{\xi=x}.$$

In this way,

$$g(x) \pm G(x) = CB(x) - CB(a)$$

gives a new meaning to Theorem 2: the sum (difference) of the arc length integrals equals the difference of two straight segments. Bernoulli was proud to declare that this sum (difference) could be measured on a *straight line.*

Parabolas In Bernoulli's language, a *parabola* is a curve defined by the function $y = x^q$, for q a rational number. In this section we discuss parabolas that are rectifiable by the above method. Remember that a curve $y = y(x)$ is rectifiable if its arc length integral admits an antiderivative in terms of elementary functions. Bernoulli was interested in the question of rectifiable parabolas and was aware of the following result.

THEOREM 3. *Let n be a nonzero integer. Then the parabola $y = x^{\frac{2n+1}{2n}}$ is rectifiable on* $[0, 1]$.

Proof. The arc length is

$$g(x) = \int_0^x \sqrt{1 + \left(\frac{2n+1}{2n}\right)^2 \xi^{1/n}}\, d\xi,$$

and the substitution $u(\xi) = 1 + (\frac{2n+1}{2n})^2\xi^{1/n}$ yields

$$g(x) = n\left(\frac{2n}{2n+1}\right)^{2n} \int_1^{u(x)} \sqrt{u}(u-1)^{n-1}\, du,$$

which can be evaluated by expanding $(u-1)^{n-1}$ using the binomial theorem. ∎

The reader may recognize that this result is the source of most arc length exercises in textbooks. Our first example corresponds to $n = 1$. Moreover, the presence of the factor $4\sqrt{2}/3$ is not essential to the solution of the problem: it is window dressing.

We can now use Theorem 1 to assert that every parabola can be rectified by adding the arc length of another (conveniently chosen) parabola.

THEOREM 4. *Any parabola $y = x^q$, $q \neq 2/3$, can be rectified by adding (subtracting) to its arc length the arc length of the auxiliary parabola*

$$Y = \frac{3q-2}{2q-1} q^{\frac{1}{2-3q}} X^{\frac{2q-1}{3q-2}},$$

where $X = q^3x^{3q-2}$. In particular, the usual quadratic (Archimedean) parabola $y = x^2$ is rectified by adding the arc length of the biquadratic-cubic parabola $Y = \frac{4}{3}2^{-1/4}X^{3/4}$.

Proof. Note that

$$\left(\frac{dy}{dx}\right)^2 + 3x\frac{dy}{dx}\frac{d^2y}{dx^2} = (3q-2)q^2x^{2(q-2)}.$$

The rest of the proof is a straightforward calculation. ∎

Integral transformations Many interesting questions can be formulated at this point. For instance: *Under what circumstances does the degree of the auxiliary parabola equal the degree of the original parabola?* The answer is clearly given by the fixed points of the rational transformation $b(q) = (2q-1)/(3q-2)$, $q \neq 2/3$, namely, $q = 1, 1/3$. Since the first value gives a trivial answer, Bernoulli considered only the second value, which corresponds to the *primary cubic* parabola $y = x^{1/3}$. This case is important because it yields $Y = X^{1/3}$, and consequently

$$g(x) - G(x) = \int_0^x \sqrt{1+\frac{1}{9\xi^{4/3}}}\,d\xi - \int_0^{\frac{1}{27x}} \sqrt{1+\frac{1}{9X^{4/3}}}\,dX$$
$$= \int_{\frac{1}{27x}}^x \sqrt{1+\frac{1}{9\xi^{4/3}}}\,d\xi = x\left(1+\frac{1}{9x^{4/3}}\right)^{3/2},$$

an actual arc length integral formula!

It is interesting to study the sequence defined recursively by $p_0 = q$, $p_n = bp_{n-1}$, $n = 1, 2, \ldots$, for a given starting value q. For example, if $q = 2$, then $p_1 = 3/4$ and $p_2 = 2$ (again). This implies that if the original parabola is the usual $y = x^2$, for which $X = 8x^7$ and $Y = \frac{4}{3}2^{-1/4}X^{3/4}$, then

$$\int_0^x \sqrt{1+4\xi^2}\,d\xi + \int_0^{8x^7} \sqrt{1+\frac{1}{\sqrt{2X}}}\,dX = x(1+4x^2)^{3/2}.$$

But applying the transformation from Theorem 1 once more to the auxiliary parabola, we obtain $\mathcal{X} = 2^{-3/4}X^{1/4} = x^{7/4}$ and $\mathcal{Y} = 2^{-3/2}\sqrt{X} = \mathcal{X}^2$. Thus

$$\int_0^{8x^7} \sqrt{1+\frac{1}{\sqrt{2X}}}\,dX + \int_0^{x^{7/4}} \sqrt{1+4\mathcal{X}^2}\,d\mathcal{X} = 8x^7\left(1+\frac{1}{4x^{7/2}}\right)^{3/2},$$

from which

$$\int_x^{x^{7/4}} \sqrt{1+4\xi^2}\,d\xi = 8x^7\left(1+\frac{1}{4x^{7/2}}\right)^{3/2} - x(1+4x^2)^{3/2}.$$

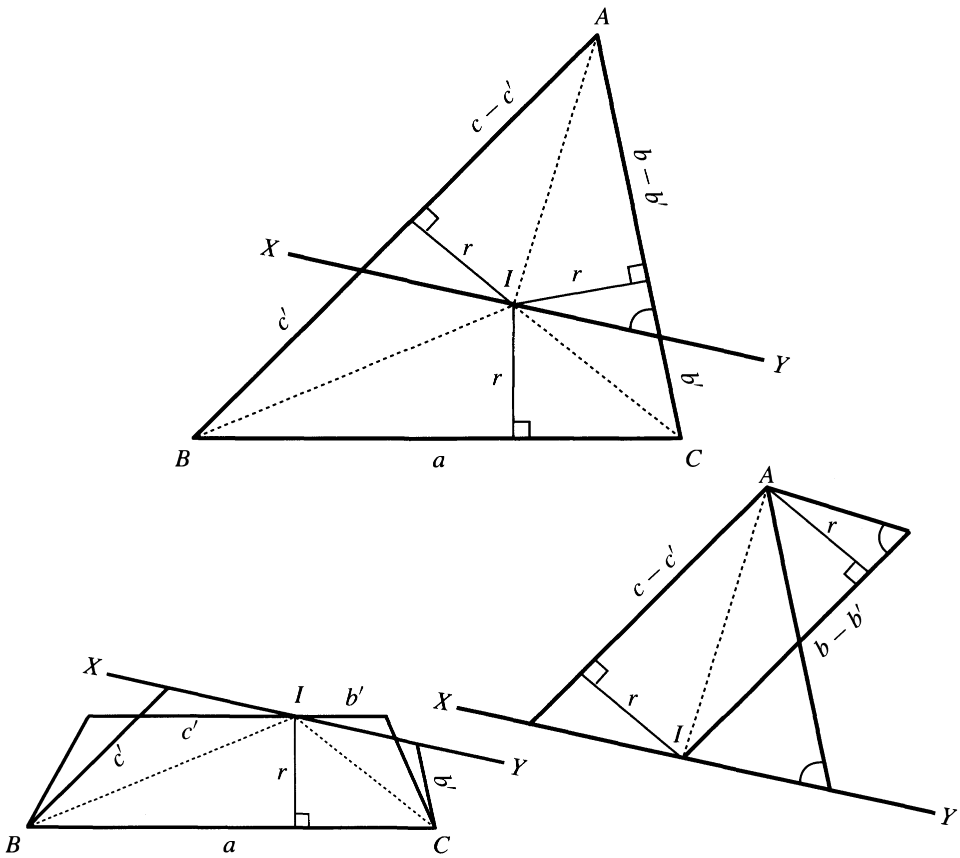
Finally, we may ask these questions: *How many different values can a sequence p_n take and still lead to an arc length integral formula? Is there any relation between the convergence of this type of sequence and new arc length integral formulas?*

REFERENCES

1. J. Bernoulli, Theorema universale rectificationi linearum curvarum inserviens. Nova Parabolarum propietas. Cubicalis primaria arcum mensura, etc., *Opera Omnia I*4, Georg Olms Verlagbuchhandlung, Hildesheim, 1968, pp. 249–25. English translation: Universal Theorem useful for the rectification of curves by J. L. Nowalsky, G. Roa, L. Solanilla, preprint, 1999.
2. V. Moll, J. Nowalsky, and L. Solanilla, The story of Landen, the hyperbola and the ellipse, *Elem. Math.* **57** (2002), 1–7.
3. J. Stillwell, *Mathematics and its History*, Undergraduate Texts in Mathematics, Springer Verlag, 1989.
4. G. Thomas and R. Finney, *Calculus and Analytic Geometry*, 9th ed., Addison Wesley, 1996.

Proof Without Words: A Line through the Incenter of a Triangle

A line passing through the incenter of a triangle bisects the perimeter if and only if it bisects the area.



$$A_{\text{top}} = \frac{(b - b' + c - c')r}{2} \qquad A_{\text{bottom}} = \frac{(a + b' + c')r}{2}$$

$$A_{\text{top}} = A_{\text{bottom}} \quad \Leftrightarrow \quad a + b' + c' = \frac{a + b + c}{2}$$

——Sidney H. Kung
University of North Florida
Jacksonville, FL 32224

Finite Groups That Have Exactly n Elements of Order n

CARRIE E. FINCH
University of South Carolina
Columbia, SC 29208

RICHARD M. FOOTE
University of Vermont
Burlington, VT 05405

LENNY JONES
Shippensburg University
Shippensburg, PA 17257

DONALD SPICKLER, JR.
Salisbury State University
Salisbury, MD 21801

Several years ago one of the authors placed the following rather innocuous question on a group theory exam: Can a finite group have exactly two elements of order two? While the correct answer of "no" can be proven fairly easily by a variety of techniques, depending on the sophistication of the solver, the authors discovered that the generalization from "two" to n is not as painless, nor is the answer always negative. In this note we investigate the answer to the question: For which n do there exist finite groups that have exactly n elements of order n? We present the details of the solution in the abelian case while only stating the result in the nonabelian case.

Preliminaries For $n = 1$, the question is easily answered since all finite groups contain exactly one element of order one, namely the identity element. So assume that $n > 1$ and write $n = \prod_{i=1}^{m} p_i^{a_i}$ where the p_i are distinct primes. Suppose that G has exactly n elements of order n.

Define an equivalence relation on the set of elements of order n by saying that elements x and y are related if and only if they generate the same cyclic subgroup. The number, k, of equivalence classes is the number of distinct cyclic subgroups of G of order n. Each such subgroup of G contains exactly $\phi(n) = \prod_{i=1}^{m} p_i^{a_i-1}(p_i - 1)$ elements of order n, where ϕ is Euler's totient function. This means that the number of distinct cyclic subgroups of G of order n is

$$k = \frac{n}{\phi(n)} = \prod_{i=1}^{m} \frac{p_i}{p_i - 1}.$$

Consequently, $m = 1$ or $m = 2$. If $m = 1$, then $p_1 = 2$. If $m = 2$, then $p_1 = 2$ and $p_2 = 3$. Hence we have the following:

PROPOSITION 1. *Let $n > 1$. Then a finite group G has exactly n elements of order n if and only if*

either $n = 2^a$, and G has exactly two cyclic subgroups of order 2^a, $a \geq 1$
or $n = 2^a 3^b$, and G has exactly three cyclic subgroups of order $2^a 3^b$, $a, b \geq 1$.

Although Proposition 1 imposes severe restrictions on the structure of a group that has exactly n elements of order n, we can, nevertheless, construct infinitely many such groups.

PROPOSITION 2. *If a finite group G has exactly n elements of order n and H is a finite group such that its order is relatively prime to the order of G, then $G \times H$ also has exactly n elements of order n.*

For example, it is easy to check that $G = \mathbb{Z}_2 \times \mathbb{Z}_4$ has exactly 4 elements of order 4; and if H is any finite group such that the order of H is odd, then $G \times H$ also has exactly 4 elements of order 4. Therefore, we have

COROLLARY. There exist infinitely many groups that contain exactly n elements of order n.

In light of Proposition 2 the following definition seems natural.

DEFINITION. Let G be a finite group with exactly n elements of order n. We call G **minimal** if no proper subgroup of G has exactly n elements of order n.

The abelian case Suppose that G is an abelian minimal finite group with exactly n elements of order n. From the preceding section we have that $n = 2^a 3^b$ with $a \geq 1$ and $b \geq 0$. By the Fundamental Theorem of Finitely Generated Abelian Groups, G is isomorphic to the direct product of subgroups of prime power order, namely its Sylow p-subgroups, where p runs over the distinct primes dividing the order of G [**1**, Theorem 5, Section 5.2]. Grouping all the direct factors for primes greater than 3 together into a subgroup G_m, we obtain that $G \cong G_2 \times G_3 \times G_m$, where G_2 and G_3 are the Sylow 2- and 3-subgroups respectively. Since the order of an element in a direct product is the least common multiple of the orders of the elements in its components, it follows easily that $G_2 \times G_3$ contains all elements in G of order n (more generally, one can deduce this from Exercise 18, Section 3.2 of [**1**], using $N = G_2 \times G_3$ and H any cyclic subgroup of order n). Therefore, by Proposition 2, since G is minimal, we must have $G_m = 1$, that is, $G \cong G_2 \times G_3$; furthermore, by the Fundamental Theorem, the Sylow subgroups G_2 and G_3 decompose further into direct products of cyclic groups as:

$$G_2 \cong \underbrace{\mathbb{Z}_2 \times \cdots \times \mathbb{Z}_2}_{s_1 \text{ factors}} \times \underbrace{\mathbb{Z}_{2^2} \times \cdots \times \mathbb{Z}_{2^2}}_{s_2 \text{ factors}} \times \cdots \times \underbrace{\mathbb{Z}_{2^a} \times \cdots \times \mathbb{Z}_{2^a}}_{s_a \text{ factors}}$$

$$\text{and} \quad G_3 \cong \underbrace{\mathbb{Z}_3 \times \cdots \times \mathbb{Z}_3}_{t_1 \text{ factors}} \times \underbrace{\mathbb{Z}_{3^2} \times \cdots \times \mathbb{Z}_{3^2}}_{t_2 \text{ factors}} \times \cdots \times \underbrace{\mathbb{Z}_{3^b} \times \cdots \times \mathbb{Z}_{3^b}}_{t_b \text{ factors}}$$

with G_3 possibly being trivial. In the situation when G_3 is trivial, $n = 2^a$ and $G \cong G_2$. When $a = 1$, every nonidentity element in G has order two, and hence the number of elements of order 2^a in G is $2^{s_a} - 1$, which is obviously not equal to 2^a. For $a > 1$, one way to count the number of elements of order 2^a is to count the total number of elements in G and then subtract the number of elements of order less than 2^a. Now, in the group $\mathbb{Z}_{2^a}$, there are $\phi(2^a) = 2^{a-1}$ elements of order 2^a. If we think of an element in G as a vector with an element from $\mathbb{Z}_2$ in each of the first s_1 positions, an element from $\mathbb{Z}_{2^2}$ in each of the next s_2 positions and so on, then an element of order 2^a in G must have a generator of $\mathbb{Z}_{2^a}$ in at least one of the last s_a positions in the vector. So elements of G with order strictly less than 2^a may have any element in the first $s_1 + s_2 + \cdots + s_{a-1}$ positions and must have an element of order strictly less than 2^a

in each of the last s_a positions in the vector. Using these facts, we find that the number of elements of order 2^a in G is

$$2^{\left(\sum_{k=1}^{a-1} ks_k\right)+(a-1)s_a}(2^{s_a}-1).$$

Since we want this to equal 2^a, we get that $s_a = 1$, $s_1 = 1$, and $s_k = 0$ for $k = 2, \ldots, a-1$. Hence, $G \cong \mathbb{Z}_2 \times \mathbb{Z}_{2^a}$. Now assume that G_3 is nontrivial, so that $b \geq 1$. Then, equating the actual count of the number of elements in G of order $2^a 3^b$ to $2^a 3^b$ gives the equation

$$2^{\left(\sum_{k=1}^{a-1} ks_k\right)+(a-1)s_a}(2^{s_a}-1)3^{\left(\sum_{k=1}^{b-1} kt_k\right)+(b-1)t_b}(3^{t_b}-1) = 2^a 3^b \tag{1}$$

where we adopt the convention that the summation is zero if the upper limit is smaller than the lower limit. From (1) we see that $2^{s_a}-1$ must be a power of 3 and $3^{t_b}-1$ must be a power of 2. Therefore we have to solve the Diophantine equations

$$2^{s_a}-1 = 3^x \quad \text{and} \tag{2}$$

$$3^{t_b}-1 = 2^y. \tag{3}$$

When s_a is odd, 2^{s_a} is congruent to 2 modulo 3. So (2) is impossible modulo 3 unless $x = 0$, and hence $s_a = 1$. When s_a is even, we can factor the left side of (2) as the difference of two squares and conclude that the only solution is $s_a = 2$. Similarly, when t_b is even we get the solution $t_b = 2$; and when t_b is odd, 3^{t_b} is congruent to 3 modulo 4. Reduction of (3) modulo 4 produces the one solution $y = 1$ and consequently, $t_b = 1$. Hence, there are four cases to consider and in each case we equate exponents in (1) to determine if a group exists. The four cases are listed below.

- $s_a = t_b = 1$
  Equating exponents in (1) we get

  $$\sum_{k=1}^{a-1} ks_k = 0 \quad \text{and} \quad \sum_{k=1}^{b-1} kt_k = 1,$$

  which implies that $b \geq 2$, $s_1 = s_2 = \cdots = s_{a-1} = t_2 = t_3 = \cdots = t_{b-1} = 0$ and $t_1 = 1$. Hence,

  $$G \cong \mathbb{Z}_{2^a} \times \mathbb{Z}_3 \times \mathbb{Z}_{3^b}.$$

- $s_a = 2, t_b = 1$
  Equating exponents in (1) we get

  $$\sum_{k=1}^{a-1} ks_k = 1-a \quad \text{and} \quad \sum_{k=1}^{b-1} kt_k = 0,$$

  which implies that $a = 1$ and $t_1 = t_2 = t_3 = \cdots = t_{b-1} = 0$. Hence,

  $$G \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_{3^b}.$$

- $s_a = 1, t_b = 2$
  Equating exponents in (1) we get

  $$\sum_{k=1}^{a-1} ks_k = -2 \quad \text{and} \quad \sum_{k=1}^{b-1} kt_k = 2-b,$$

  which is impossible so no group exists for this case.

- $s_a = t_b = 2$
  Equating exponents in (1) we get
  $$\sum_{k=1}^{a-1} ks_k = -a - 1 \quad \text{and} \quad \sum_{k=1}^{b-1} kt_k = 1 - b,$$
  which again is impossible and no group exists for this case.

Therefore we have proven the following:

THEOREM 1. *An abelian group G is a minimal finite group with exactly $n > 1$ elements of order n if and only if $n = 2^a 3^b$ and G is isomorphic to one of the following groups:*

- $\mathbb{Z}_2 \times \mathbb{Z}_{2^a}, \quad a \geq 2, b = 0$
- $\mathbb{Z}_{2^a} \times \mathbb{Z}_3 \times \mathbb{Z}_{3^b}, \quad a \geq 1, b \geq 2$
- $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_{3^b}, \quad a = 1, b \geq 1$

Theorem 1 shows that there exists an abelian group with exactly $2^a 3^b$ elements of order $2^a 3^b$ except when $a \geq 2$ and $b = 1$. We see next that these curious exceptions are not duplicated in the nonabelian case.

The nonabelian case Of course from Proposition 2 we can easily manufacture nonabelian groups that have exactly n elements of order n by letting G be any of the groups from Theorem 1 and letting H be any nonabelian group such that the order of H is relatively prime to the order of G. This construction, however, violates our definition of minimality. To guarantee minimality we can assume that the nonabelian group G with exactly n elements of order n is generated by its subgroups of order n. This formulation allows us to use generators and relations to construct such groups. Theorem 2 gives the complete classification of minimal nonabelian groups that have exactly n elements of order n. For the details of the proof, see [**2**].

THEOREM 2. *A nonabelian group G is a minimal finite group with exactly $n > 1$ elements of order n if and only if $n = 2^a 3^b$ and G is isomorphic to one of the following groups:*

- $\langle x, y \mid x^{2^a} = y^3 = 1, x^{-1}yx = y^{-1}\rangle \times \mathbb{Z}_{3^b}$ *with* $a \geq 1$, $b \geq 1$.
- $\langle x, y \mid x^{2^a} = y^2 = 1, y^{-1}xy = x^{2^{a-1}+1}\rangle$ *with* $a \geq 3$, $b = 0$.
- $\mathbb{Z}_{2^a} \times \langle x, y \mid x^{3^b} = y^3 = 1, y^{-1}xy = x^{3^{b-1}+1}\rangle$ *with* $a \geq 1$, $b \geq 2$.
- $Q_8 \times \mathbb{Z}_{3^b}$ *with* $a = 2$, $b \geq 1$ *where* Q_8 *is the quaternion group of order 8.*
- $S_4 \times \mathbb{Z}_{3^b}$ *with* $a = 2$, $b \geq 1$ *where* S_4 *is the symmetric group on four letters.*
- $GL_2(3) \times \mathbb{Z}_{3^b}$ *with* $a = 3$, $b \geq 1$ *where* $GL_2(3)$ *is the group of* 2×2 *invertible matrices with entries from the field* $\mathbb{Z}_3$.
- $GL_2^*(3) \times \mathbb{Z}_{3^b}$ *with* $a = 3$, $b \geq 1$ *where* $GL_2^*(3)$ *is the group of order 48, which has generalized quaternion Sylow 2-subgroups and contains* $SL_2(3)$, *the group of all* 2×2 *matrices of determinant 1 with entries from the field* $\mathbb{Z}_3$, *as a subgroup of index 2.*

Open questions The following is a list of some unanswered questions for future investigation that have arisen from the work in this paper.

- For a given n, is it possible to determine what values of m are possible such that a finite group has exactly m elements of order n? What relationship can be found between m and n?

- For a given pair m and n, is it possible to classify all minimal finite groups having exactly m elements of order n and does this classification provide any insight into the groups themselves?

Acknowledgment. The authors thank the referees for the valuable suggestions.

REFERENCES

1. D. Dummit, R. Foote, *Abstract Algebra*, 2nd ed., John Wiley & Sons, 1999.
2. C. Finch, R. Foote, L. Jones, The classification of minimal finite groups that have exactly n elements of order n, preprint.

Rootless Matrices

BERTRAM YOOD
Pennsylvania State University
University Park, PA 16802

Let A be an $n \times n$ matrix over the complex field, $n \geq 2$. An rth root of A is a matrix S such that $S^r = A$. For example, $S^2 = W$ where

$$S = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}, \quad \text{and} \quad W = \begin{bmatrix} 1 & -3 \\ 0 & 4 \end{bmatrix},$$

so that S is the *square root* of W. It is natural to ask when a matrix does or does not have roots. We say that A is *rootless* if there is no matrix S and no positive integer $r \geq 2$ such that $S^r = A$.

This study started with the rather accidental discovery that the matrix

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

is rootless. To see this, suppose that $S^r = T$ for some $r \geq 2$ and

$$S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

The null space of T, that is, the set of all vectors $\binom{x}{y}$ for which $T\binom{x}{y} = \binom{0}{0}$, is the set of all vectors of the form $\binom{x}{0}$. The null space of S is contained in that of T (which is a rank-one matrix), and therefore the null space of S has dimension zero or dimension one. It cannot be zero-dimensional, for in that case S, and hence T, would be one-to-one, and hence invertible. Thus, $S\binom{1}{0} = \binom{0}{0}$, so that $a = c = 0$. We then have

$$\begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}^r = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

It is readily shown by induction that

$$\begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}^r = \begin{bmatrix} 0 & bd^{r-1} \\ d^r & 0 \end{bmatrix}.$$

Therefore, $bd^{r-1} = 1$ and $d^r = 0$, which is impossible, as $r \geq 2$.

It is easy to check that our matrix T above satisfies $T^2 = 0$. That is, T is a 2×2 nilpotent matrix such that $T^1 \neq 0$. Recall, an $n \times n$ matrix is A called *nilpotent* if $A^s = 0$ for some positive integer $s \geq 2$. We sought a more general result for which this example would be a special case:

THEOREM 1. *Let A be an $n \times n$ matrix over the complex field, $n \geq 2$. Then A is rootless if $A^{n-1} \neq 0$ and $A^n = 0$.*

For the following clever proof we are indebted to one of the referees.

Proof. The proof is by contradiction. Suppose that $A = S^r$, for $r \geq 2$. Then $S^{rn} = A^n = 0$ so that S is an $n \times n$ nilpotent matrix. It is well known that if some power of an $n \times n$ matrix is 0, then its nth power is zero (this will be proved independently below). Therefore, $S^k = 0$ for all positive integers $k \geq n$. But we also have

$$S^{r(n-1)} = A^{n-1} \neq 0.$$

Now $n \geq 2$ and $r \geq 2$, hence,

$$\frac{n}{n-1} \leq 2 \leq r, \quad \text{so that} \quad n \leq r(n-1).$$

Therefore $S^{r(n-1)} = 0$ or $A^{n-1} = 0$, which is contrary to the hypotheses on A. Hence A is rootless. ∎

We point out that A need not be rootless if $A^{k-1} \neq 0$ and $A^k = 0$ for some positive integer $k < n$. An example with $n = 3$ and $k = 2$ is the following:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

So the right-hand 3×3 matrix has a square root, yet its square vanishes.

The problem of solving the matrix equation $X^m = A$, where A is a given matrix, has been examined with care. Such a solution always exists if A is a self-adjoint matrix (that is, A is equal to its conjugate transpose). This is a consequence of the spectral theorem for self-adjoint matrices [**2**, Ch. 5]. But as we saw above, there do exist rootless matrices. If A is a nonsingular matrix (one with an inverse), solutions always exist. We cite the classical reference by Wedderburn [**3**, pp. 96–97]. Considerable attention has been given to special cases of A where the roots are polynomials in A. For reference, we cite MacDuffee [**1**, pp. 119–120].

The class of rootless matrices given in our theorem above is described in a *top-down* manner. We would like to add to this a *bottom-up* characterization which gives a more precise description of the shape and construction of the rootless matrices of the theorem.

We say that a matrix B is *upper triangular* if all entries of B below the diagonal are zero. That is, if the b_{ij} (the entry in the ith-row and jth-column of B) satisfy $b_{ij} = 0$ if $i > j$. A matrix B is said to be *strictly upper triangular* if every entry of B *on or below the diagonal* is zero (that is, $b_{ij} = 0$ if $i \geq j$). Lastly, recall that the *superdiagonal* of a matrix is the collection of entries immediately above and to the right of the diagonal (that is, entries b_{ij} such that $j = i + 1$). Then we have the following:

THEOREM 2. *Let A be an $n \times n$ matrix over the complex field, $n \geq 2$. Suppose that A is of the form $S^{-1}BS$ where S is an invertible matrix, and B is a strictly upper triangular matrix with all nonzero entries on its superdiagonal. Then A is rootless.*

This theorem will be proved as soon as it is shown that the "top-down" characterization ($A^n = 0$, $A^{n-1} \neq 0$) is equivalent to the "bottom-up" hypotheses of Theorem 2. That is,

THEOREM 3. *Let A be an $n \times n$ matrix over the complex field, $n \geq 2$. Then A is of the form $S^{-1}BS$ where S is an invertible matrix, and B is a strictly upper triangular matrix with all nonzero entries on its superdiagonal, if and only if A satisfies $A^n = 0$ and $A^{n-1} \neq 0$.*

We will prove Theorem 3 (and hence Theorem 2) by a series of lemmas. We first cite the well known result:

LEMMA 1. *Any $n \times n$ matrix A, $n \geq 2$, is similar to an upper triangular matrix.*

The proof for this may be found in many standard references, such as [**2**, Ch. 5]. Therefore, in our study of an $n \times n$ matrix A, we may assume that A is upper triangular. But we can say more, since we assume as well that $A^n = 0$.

LEMMA 2. *Suppose that A is an upper-triangular matrix such that $A^n = 0$. Then A is strictly upper triangular.*

Proof. For any complex number $\lambda \neq 0$, the matrix $\lambda I - A$ is invertible (where I is the identity matrix) since its inverse is $\lambda^{-1}I + \sum_{k=1}^{n-1} \lambda^{-k}A^k$. Thus zero is the only possible eigenvalue for A. As the diagonal elements of A are its eigenvalues (since it is assumed A is upper triangular), the diagonal elements are all zero. Hence A is strictly upper triangular. ■

Let Γ_n denote the set of all $n \times n$ strictly upper triangular matrices.

LEMMA 3. *Let V be the product of k matrices in Γ_n, $1 \leq k \leq n$. Then the first k columns and the last k rows of V are zero.*

Proof. We give a proof by mathematical induction. The statement is true by definition if $k = 1$. Let $1 \leq k < n$. We assume our result for any product $V = (v_{ij})$ of k elements of Γ_n.

Let $W = (w_{ij})$ be the product of $k + 1$ elements of Γ_n. We can express W as either BV_1 or V_2B for $B = (b_{ij})$ an element of Γ_n and V_1, V_2 each a product of k elements of Γ_n. We describe each of V_1 and V_2 in turn by (v_{ij}). By the inductive hypothesis $v_{ij} = 0$ if $j \leq k$ and $i \geq n - k$.

From $W = V_1B$ we see that the last k rows of W are zero and from $W = BV_2$ we see its first k columns are zero. We now extend this to show W is zero in column $k + 1$ and row $n - (k + 1) = n - 1 - k$.

From $W = V_1B$ we have, for any row i

$$w_{i,k+1} = \sum_{j=1}^{n} v_{ij}b_{j,k+1} = \sum_{j>k} v_{ij}b_{j,k+1}.$$

But $b_{j,k+1} = 0$ if $j \geq k + 1$, so we have $w_{i,k+1} = 0$. Thus W is zero in column $k + 1$, hence in its first $k + 1$ columns.

Similarly from $W = BV_2$ we have, for any column r

$$w_{n-k-1,r} = \sum_{j=1}^{n} b_{n-k-1,j}v_{jr}.$$

But $v_{jr} = 0$ for $j \geq n - k$ and $b_{n-k-1,j} = 0$ for $j \leq n - k - 1$ so that $w_{n-k-1,r} = 0$, and W is zero in row $n - (k + 1)$, hence zero in its last $k + 1$ rows. ■

As an immediate consequence of the above lemma we consider the cases $k = n$ and $k = n - 1$, and conclude:

LEMMA 4. *For $B \in \Gamma_n$, $B^n = 0$, that is, all elements of Γ_n are nilpotent. Further, let $B^{n-1} = (w_{ij})$. Then every $w_{ij} = 0$ except possibly w_{1n}.*

We next determine the value of that w_{1n}. Given an $n \times n$ matrix $B = (b_{ij})$ let $\pi(B)$ be the product of the entries on the superdiagonal: $\pi(B) := \prod_{i=1}^{n-1} b_{i,i+1}$.

LEMMA 5. *For $B \in \Gamma_n$ let $B^{n-1} = (w_{ij})$. Then $w_{1n} = \pi(B)$.*

Proof. Again, we argue by induction on n, the size of the matrix. The statement of the lemma is true for $n = 2$, since in that case

$$B = \begin{bmatrix} 0 & b_{12} \\ 0 & 0 \end{bmatrix}.$$

Suppose that the lemma is valid for all matrices in Γ_k where k is some particular integer, $1 \leq k < n$. We then must show that it is valid for any $V \in \Gamma_{k+1}$. Suppose $V \in \Gamma_{k+1}$, $V = (v_{ij})$. Since V is a strictly upper triangular matrix, its first column and last row are zero. Furthermore, $v_{i2} = 0$ for $i \geq 2$ and $v_{k+1,j} = 0$ for $1 \leq j \leq k+1$. Hence, if w let Q be the matrix obtained from V by deleting the first row and first column we have

$$Q = \begin{bmatrix} 0 & v_{23} & \cdots & v_{2,k+1} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & v_{k,k+1} \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Note that $Q \in \Gamma_k$. By construction, we have

$$V = \left[\begin{array}{c|ccc} 0 & v_{1,2} & \cdots & v_{1,k+1} \\ \hline \vdots & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right].$$

Since the first column of V is zero, we see that

$$V^{k-1} = \begin{bmatrix} 0 & \cdots & \cdots & \cdots \\ \vdots & & & \\ \vdots & & Q^{k-1} & \\ 0 & & & \end{bmatrix}.$$

Let $V^{k-1} = (w_{ij})$. By Lemma 4 and our induction hypothesis the entries in Q^{k-1} are all zero except the entry in the upper right corner; that is, $w_{2,k+1} = \pi(Q) = v_{23} \cdots v_{k,k+1}$. Note that $w_{i,k+1} = 0$ for $2 < i \leq k+1$. As the first row of V is $(0, v_{12}, \ldots)$ we see that the sole nonzero entry of $V^k = V \cdot V^{k-1}$ must be the $(1, k+1)$ entry with value $\pi(V)$. This completes our inductive argument. ∎

With this, the proof of Theorem 3 is nearly complete. Suppose A is an $n \times n$ matrix such that $A^n = 0$ and $A^{n-1} \neq 0$. By Lemmas 1 and 2 we may assume A is strictly upper triangular. But then by Lemma 4 the sole entry of A^{n-1} not known in advance to be zero is $a_{1,n}$, which must equal $\pi(A) = a_{1,2} \cdots a_{n-1,n}$ by Lemma 5, and this entry must be nonzero since $A^{n-1} \neq 0$. Therefore each $a_{i,i+1} \neq 0$ if $A^{n-1} \neq 0$. Similarly, if a matrix satisfies the hypotheses of Theorem 2, then it must also satisfy the hypotheses of Theorem 1.

There are some further questions the reader might like to consider. We have shown that nilpotent $n \times n$ matrices A such that $A^{n-1} \neq 0$ are rootless. Such nilpotent matrices are of greatest possible rank (here, rank $n - 1$). These have been termed in the literature *principal nilpotents* and are part of many interesting problems in matrix theory. All of the rootless matrices shown here are principal nilpotents, but there are nonnilpotent rootless matrices. The reader is encouraged to work this out to show the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

is rootless, but not nilpotent. Also, we saw an example of a nilpotent matrix that was *not* principal and had a square root. Is this always the case? That is, are there nilpotent matrices of less than maximal rank that are still rootless? With these interesting questions worked out, the reader should try to give a complete description of *all* rootless matrices, and we hope that our remarks will help you on your way.

REFERENCES

1. C. C. Macduffie, *The Theory of Matrices*, Chelsea Publ. Co., New York, 1946.
2. G. Strang, *Linear Algebra and its Applications*, 3rd ed., Harcourt, Brace, Jovanovich, San Diego, CA, 1988.
3. J. H. M. Wedderburn, Lectures on matrices, *Amer. Math. Soc. Coll. Publ.*, vol. **17**, Providence, RI, 1934.

Lots of Smiths

PATRICK COSTELLO
Eastern Kentucky University
Richmond, KY 40475
pat.costello@eku.edu

KATHY LEWIS
Somerset Community College
Somerset, KY 42501

Undoubtedly, many people had called Dr. Harold Smith at 493-7775 without thinking much about his phone number. Dr. Smith's brother-in-law, Albert Wilansky, however, noticed something very interesting about this phone number. When written as the single number 4937775, it is a composite number where the sum of the digits in its prime factorization is equal to the digit sum of the number. Adding the digits in the number and the digits of its prime factors 3, 5, 5, and 65837 resulted in identical sums of 42. Wilansky, a mathematics professor at Lehigh University, termed numbers having this property to be *Smith numbers* [**5**]. It turns out that the terminology was an appropriate choice because we will show that Smith numbers are very common, about as common as the name Smith in most American phone books.

The number 4 is the smallest Smith number because it is composite, it has a digit sum of 4, and the sum of the digits in its prime factorization is $2 + 2 = 4$. In his article, Wilansky provided two other slightly larger examples of Smith numbers: 9985 and 6036. He also told how many Smith numbers lie between 0 and 9999; as you can check, there are 376 of them. Because these numbers seemed to occur fairly frequently, Wilansky raised the question of whether there are infinitely many Smith numbers.

In 1987, Wayne McDaniel [3] succeeded in showing that infinitely many Smith numbers do in fact exist. McDaniel's approach was through a generalization of the problem. He defined a *k-Smith* number to be a composite integer where the sum of the digits in the prime factorization is equal to k times the digit sum of the number. In his article, McDaniel produced an infinite sequence of k-Smith numbers for each positive integer k. Since $k = 1$ corresponds to Wilansky's definition of a simple Smith number, McDaniel has shown that there are infinitely many Smith numbers. We will give further evidence of their abundance by producing yet another infinite sequence of Smith numbers.

Notation and basic theorems For any positive integer n, we let $S(n)$ denote the sum of the digits of n and $S_p(n)$ denote the sum of the digits of the prime factorization of n. A number is Smith when these two quantities are equal. For example, $S(27) = 2 + 7 = 9$ and $S_p(27) = S_p(3 * 3 * 3) = 3 + 3 + 3 = 9$. Hence 27 is a Smith number.

For any positive integer n, we let $N(n)$ denote the number of digits of n. For example, $N(27) = 2$. An algebraic formula for $N(m)$ is $N(m) = [\log_{10} m] + 1$ where $[x]$ is the greatest integer in x.

A *repunit*, denoted R_n, is a number consisting of a string of n ones. For example, $R_4 = 1111$. An algebraic formula for R_n is $R_n = (10^n - 1)/9$.

We now relate a few of the known results about the functions S and S_p that are pertinent to the construction of our infinite sequence of Smith numbers. In his paper, McDaniel [3] stated the following theorem without proof. (A detailed proof of the theorem is supplied in [2].)

THEOREM 1. *If t is a positive integer with $t < 9R_n$, then $S(t * (9R_n)) = S(9R_n) = 9n$.*

This theorem gives a way to know the digit sum of certain large numbers without having to expand the number into all its digits. For example, suppose we start with $9R_5 = 99999$ and arbitrarily choose $t = 44599$. Since $44599 < 99999$, it follows from Theorem 1 that $S(44599 * 99999) = 9 * 5 = 45$. We can check by expanding the product $44599 * 99999$. The product gives 4459855401 and in fact $S(4459855401) = 45$.

The following theorem follows directly from the definition of the function S_p.

THEOREM 2. *If m and n are positive integers, then $S_p(mn) = S_p(m) + S_p(n)$.*

This theorem says that the function S_p is an additive function, a fact we will use often.

In 1983, Keith Wayland and Sham Oltikar [4] provided another useful theorem.

THEOREM 3. *If $S(u) > S_p(u)$ and $S(u) \equiv S_p(u) (\mathrm{mod}\, 7)$, then $10^k u$ is a Smith number where $k = (S(u) - S_p(u))/7$.*

The essence of this theorem is that padding a zero on the end of a number does not change its digit sum, but it does increase the digit sum of its primes by 7. Each new zero on the end (which is achieved by multiplying by another factor of $10 = 2 * 5$) adds $2 + 5 = 7$ to the $S_p(u)$ value until it equals the $S(u)$ value. Theorem 3 was used by Oltikar and Wayland to say that every prime whose digits are all 0 and 1 has some multiple that qualifies as a Smith number. For example, some multiple of the prime number 1010111111 must be a Smith number. (Try multiplying the prime by 6 and then apply the Theorem.) Since there are lots of primes containing just 0s and 1s, this gave further circumstantial evidence (before McDaniel's proof) that there are infinitely many Smith numbers.

McDaniel [3] also gave an upper bound for $S_p(m)$ that does not involve the value of specific prime factors of m, as follows:

THEOREM 4. *If* $p_1, \ldots, p_r$ *are prime numbers, not necessarily distinct, and if* $m = p_1 p_2 \cdots p_r$, *then* $S_p(m) < 9N(m) - .54r$.

The new infinite sequence With the help of the previous theorems, we begin to construct a new infinite sequence of Smith numbers. Start with an integer n greater than 7. Let $m = 9\mathrm{R}_n = 10^n - 1$. By Theorem 4, $S_p(m) < 9N(m) - .54r$, where r is the number of prime factors of m. Since 3^2 divides m and $\mathrm{R}_n > 1$, m has at least three primes in its factorization. This means that $r > 2$ and $.54r > 1$. Since m has n digits, Theorem 4 gives us $S_p(m) < 9n - 1$. But $9n$ is actually the digit sum of m and so $S_p(m) < S(m) - 1$. This implies that $0 < S(10^n - 1) - S_p(10^n - 1)$. Then let x be the least residue of $S(10^n - 1) - S_p(10^n - 1)$ modulo 7.

We now show that if we multiply $10^n - 1$ by a power of 11 that involves the computed least residue x, we get a number where the S and S_p values are congruent mod 7. This will be the main ingredient used in generating the new infinite sequence of Smith numbers.

THEOREM 5. *Let x be the least residue of* $S(10^n - 1) - S_p(10^n - 1)$ *modulo 7. Let j be the least residue of* $4x$ *modulo 7. Then*

$$S(11^j(10^n - 1)) \equiv S_p(11^j(10^n - 1)) (\mathrm{mod}\, 7).$$

Proof. First observe that $S(11^j(10^n - 1)) - S_p(11^j(10^n - 1)) = S(10^n - 1) - S_p(11^j(10^n - 1))$ by Theorem 1 and the choice of $n \geq 8$ so that $11^j < 10^n - 1$. This is equal to

$$\begin{aligned} & S(10^n - 1) - S_p(11^j) - S_p(10^n - 1) \quad \text{by Theorem 2} \\ &= S(10^n - 1) - 2j - S_p(10^n - 1) \quad \text{by Theorem 2} \\ &= (S(10^n - 1) - S_p(10^n - 1)) - 2j \\ &\equiv x - 2j (\mathrm{mod}\, 7) \end{aligned}$$

But since $4x \equiv j (\mathrm{mod}\, 7)$, $8x \equiv 2j (\mathrm{mod}\, 7)$, and the expression above is congruent to zero. Hence $S(11^j(10^n - 1)) - S_p(11^j(10^n - 1)) \equiv 0 (\mathrm{mod}\, 7)$. ∎

Using Theorems 3 and 5, we can now construct the infinite sequence of Smith numbers. Let n be an integer greater than 7. Compute x as the least residue of $S(10^n - 1) - S_p(10^n - 1)$ modulo 7. Compute j to be the least residue of $4x$ mod 7. Then $S(11^j(10^n - 1)) - S_p(11^j(10^n - 1)) \equiv 0 (\mathrm{mod}\, 7)$ by Theorem 5. So let $k = (S(11^j(10^n - 1)) - S_p(11^j(10^n - 1)))/7$. Then the number $a_n = 10^k \cdot 11^j(10^n - 1)$ is a Smith number by Theorem 3. Since each integer $n \geq 8$ gives a Smith number, there must be infinitely many Smith numbers.

Examples We now show the computations needed to produce two specific Smith numbers in our infinite sequence.

EXAMPLE 1. Let $n = 8$. Then $10^8 - 1 = 99999999$. In this case, $S(10^8 - 1) = 8 * 9 = 72$ and $S_p(10^8 - 1) = S_p(3 * 3 * 11 * 73 * 101 * 137) = 31$ so that $S(10^8 - 1) - S_p(10^8 - 1) = 72 - 31 = 41 \equiv 6 (\mathrm{mod}\, 7)$. Then $x = 6$ in Theorem 5 and $4 * 6 \equiv 3 (\mathrm{mod}\, 7)$ which gives us $j = 3$. We let $k = (S(11^3(10^8 - 1) - S_p(11^3(10^8 - 1)))/7$. Then $k = (S(133099998669) - S_p(3 * 3 * 11 * 11 * 11 * 11 * 73 * 101 * 137))/7 = (72 - 37)/7 = 35/7 = 5$. Finally, $10^5 * 11^3(10^8 - 1) = 13309999866900000$ is the first Smith number in our sequence.

In 1925, Lt.-Col. Allan J. C. Cunningham and H. Woodall published a small volume of tables of the factorizations of $b^n \pm 1$ for the bases $b = 2, 3, 5, 6, 7, 10, 11, 12$ to various powers of n. The authors left blanks in the tables where new factors could be entered. They put question marks on numbers of unknown character. Most importantly, they gave credit to those who had discovered notable factors in the past. All of these techniques stimulated work on the remaining composite numbers in the tables. The ongoing work on the Cunningham-Woodall tables has usually been referred to as the Cunningham project.

Factorizations of $b^n \pm 1$ for the bases $b = 2, 3, 5, 6, 7, 10, 11, 12$ to various high powers n are easily available [**1**]. In fact, for any value of $10^n - 1$, $n \geq 8$ that has been completely factored in the table, we can find the corresponding Smith number belonging to our sequence.

EXAMPLE 2. Let $n = 44$. The factorization of $10^{44} - 1$ is given [**1**] as $10^{44} - 1 = 3^2 * 11^2 * 23 * 89 * 101 * 4093 * 8779 * 21649 * 513239 * 1052788969 * 1056689261$. Adding up the digits in the prime factors, we get $S_p(10^{44} - 1) = 225$. Since $10^{44} - 1 = 9R_{44}$, we have that $S(10^{44} - 1) = 44 * 9 = 396$. Then $S(10^{44} - 1) - S_p(10^{44} - 1) = 396 - 225 = 171 \equiv 3 (\mathrm{mod}\, 7)$. So $x = 3$ in Theorem 5 and $4 * 3 \equiv 5 (\mathrm{mod}\, 7)$ which gives us $j = 5$.

We let $k = ((S(11^5(10^{44} - 1) - S_p(11^5(10^{44} - 1)))/7 = (396 - (225 + 5 * 2))/7 = (396 - 235)/7 = 161/7 = 23$. Thus the 73-digit number $10^{23} 11^5 (10^{44} - 1)$ is the Smith number in our sequence corresponding to $n = 44$.

The Smith numbers that McDaniel produces in his infinite sequence have the form $t(10^n - 1)10^m$, where t is chosen from the set $\{2, 3, 4, 5, 7, 8, 15\}$. Our Smith numbers replace the t value with a power of 11 and sometimes alter the m value. When $n = 8$, McDaniel's procedure gives $8(10^8 - 1)10^5$; when $n = 44$, it gives $3(10^{44} - 1)10^{24}$. Our slight change has produced an entirely different infinite sequence of Smith numbers. We leave the reader with a challenge. Since there seem to be lots of Smith numbers, can you find another infinite sequence of Smiths? (Hint: Look back at Theorem 5 and see what role the digit sum of 11 played. The key is that 2 is relatively prime to 7.)

REFERENCES

1. J. Brillhart, D. H. Lehmer, J. Selfridge, B. Tuckerman, and S. Wagstaff, Jr., *Factorizations of* $b^n \pm 1$, $b = 2, 3, 5, 6, 7, 10, 11, 12$ *Up to High Powers*, 2nd ed., American Mathematical Society, Providence, Rhode Island, 1988.
2. K. Lewis, *Smith Numbers: An Infinite Subset of* $\mathbb{N}$, M.S. thesis at Eastern Kentucky University, 1994.
3. W. McDaniel, The existence of infinitely many k-Smith numbers, *Fibonacci Quarterly* **25** (1987), 76–80.
4. S. Oltikar and K. Wayland, Construction of Smith numbers, this MAGAZINE **56** (1983), 36–37.
5. A. Wilansky, Smith numbers, *Two-Year College Math Journal* **13** (1982), 21.

PROBLEMS

ELGIN H. JOHNSTON, *Editor*
Iowa State University

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

Proposals

To be considered for publication, solutions should be received by November 1, 2002.

1648. *Proposed by Erwin Just (Emeritus) Bronx Community College, New York, NY.*

Prove that there exist an infinite number of integers, none of which is expressible as the sum of a prime and a perfect square.

1649. *Proposed by K. R. S. Sastry, Bangalaore, India.*

Prove that if a right triangle has all sides of integral length, then it has at most one angle bisector of integral length.

1650. *Proposed by M. N. Deshpande, Nagpur, India.*

Let $R(\theta)$ denote the rhombus with unit side and and a vertex angle of θ, and let $n \geq 2$ be a positive integer. Prove that a regular $4n$-gon of unit side can be tiled with the collection of $n(2n-1)$ rhombi consisting of n copies of $R(\frac{\pi}{2})$ and $2n$ copies of each of $R(\frac{\pi k}{2n})$, $1 \leq k \leq n-1$.

1651. *Proposed by Juan-Bosco Romero Márquez, Universidad de Valladolid, Valladolid, Spain.*

Prove that for $x \geq 2$,

$$\left(\frac{x}{e}\right)^{x-1} \leq \Gamma(x) \leq \left(\frac{x}{2}\right)^{x-1},$$

where Γ is the gamma function.

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

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

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

1652. *Proposed by Razvan A. Satnoianu, City University, London, United Kingdom.*

In triangle ABC, let r denote the radius of the inscribed circle, R the radius of the circumscribed circle, and p the semiperimeter. Prove the following inequalities, and show that in each case the constant on the right is the best possible:

(a) $\dfrac{R}{p}+\dfrac{p}{R}\geq 2.$

(b) $\dfrac{r}{p}+\dfrac{p}{r}\geq \dfrac{28\sqrt{3}}{9}.$

(c) $\dfrac{r}{p}+\dfrac{p}{r}\geq \dfrac{56}{31}\left(\dfrac{R}{p}+\dfrac{p}{R}\right).$

Quickies

Answers to the Quickies are on page 233.

Q921. *Proposed by Kent Holing, Statoil Research Center, Trondheim, Norway.*

Let m and n be relatively prime positive integers. Show that the numbers $\sqrt[3]{m/n}$ and $\sqrt[3]{m+n}$ are not both constructible with straightedge and compass.

Q922. *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, AB, Canada.*

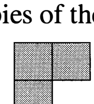
Two directly homothetic triangles are such that the incircle of one of them is the circumcircle of the other. If the ratio of their areas is 4, prove that the triangles are equilateral.

Solutions

Tromino Tiles **June 2001**

1623. *Proposed by Emeric Deutsch, Polytechnic University, Brooklyn, NY.*

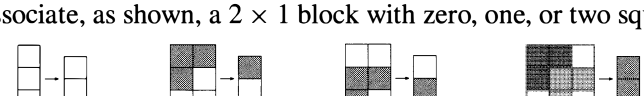
Find the number of ways that k copies of the tromino



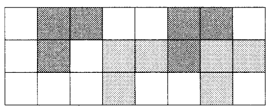
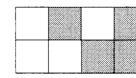
can be placed, with the orientation shown and without overlapping, on a $3\times n$ rectangle.

I. *Solution by Stephen Blair, Portland State University, Portland, OR.*

Any configuration of k trominos on a $3\times n$ rectangle can be described as an ordered juxtaposition of four types of column structures. These types are blank columns: , pairs of columns containing one tromino:  or , and sets of three adjacent columns containing two trominos: . With each of these four column sets we associate, as shown, a 2×1 block with zero, one, or two squares shaded:



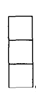
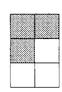
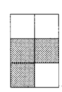
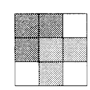
Note that the number of squares shaded in the 2×1 block is equal to the number of trominos in the associated column structure. For a given configuration of k trominos on a $3 \times n$ board, we associate a $2 \times (n-k)$ board on which k squares are shaded. This is done by partitioning the k tromino configuration into the four types of column structures, then replacing each column structure with its associated 2×1 block. For example, the 3×8, four-tromino configuration

 is associated with the $2 \times (8-4)$ board .

It is not hard to see that this mapping scheme defines a bijection between the configurations of k trominos on the $3 \times n$ board and the set of $2 \times (n-k)$ boards with k cells colored. Because there are $\binom{2(n-k)}{k}$ ways to color k squares on a $2 \times (n-k)$ board, there are also $\binom{2(n-k)}{k}$ ways to place k trominos on a $3 \times n$ board.

II. *Solution by the proposer.*

Imagine that the uncovered squares of the $3 \times n$ rectangle are covered by 1×1 squares, so the $3 \times n$ rectangle is tiled by $3n - 3k$ pieces  and k pieces . The only tilings that are not concatenations of tilings of smaller rectangles are

, , , and . $(*)$

Furthermore, if a tiling is not one of the tilings in $(*)$, then it is a concatenation of a finite sequence of such tilings.

Now let $a_{n,k}$ be the number of ways to tile a $3 \times n$ rectangle using k tromino pieces and $3n - 3k$ single square pieces, and let $G(t, z) = \sum_{n,k \geq 0} a_{n,k} t^k z^n$ be the generating function for the $a_{n,k}$, where we set $a_{0,0} = 1$. Because the generating function for the tilings $(*)$ is

$$P(t, z) = z + 2tz^2 + t^2 z^3 = z(1 + tz)^2,$$

and any tiling is a concatenation of tilings from $(*)$, it follows that

$$G(t, z) = \sum_{j=0}^{\infty} P(t, z)^j = \sum_{j=0}^{\infty} z^j (1 + tz)^{2j} = \sum_{i,j \geq 0} \binom{2j}{i} t^i z^{i+j}.$$

The coefficient of $t^k z^n$ is $a_{n,k} = \binom{2n-2k}{k}$.

Also solved by D. Bednarchak, Agnes Benedek (Argentina), Robert E. Bernstein, Jany C. Binz (Switzerland), Tom Boerkoel, Marc Brodie, Knut Dale (Norway), Daniele Donini (Italy), Marty Getz and Dixon Jones, Jerrold W. Grossman, Tom Jager, S. C. Locke, Reiner Martin, Carl P. McCarty and Loretta McCarty, Rob Pratt, Les Reid, William Tressler, LeRoy Wenstrom, WMC Problems Group, Michel Woltermann, and Li Zhou. There were two incorrect submissions.

An Ellipsoid Tangent to a Tetrahedron — June 2001

1624. *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, AB, Canada.*

An ellipsoid is tangent to each of the six edges of a tetrahedron. Prove that the three segments joining the points of tangency of opposite edges are concurrent.

Solution by Michel Bataille, Rouen, France.

Under a suitable affine transformation, the ellipsoid becomes a sphere, and concurrency and tangency are preserved. Thus we need only consider the case in which the ellipsoid is a sphere that is tangent at points R, S, T, U, V, and W to sides BC, CA, AB, DA, DB, and DC, respectively, of tetrahedron $ABCD$. Because all segments of tangents from a vertex to the point of tangency on the sphere have the same length, we can set $x = AS = AT = AU$, $y = BT = BR = BV$, $z = CR = CS = CW$, and $t = DU = DV = DW$. Denoting by $\mathbf{M}$ the vector from the origin to the point M, let I be the point determined by

$$m\mathbf{I} = yzt\mathbf{A} + ztx\mathbf{B} + txy\mathbf{C} + xyz\mathbf{D},$$

where $m = yzt + ztx + txy + xyz$. Then

$$m\mathbf{I} = zt(y\mathbf{A} + x\mathbf{B}) + xy(t\mathbf{C} + z\mathbf{D}) = zt(y + x)\mathbf{T} + xy(t + z)\mathbf{W}.$$

Because $zt(y + x)$ and $xy(t + z)$ are positive and sum to m, it follows that I lies on segment TW. Similarly,

$$m\mathbf{I} = yz(t\mathbf{A} + x\mathbf{D}) + tx(z\mathbf{B} + y\mathbf{C}) = yz(t + x)\mathbf{U} + tx(z + y)\mathbf{R},$$

and

$$m\mathbf{I} = ty(z\mathbf{A} + x\mathbf{C}) + zx(t\mathbf{B} + y\mathbf{D}) = ty(z + x)\mathbf{S} + zx(t + y)\mathbf{V},$$

showing that I lies on segments UR and SV as well. Thus the three segments joining points of tangency of opposite edges are concurrent at I.

Also solved by Daniele Donini (Italy), Ovidiu Furdui, Michael Golomb, Joel Schlosberg, Peter Y. Woo, and the proposer.

A Product of Powers and a Power of Products — June 2001

1625. *Proposed by Mihàly Bencze, Romania.*

Let $x_1, x_2, \ldots, x_n$ be positive real numbers and let $a_1, a_2, \ldots, a_n$ be positive integers. Prove that

$$\prod_{k=1}^{n}\left(1 + x_k^{1/a_k}\right)^{a_k} \geq \left(1 + \left(\prod_{k=1}^{n} x_k\right)^{1/\sum_{k=1}^{n} a_k}\right)^{\sum_{k=1}^{n} a_k}.$$

Solution by Robert R. Burnside, University of Paisley, Scotland.

We prove a more general version of the inequality. Let b_k and y_k, $k = 1, 2, \ldots, n$, be fixed positive real numbers, with $\sum_{k=1}^{n} b_k = 1$. For $s \geq 0$, define $f(s) = \prod_{k=1}^{n}(s + y_k)^{b_k}$. Then

$$\frac{f'(s)}{f(s)} = \sum_{k=1}^{n} \frac{b_k}{s + y_k} \geq \frac{1}{\prod_{k=1}^{n}(s + y_k)^{b_k}} = \frac{1}{f(s)},$$

where we have used the weighted arithmetic mean-geometric mean inequality. It follows that $f'(s) \geq 1$ and hence that $f(x) - f(0) \geq x$ for all $x \geq 0$. Consequently, $\prod_{k=1}^{n}(x + y_k)^{b_k} \geq x + \prod_{k=1}^{n} y_k^{b_k}$, with equality for x nonzero, if and only if $y_1 = y_2 = \cdots = y_n$.

The inequality in the problem statement is obtained by taking $x = 1$, $y_k = x_k^{1/a_k}$, and $b_k = a_k / \sum_{k=1}^{n} a_k$, with all $a_k > 0$. If all $a_k < 0$, then the inequality sign is reversed.

Also solved by Michel Bataille (France), Jean Bogaert (Belgium), Knut Dale (Norway), Minh Can, Daniele Donini (Italy), Costas Efthimiou, Ovidiu Furdui, Tom Jager, Reiner Martin, Michael G. Neubauer, Joel Schlosberg, Heinz-Jürgen Seiffert (Germany), Beiment Teclezghi and Tewodros Amdeberhan, Xianfu Wang (Canada), Li Zhou, and the proposer.

A Condition Implying Additivity **June 2001**

1626. *Proposed by Ho-joo Lee, student, Kwangwoon University, Seoul, South Korea.*

Let $f, g, h : \mathbb{R} \to \mathbb{R}$ be functions such that $f(g(0)) = g(f(0)) = h(f(0)) = 0$ and

$$f\,(x + g(y)) = g\,(h\,(f(x))) + y \qquad (*)$$

for all $x, y \in \mathbb{R}$. Prove that $h = f$ and that $g(x + y) = g(x) + g(y)$ for all $x, y \in \mathbb{R}$.

Solution by Michael K. Kinyon, Western Michigan University, Kalamazoo, MI.

Setting $x = 0$ in $(*)$ gives

$$f\,(g(y)) = g\,(h\,(f(0))) + y = g(0) + y,$$

for all y. Setting $y = 0$ in this expression gives

$$g(0) = f\,(g(0)) = 0,$$

and it follows that $f(g(y)) = y$ for all y. Thus f is surjective and is a left inverse of g. Setting $y = 0$ in $(*)$ we have

$$f(x) = f\,(x + g(0)) = g\,(h\,(f(x)))\,.$$

Because f is surjective, it follows that g is surjective and h is a right inverse of g. The left and right inverses of g must be equal, so we have $g^{-1} = f = h$. Using this in $(*)$ we obtain

$$f\,(x + g(y)) = f(x) + y.$$

Substituting $x = g(u)$ in this last expression, then applying g to both sides yields

$$g(u) + g(y) = g(u + y),$$

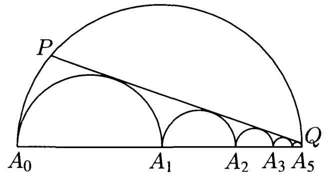
for all $u, y \in \mathbb{R}$. Note that the condition $g(f(0)) = 0$ was not used.

Also solved by Hamza Ahmad, Claudi Alsina (Spain), Geta Techanie Ayele, S. Floyd Barger, Michel Bataille (France), Brian D. Beasley, Anthony C. Blackman and Eduard S. Belinsky (Barbados), Jean Bogaert (Belgium), Marc Brodie, Minh Can, Ron Martin Carroll, Con Amore Problem Group (Denmark), Knut Dale (Norway), Richard Daquila, Charles R. Diminnie, Daniele Donini (Italy), Tim Edwards, Costas Efthimiou, Ovidiu Furdui, Michel Golomb, Kazuo Goto (Japan), Lee O. Hagglund, Tracy Dawn Hamilton and Howard B. Hamilton, Damion J. Hammock, Brian Hogan, Joel Iiams, Tom Jager, J. Todd Lee and Paula Grafton Young, S. C. Locke, Hieu D. Nguyen, Stephen Noltie, Perry and the Masons Solving Group (Spain), Victor Pambuccian, David R. Patten, Sam L. Robinson and Gerald Thompson, Richard F. Ryan, Grigor Sargsyan, Joel Schlosberg, Heinz-Jürgen Seiffert, Laishram Shanta Singh and Ritumoni Sarma (India), Shing S. So, Beiment Teclezghi and Tewodros Amdeberhan, Nora S. Thornber, Thomas Vanden Eynden, Gregory P. Wene, LeRoy Wenstrom, Western Maryland College Problems Group, Li Zhou, and the proposer.

A Generalization of the Arbelos **June 2001**

1627. *Proposed by Jiro Fukata, Shinsei-cho, Gifu-ken, Japan.*

Semicircle C has diameter A_0A_n. Semicircles $C_1, C_2, \ldots, C_n$ are drawn so that C_k has diameter $A_{k-1}A_k$ on A_0A_n. In addition, C_1 is internally tangent to C at A_0 and externally tangent to C_2 at A_1, C_n is internally tangent to C at A_n and externally tangent to C_{n-1} at A_{n-1}, for $2 \le k \le n-1$, C_k is externally tangent to C_{k-1} and C_{k+1} at A_{k-1} and A_k respectively, and each C_k, $1 \le k \le n$ is tangent to a chord PQ of C. The case $n = 5$ is illustrated in the accompanying figure.



(a) Let A_1A_1' and $A_{n-1}A_{n-1}'$ be perpendicular to A_0A_n at A_1 and A_{n-1}, respectively. Let circle X be externally tangent to C_2, internally tangent to C and tangent to A_1A_1' on the side opposite C_1, and let circle Y be externally tangent to C_{n-1}, internally tangent to C and tangent to $A_{n-1}A_{n-1}'$ on the side opposite C_n. Prove that circle X is congruent to circle Y.

(b) Suppose C_0 is a semicircle with diameter on A_0A_n and tangent to PQ. Let D and E be the endpoints of its diameter. Lines DD' and EE' are drawn perpendicular to A_0A_n Let Z be the circle tangent to each of DD' and EE' and internally tangent to C. Show that Z is tangent to the circle with radius A_1A_{n-1}.

Solution by Marty Getz and Dixon Jones, University of Alaska, Fairbanks, AK.

(a) We show that circles X and Y each have diameter $\frac{(A_0A_{n-1})(A_1A_n)}{A_0A_{n-1}+A_1A_n}$. First observe that that

$$\frac{A_jA_{j+1}}{A_{j-1}A_j} = \frac{1-\sin\theta}{1+\sin\theta}, \qquad j = 1, 2, \ldots, n-1,$$

where θ is the angle determined by the extended chord PQ and the extended diameter A_0A_n. In particular, with $a = \frac{1-\sin\theta}{1+\sin\theta}$, we have

$$\frac{A_1A_n}{A_0A_{n-1}} = \frac{A_1A_2 + A_2A_3 + \cdots + A_{n-1}A_n}{A_0A_1 + A_1A_2 + \cdots + A_{n-2}A_{n-1}} = a.$$

Let B be the point at which circle X touches C, and let line A_nB intersect line A_1A_1' in T. See Figure 1. Let A_0B meet X in R and A_nB meet X in S. Because $\angle A_0BA_n$ is a right angle, it follows that RS is a diameter of X and is parallel to A_0A_n. In particular, R is the point at which circle X touches A_1A_1'. Because $\angle TSA_1 = 180° - \angle BRA_2 = \angle A_0RA_2$ and $\angle A_1TS = \angle RA_0A_2$, it follows that $\triangle TSA_1 \sim \triangle A_0RA_2$. Hence $\frac{RA_1}{TR} = \frac{A_1A_2}{A_0A_1} = a$. Furthermore, because

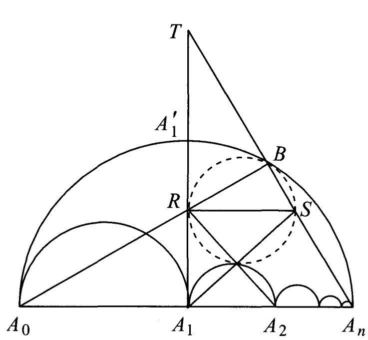


Figure 1

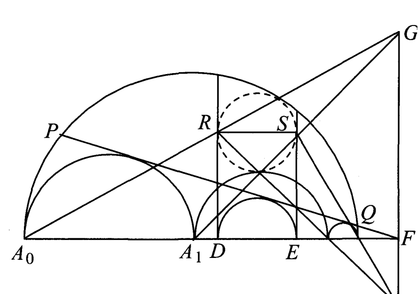


Figure 2

$\triangle TRS \sim \triangle TA_1A_n$, we have

$$\frac{RS}{A_1A_n} = \frac{TR}{TA_1} = \frac{TR}{TR+RA_1} = \frac{1}{1+a} = \frac{A_0A_{n-1}}{A_0A_{n-1}+A_1A_n}.$$

Thus, $RS = \frac{(A_0A_{n-1})(A_1A_n)}{A_0A_{n-1}+A_1A_0}$. A symmetric argument gives the same result for the diameter of Y.

(b) Figure 2 shows circle Z with diameter RS parallel to A_0A_n. Let F be the intersection of lines PQ and A_0A_n. Because $\frac{A_0A_1}{RS} = \frac{A_0A_1}{DE}$, the intersection G of lines A_0R and A_1S lies on the line through F and perpendicular to line A_0F. By a similar argument, the intersection H of lines $A_{n-1}R$ and A_nS also lies on line GF. Thus RS and HS lie along two of the altitudes of $\triangle RGH$. Because GS is concurrent with these two lines, it lies along the third altitude of $\triangle RGH$. Thus, SA_1 is perpendicular to RA_{n-1}, and it follows that circle Z is tangent to the circle with diameter A_1A_{n-1}.

Also solved by Herb Bailey, Michel Bataille (France), Jany C. Binz (Switzerland), Daniele Donini (Italy), Joel Schlosberg, and the proposer.

Answers

Solutions to the Quickies from page 228.

A921. It is known that $r = \sqrt[3]{m/n}$ is constructible if and only if r is rational and that $s = \sqrt[3]{m+n}$ is constructible if and only if s is an integer. (See George E. Martin, *Geometric Constructions*, Springer Verlag, 1998.) Furthermore, r is a rational number if and only if both m and n are cubes, and s is an integer if and only if $m+n$ is a cube.

Now assume that r is constructible. Then $m = p^3$ and $n = q^3$ for integers p and q, and $m+n = p^3+q^3$. Thus, by Fermat's Theorem, $m+n$ cannot also be a cube. Therefore $\sqrt[3]{m+n}$ cannot be constructed.

A922. Let the sides, area, circumradius, and inradius of the larger triangle be a, b, c, F, R, and r, respectively, and let the corresponding sides and area of the smaller triangle be a', b', c', and F'. We then have

$$\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'} = 2, \qquad 4FR = abc, \qquad \text{and} \qquad 4F'r = a'b'c'.$$

It follows that

$$\frac{FR}{F'r} = \frac{abc}{a'b'c'} = 8,$$

and hence that $R = 2r$. However, it is known that $R \geq 2r$ with equality if and only if the triangle is equilateral.

REVIEWS

PAUL J. CAMPBELL, *Editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles and books 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.

Johnson, Valen E., An A is an A is an A ... and that's the problem, *New York Times* (14 April 2002), Section 4A (Education Life), 14; `http://www.nytimes.com/2002/04/14/edlife/14EDVIEW.html`. Special section: The grade inflation problem, *The UMAP Journal* 19 (3) (1998) 279–336. Johnson, Valen E., An alternative to traditional GPA for evaluating student performance, *Statistical Science* 12 (4) 251–278; `ftp://ftp.isds.duke.edu/pub/WorkingPapers/96-20.ps`.

Last week, a statistically sophisticated colleague in another department approached me for help in devising a system to keep up with the current level of grade inflation, so that his students would not be "penalized" relative to others. Author Johnson (Statistics and Decision Sciences, Duke University) was the main proponent of a 1996 proposal to revise the calculation of grade-point averages (GPAs) at Duke to take into account difficulty of the course, quality of students in the course, and the instructor's grading history. Duke rejected the proposal, but Johnson returns here with further data on ramifications—not of grade inflation per se, but of grading inequity, which can be a consequence of grade inflation not being uniform among departments or instructors. He examined the effect of grades on student evaluations of courses and instructors and on student choice of courses, at Duke. The results were what you might expect: Students are much more likely to give low evaluations in courses where they expect lower grades than usual, and students are much more likely to enroll with instructors who grade higher. Apart from the consequences for retention and advancement of individual faculty (which can be dire), there are implications for departments and the institution as a whole. Differences in grading can result in—in fact, probably already have at your institution—shifts in enrollments and allocation of resources. If your institution is like Duke (and most others), *you* as a mathematics instructor have a particular problem, *since your department grades the lowest* (or nearly). "Uneven grading practices allow students to manipulate their grade point averages and honors status by selecting certain courses, and discourage them from taking courses that would benefit them [think mathematics courses]. By rewarding mediocrity, excellence is discouraged." The Special Section in *The UMAP Journal* contains three Outstanding entries in COMAP's 1998 Mathematical Contest in Modeling on the Grade Inflation Problem, along with a commentary by Johnson; his article in *Statistical Science* details the Duke proposal. Watch for his forthcoming book *College Grading: A National Crisis in Undergraduate Education.*

Primus: Problems, Resources, and Issues in Mathematics Undergraduate Studies. Special Issues. The Undergraduate Seminar in Mathematics. Part 1: September 2001; Part 2: December 2001; 11 (3 and 4) 193–257, 289–369.

These special issues of *Primus* offer 11 interpretations of what a seminar in mathematics can be about, from presenters at the New Orleans Joint Mathematics Meetings in January 2001. The ideas and experiences vary in level (freshman to senior), focus (communication skills, integration of mathematical ideas), faculty role, and grading, but all feature student involvement at a fundamentally different level than in other courses. Does your department offer such a seminar? Do you need ideas or new ideas? Take a look here for inspiration.

Chown, Marcus, Smash and grab, *New Scientist* (6 April 2002) 24–28. Calude, C.S., and B. Pavlov, Coins, quantum measurements, and Turing's barrier, *Quantum Information Processing* (in press); `http://www.cs.auckland.ac.nz/CDMTCS/researchreports/170boris.pdf` .

Will quantum computers really make a difference? Will they make a difference to mathematics? Cristian S. Calude (University of Auckland) thinks that "quantum computing is theoretically capable of computing uncomputable functions." Taking the halting problem as the uncomputable function, the key ideas are to superimpose simultaneously an infinite number of quantum states (via Hilbert space) and then to detect and measure the probability of a program halting. After some finite amount of time, you get an answer: not an absolute yes-or-no answer but an answer with an accompanying (tiny) probability (want a larger probability? run the quantum program again for longer). "Because of these new computational models, the idea of 'proof' might ... change."

Matthews, Robert, $1 million mathematical mystery "solved," NewScientist.com news service; `http://www.newscientist.com/news/news.jsp?id=ns99992143` . Matthews, Robert, British professor chases solution to $1m maths prize, *Daily Telegraph* (13 April 2002). Dunwoody, Martin, A proof of the Poincaré conjecture? (revised version eight, 11 April 2002); `http://www.maths.soton.ac.uk/ mjd/Poin.pdf` .

Martin Dunwoody (Southampton University) claimed to have proved the famous Poincaré conjecture, which states that if every loop on a compact 3-D manifold can be shrunk to a point, the manifold is topologically equivalent to a sphere? This is one of the seven Clay Mathematics Institute million-dollar mathematical questions. The press (at least in Great Britain) got excited. Mathematicians found a gap. (Does this sequence of events sound familiar?) At this writing, Dunwoody has added a question mark at the end of the title of his preprint, which is more of a blueprint for a proof than a proof itself. Stay tuned, but don't hold your breath.

Flannery, Sarah, with David Flannery, *In Code: A Mathematical Journey*, Workman, 2001; ix + 341 pp, $24.95. ISBN 0–19628–12384–8.

High-school student Sarah Flannery invented a new public-key cryptographic algorithm (which she calls the Cayley-Purser algorithm) as her project entry in the Irish Young Scientist competition. Her algorithm uses matrices but not modular exponentiation, hence it is 20 to 30 times as fast in practice as the RSA algorithm. This book, written in part by her mathematician father, details how she came to enter the contest, the fame that winning it brought her, the accompanying stress (due to press publicity of the potential of her becoming wealthy from selling her idea), and (in an appendix) all the mathematics behind the algorithm. There is enough exposition in the text itself of the elementary aspects of public-key cryptography and of matrices so that the reader gets the flavor of the subject and her work. Her tale rambles; but on the whole it is inspiring, and the personal nature of the writing may help add to its appeal to young readers.

Weibel, Ewald R., *Symmorphosis: On Form and Function in Shaping Life*, Harvard University Press, 2000; xiii + 263 pp, $45. ISBN 0–674–00068–4.

"This book addresses a simple question: Are animals designed economically?" The "symmorphosis" of the title refers to sizing of parts of a system to its function, including providing some margin of safety. Author Weibel works out "the quantitative relations between form and function" in various physiological settings: cell, muscle, lung, and circulation. Much of the modeling is traditional, but the last page mentions the "Koch tree" as a model of the airway tree and remarks on its relation to the Mandelbrot set. This book may not interest you as a mathematician directly, unless you are involved in mathematical modeling of physiological processes; but, like its predecessor D'Arcy Thompson's *On Growth and Form*, it may inspire biology students to study mathematics with you.

THIS MAGAZINE extends its appreciation to Prof. Campbell on his completing 25 years of service as Associate Editor and Reviews Editor. He in turn thanks Assistant Editor Eric Rosenthal for prodigious assistance over the years.

NEWS AND LETTERS

42^{nd} International Mathematical Olympiad

Washington, D.C., United States of America

July 9 and 10, 2001

edited by Titu Andreescu and Zuming Feng

Problems

1. Let ABC be an acute-angled triangle with O as its circumcenter. Let P on line BC be the foot of the altitude from A. Assume that $\angle BCA \geq \angle ABC + 30°$. Prove that $\angle CAB + \angle COP < 90°$.
2. Prove that
$$\frac{a}{\sqrt{a^2+8bc}} + \frac{b}{\sqrt{b^2+8ca}} + \frac{c}{\sqrt{c^2+8ab}} \geq 1$$
for all positive real numbers a, b, and c.
3. Twenty-one girls and twenty-one boys took part in a mathematical competition. It turned out that
   (a) each contestant solved at most six problems, and
   (b) for each pair of a girl and a boy, there was at least one problem that was solved by both the girl and the boy.

   Prove that there is a problem that was solved by at least three girls and at least three boys.
4. Let n be an odd integer greater than 1 and let $c_1, c_2, \ldots, c_n$ be integers. For each permutation $a = (a_1, a_2, \ldots, a_n)$ of $\{1, 2, \ldots, n\}$, define $S(a) = \sum_{i=1}^{n} c_i a_i$. Prove that there exist permutations b and c, $b \neq c$, such that $n!$ divides $S(b) - S(c)$.
5. In a triangle ABC, let segment AP bisect $\angle BAC$, with P on side BC, and let segment BQ bisect $\angle ABC$, with Q on side CA. It is known that $\angle BAC = 60°$ and that $AB + BP = AQ + QB$. What are the possible angles of triangle ABC?
6. Let $a > b > c > d$ be positive integers and suppose
$$ac + bd = (b + d + a - c)(b + d - a + c).$$
Prove that $ab + cd$ is not prime.

Solutions

Note: For interested readers, the editors recommend the *USA and International Mathematical Olympiads 2001*. There many of the problems are presented together with a collection of remarkable solutions developed by the examination committees, contestants, and experts, during or after the contests.

1. Let $\alpha = \angle CAB$, $\beta = \angle ABC$, and $\gamma = \angle BCA$. Let ω be the circumcircle of triangle ABC, and let R denote the circumradius of triangle ABC. Also, let M be the midpoint of side BC. Because $90^\circ > \gamma > \beta$, line AP is closer to C than to B. Then P is on segment CM. Moreover, since triangle ABC is acute, O is inside triangle ABC and triangles BOP, COP, OPM are all nondegenerate. Because $90^\circ > \gamma - \beta \geq 30^\circ$,

$$\sin(\gamma - \beta) \geq \frac{1}{2}. \tag{1}$$

Note that $\angle COB = 2\angle CAB = 2\alpha$ and $\angle PCO = \angle BCO = \angle OBC = (180^\circ - \angle COB)/2 = 90^\circ - \angle CAB = 90^\circ - \alpha$. It suffices to prove that $\angle COP < \angle PCO$, or to proving that $OP > PC$.

Because $\angle AOC = 2\beta$, we have $\angle CAO = 90^\circ - \beta$. Note that in right triangle APC, $\angle CAP = 90^\circ - \gamma$, so $\angle PAO = \angle CAO - \angle CAP = \gamma - \beta$. By (1), $\sin \angle PAO \geq 1/2$. Let N be the foot of perpendicular from O to segment AP. Then $MPNO$ is a rectangle, so $MP/OA = ON/OA = \sin \angle PAO \geq 1/2$, or $2MP \geq OA = OC$. In right triangles OCM and OPM, $OC > CM$ and $OP > MP$. Therefore, $PC - MP = MC - 2MP \leq MC - OC < MC - MC = 0$. We obtain $OP > MP > PC$, as desired.

2. By multiplying a, b, and c by a suitable factor, we reduce the problem to the case when $a + b + c = 1$. Note that the function $f(t) = \frac{1}{\sqrt{t}}$ is **convex** for $t > 0$ as $f''(t) = \frac{3}{4\sqrt{t^5}}$. Thus, by **Jensen's Inequality**, we obtain

$$\frac{a}{\sqrt{x}} + \frac{b}{\sqrt{y}} + \frac{c}{\sqrt{z}} \geq \frac{1}{\sqrt{ax + by + cz}},$$

for all $x, y, z > 0$. Setting $x = a^2 + 8bc$, $y = b^2 + 8ca$, $z = c^2 + 8ab$ in the last inequality, we obtain

$$\begin{aligned}\frac{a}{\sqrt{a^2 + 8bc}} + \frac{b}{\sqrt{b^2 + 8ca}} + \frac{c}{\sqrt{c^2 + 8ab}} &\geq \frac{1}{\sqrt{a^3 + b^3 + c^3 + 24abc}} \\ &\geq \frac{1}{\sqrt{(a + b + c)^3}} = 1,\end{aligned}$$

as

$$\begin{aligned}(a + b + c)^3 &= a^3 + b^3 + c^3 + 3\sum_{\text{cyc}}(a^2b + b^2a) + 6abc \\ &\geq a^3 + b^3 + c^3 + 18\sqrt[6]{a^6b^6c^6} + 6abc = a^3 + b^3 + c^3 + 24abc.\end{aligned}$$

3. Assign each problem a unique letter, and also number the boys $1, 2, \ldots, 21$ and number the girls $1, 2, \ldots, 21$. Construct a 21×21 matrix of letters as follows: in the ith row and jth column, write the letter of any problem that both the ith girl and the jth boy solved—at least one such problem exists by condition (b). If we consider the ith row, each letter in that row corresponds to a problem that the ith girl solved. Since each girl solved at most six problems, each row contains at most 6 distinct letters. Similarly, each column contains at most 6 distinct letters.

We have following key observation: *In each row (resp. column), consider the letters which appear at least three times. At least 11 squares in the row (resp. column) contain one of these letters.* Indeed, there are at most 6 different letters, and they cannot all appear at most twice, since there are $21 > 12$ letters total. So at most 5 different letters appear at most twice, giving a total of at most 10 squares containing letters appearing at most twice. Then at least 11 other squares each contain a letter that appears at least three times.

In the matrix, color all the squares which contain letters appearing at least three times in the same row (resp. column) in red (resp. blue). By the above observation, each row contains at least 11 red squares, so the total number of red squares is at least 21×11. Similarly, each column contains at least 11 blue squares, so the total number of blue squares is at least 21×11. Since there are only $21 \times 21 < 21 \times 11 + 21 \cdot 11$ total squares, some square is colored both red and blue. Because the letter in this square appears in three different columns and three different rows, at least three boys and three girls solved the corresponding problem. Thus, we find the problem satisfying the desired property.

4. Let $\sum_a$ denote the sum over all $n!$ permutations $a = (a_1, a_2, \ldots, a_n)$. We compute $\sum_a S(a)$ modulo $n!$ in two ways, one of which assuming that the desired conclusion is false, and reach a contradiction.

Suppose, for the sake of contradiction, that the claim is false. Then each $S(a)$ must have a different remainder mod $n!$. Since there are exactly $n!$ such permutations a, there exists exactly one permutation a such that $S(a) \equiv s \pmod{n!}$ for each $s = 1, 2, \ldots, n!$. Since $n > 1$, $n!$ is even and $n! + 1$ is odd. Hence,

$$\sum_a S(a) \equiv \sum_{s=1}^{n!} s \equiv \frac{n!}{2} \cdot (n! + 1) \pmod{n!},$$

or

$$\sum_a S(a) \equiv \frac{n!}{2} \pmod{n!}. \tag{1}$$

On the other hand, for $i, k \in \{1, \ldots, n\}$, we have $a_i = k$ in exactly $(n-1)!$ permutations a. Thus, for $1 \le i \le n$,

$$\sum_a a_i = (n-1)!\,(1 + 2 + \cdots + n) = n! \cdot \frac{n+1}{2}.$$

Hence,

$$\sum_a S(a) = \sum_a \sum_{i=1}^{n} c_i a_i = \sum_{i=1}^{n} \sum_a c_i a_i = \sum_{i=1}^{n} \left(c_i \sum_a a_i \right).$$

Because $n + 1$ is even, $n!$ divides $\sum_a a_i = n! \cdot \frac{n+1}{2}$ for each i. It follows that $n!$ divides $\sum_a S(a)$, contradicting (1). Therefore, the initial assumption was false, and there *do* exist distinct permutations b and c such that $n!$ is a divisor of $S(b) - S(c)$.

5. Let $\angle ABC = 2x$ and $\angle BCA = y$. Then $\angle ABQ = \angle QBC = x$ and $\angle CAB + \angle ABC + \angle BCA = 60° + 2x + y = 180°$, so

$$y = 120° - 2x. \tag{1}$$

Extend segment AB through B to R so that $BR = BP$, and construct S on ray AQ so that $AS = AR$.

We claim that points B, P, S are collinear. Because $BR = BP$, triangle BPR is isosceles with base angles

$$\angle BRP = \angle RPB = (180° - \angle PBR)/2 = x = \angle QBP. \tag{2}$$

Note that $AS = AR$ and $\angle RAS = \angle BAC = 60°$, implying that triangle ARS is equilateral. Since line AP bisects $\angle RAS$, R and S are symmetric with respect to line AP. Thus,

$$PR = PS \tag{3}$$

and $\angle ARP = \angle PSA$, or $\angle BRP = \angle PSQ$. By (2), we have

$$\angle QBP = \angle BRP = \angle PSQ. \tag{4}$$

Because $AQ + QS = AB + BR = AB + BP = AQ + QB$, $QS = QB$. Hence, triangle BQS is isosceles with

$$\angle BSQ = \angle QBS. \tag{5}$$

Now, assume to the contrary that triangle is BPS is nondegenerate. Then either $AC < AS$ or $AC > AS$. In either case, combining (4) and (5) gives $\angle PBS = |\angle QBP - \angle QBS| = |\angle PSQ - \angle BSQ| = \angle PSB$, that is, triangle PBS is isosceles with $PB = PS$. By (3), it follows that $PB = PS = PR$. Hence, triangle BPR is equilateral. But then $\angle ABC = 180° - \angle CBR = 120°$, and by (1), $y = 0°$, which is absurd. Therefore, our assumption was wrong and B, P, S are collinear. Consequently, $S = C$.

Since $S = C$, by (1) and (5), we obtain $x = y = 120° - 2x$, or $x = 40°$. Therefore, $\angle ABC = 80°$, $\angle BCA = 40°$, and $\angle CAB = 60°$.

6. Let $x = b + d + a - c$. It is clear that $x > 1$. We have $c \equiv a + b + d \pmod{x}$ and $d \equiv c - a - b \pmod{x}$. These congruences, combined with the given condition, yield

$$0 \equiv ac + bd \equiv a(a + b + d) + bd \equiv (a + b)(a + d) \pmod{x}$$

and

$$0 \equiv ac + bd \equiv ac + b(c - a - b) \equiv (a + b)(c - b) \pmod{x}.$$

Hence, $x \| (a + b)(a + d)$ and $x \| (a + b)(c - b)$.

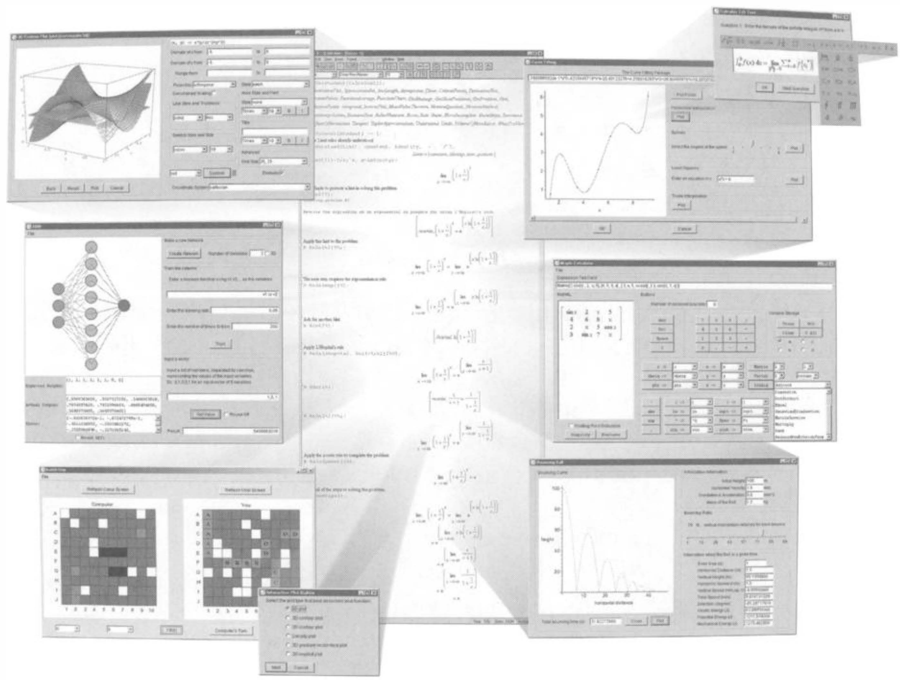
Because $a + b > (a + b) - (c - d) = x$ and $2x = 2[a + (b - c) + d] > 2a > a + b$, $a + b$ is not divisible by x. Thus, there is a prime p that divides each of x, $(a + d)$, and $(c - b)$. To finish, we only need to prove that p is a proper divisor of $ab + cd$. In fact, $ab + cd > a + d \geq p$ and

$$p \| (a + d)b + (c - b)d = ab + cd,$$

as desired.

Introducing Maple® 8

command the brilliance of a thousand mathematicians



Continuing the tradition of innovation, mathematical leadership and maximum value.

Revolution Maple® 8 features Maplets™ — a true innovation for mathematical software. Maplets lets you easily create custom applet-style user interfaces for Maple® without complicated UI programming. Built on the flexibility of Java™, Maplets are ideal for sharing your Maple work with students, colleagues, or anyone who may not need or want the full power of the worksheet environment.

Evolution Maple 8 offers a sweeping range of new mathematical functionality and environment enhancements including a new package to explore and demonstrate concepts in calculus, a library of over 13,000 scientific constants, numerical solutions to PDEs with boundary conditions, calculus of variations, code generation and connectivitity to Java, spell-checking, greater worksheet display control, a dialog-based plot builder, and much more.

For more information, and details on special upgrade pricing, call **1/800 267.6583**. Outside North America, call **1/519 747.2373** or visit **www.maplesoft.com/sales** for your local distributor.

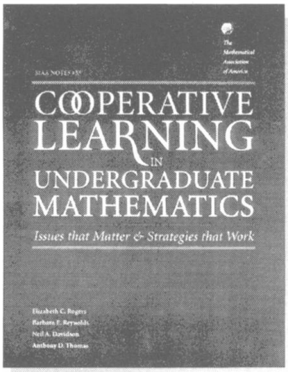
Waterloo Maple
ADVANCING MATHEMATICS

WWW.MAPLESOFT.COM | WWW.MAPLEAPPS.COM | INFO@MAPLESOFT.COM | NORTH AMERICAN SALES 1.800.267.6583

© 2002 Waterloo Maple Inc. Maple is a registered trademark of Waterloo Maple Inc. Maplets is a trademark of Waterlo Maple Inc. All other trademarks are property of their respective owners .

The Mathematical Association of America

Cooperative Learning in Undergraduate Mathematics: Issues that Matter and Strategies that Work



Elizabeth C. Rogers, Barbara E. Reynolds, Neil A. Davidson, and Anthony D. Thomas, Editors

Series: MAA Notes

This volume offers practical suggestions and strategies both for instructors who are already using cooperative learning in their classes, and for those who are thinking about implementing it. The authors are widely experienced with bringing cooperative learning into the undergraduate mathematics classroom. In addition they draw on the experiences of colleagues who responded to a survey about cooperative learning which was conducted in 1996-97 for Project CLUME (Cooperative Learning in Undergraduate Mathematics Education).

The volume discusses many of the practical implementation issues involved in creating a cooperative learning environment:

- how to develop a positive social climate, form groups and prevent or resolve difficulties within and among the groups.
- what are some of the cooperative strategies (with specific examples for a variety of courses) that can be used in courses ranging from lower-division, to calculus, to upper division mathematics courses.
- what are some of the critical and sensitive issues of assessing individual learning in the context of a cooperative learning environment.
- how do theories about the nature of mathematics content relate to the views of the instructor in helping students learn that content.

The authors present powerful applications of learning theory that illustrate how readers might construct cooperative learning activities to harmonize with their own beliefs about the nature of mathematics and how mathematics is learned.

In writing this volume the authors analyzed and compared the distinctive approaches they were using at their various institutions. Fundamental differences in their approaches to cooperative learning emerged. For example, choosing Davidson's guided-discovery model over a constructivist model based on Dubinsky's action-process-object-schema (APOS) theory affects one's choice of activities. These and related distinctions are explored.

A selected bibliography provides a number of the major references available in the field of cooperative learning in mathematics education. To make this bibliography easier to use, it has been arranged in two sections. The first section includes references cited in the text and some sources for further reading. The second section lists a selection (far from complete) of textbooks and course materials that work well in a cooperative classroom for undergraduate mathematics students.

Catalog Code: NTE-55/JR 250 pp., Paperbound, 2001 ISBN 088385-166-0 List: $ 31.95 MAA Member: $25.95

Name__________ Credit Card No.__________

Address__________ Signature__________ Exp. Date____/___

City__________ Qty_____ Price $_____ Amount $_____

State_____ Zip _____ Shipping and Handling $_____

Phone __________ **Catalog Code: NTE-55/JR** Total $_____

Shipping and Handling: USA orders (shipped via UPS): $3.00 for the first book, and $1.00 for each additional book. Canadian orders: $4.50 for the first book and $1.50 for each additional book. Canadian orders will be shipped within 2-3 weeks of receipt of order via the fastest available route. We do not ship via UPS into Canada unless the customer specially requests this service. Canadian customers who request UPS shipment will be billed an additional 7% of their total order. Overseas Orders: $4.50 per item ordered for books sent surface mail. Airmail service is available at a rate of $10.00 per book. Foreign orders must be paid in US dollars through a US bank or through a New York clearinghouse. Credit card orders are accepted for all customers. **All orders must be prepaid with the exception of books purchased for resale by bookstores and wholesalers.**

Order Via:

Phone: **1 (800) 331.1622**
Fax: **(301) 206.9789**
Mail: **Mathematical Association of America**
PO Box 91112
Washington, DC 20090-1112
Web: **www.maa.org**

CONTENTS

ARTICLES

163 Golden, $\sqrt{2}$, and π Flowers: A Spiral Story, *by Michael Naylor*

173 Pillow Chess, *by Grant Cairns*

187 Doubly Recursive Multivariate Automatic Differentiation, *by Dan Kalman*

NOTES

203 Avoiding Your Spouse at a Party Leads to War, *by Marc A. Brodie*

209 Bernoulli on Arc Length, *by Victor H. Moll, Judith L. Nowalsky, Gined Roa, and Leonardo Solanilla*

214 Proof Without Words: A Line through the Incenter of a Triangle, *by Sidney H. Kung*

215 Finite Groups That Have Exactly n Elements of Order n, *by Carrie E. Finch, Richard M. Foote, Lenny Jones, and Donald Spickler, Jr.*

219 Rootless Matrices, *by Bertram Yood*

223 Lots of Smiths, *by Patrick Costello and Kathy Lewis*

PROBLEMS

227 Proposals 1648–1652

228 Quickies 921–922

228 Solutions 1623–1627

233 Answers 921–922

REVIEWS

234

NEWS AND LETTERS

236 42nd Annual International Mathematical Olympiad

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