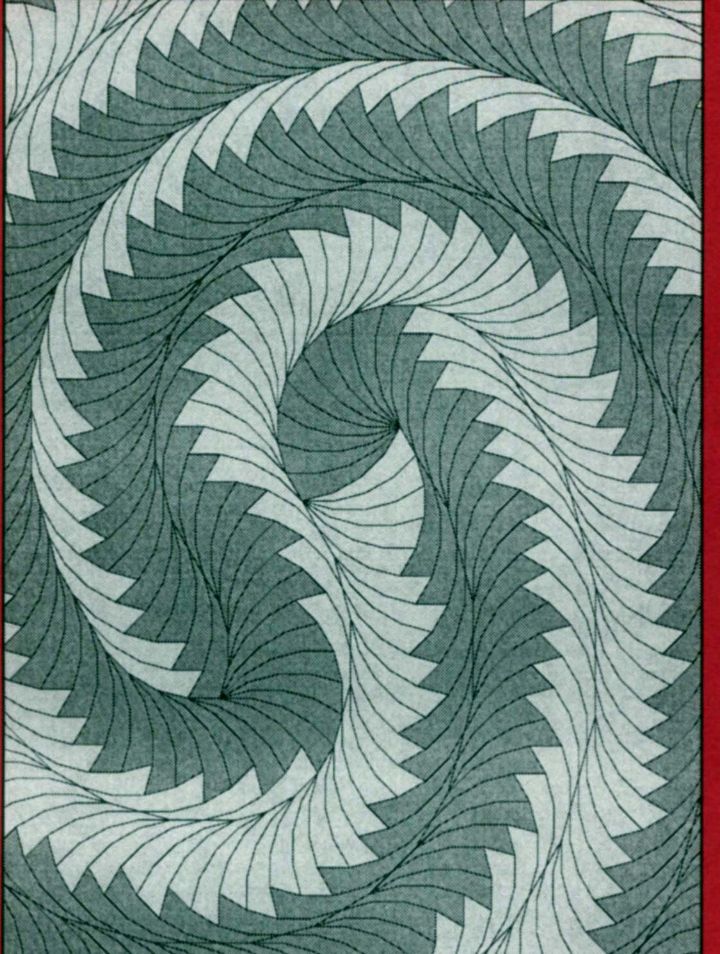
Vol. 73, No. 5, December 2000

MATHEMATICS MAGAZINE



- Odd Spiral Tilings
- Thales' Shadow
- The Solitaire Army Reinspected

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. 71, pp. 76–78, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Send new manuscripts to Frank Farris, Editor-Elect, Department of Mathematics and Computer Science, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053-0290. Manuscripts should be laser-printed, with wide line-spacing, and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit three copies and keep one copy. In addition, authors should supply the full five-symbol Mathematics Subject Classification number, as described in *Mathematical Reviews*, 1980 and later. Copies of figures should be supplied on separate sheets, both with and without lettering added.

AUTHORS

Béla Csákány and **Rozália Juhász** received their bachelor's degrees from the University of Szeged in 1955, resp. 1962. Béla did his graduate work at the Moscow State University with A. G. Kurosh as mentor, and Rozália did hers in Szeged with L. Lovász. They received their degrees (which were equivalent at that time to the Ph.D.) in 1962, resp. 1978. Except for semesters at the University of Southwestern Louisiana and the Université de Montréal, they have always been faithful faculty members of the University of Szeged. His favorite topic is universal algebra (varieties and clones); her main interest is in combinatorial geometry (Ramsey-type problems). Beyond their textbooks in algebra and geometry, Béla is the author of the book *Diszkrét matematikai játékok* (games = játékok in Hungarian), which includes, e.g., the Reiss theory of Peg Solitaire. Although Béla and Rozália were married some 36 years ago, this article is their first joint mathematical enterprise: their children became an engineer and an ophthalmologist (respectively, of course).

Lothar Redlin was born near Freiberg, Germany, grew up on Vancouver Island in British Columbia, and received his Ph.D. from McMaster University in Hamilton, Ontario, in 1978. Since 1985 he has taught at Penn State's Abington College. **Ngo Viet** was born in Hue, Vietnam, and was educated at the University of Minnesota and the University of California at Berkeley, where he received his Ph.D. in 1984. He has taught at California State University, Long Beach for the past 15 years. **Saleem Watson** was born in Zeitoun, Egypt, attended high school in Oshawa, Ontario, received his B.A. at Andrews University in Michigan, and his Ph.D. at McMaster University in 1978. He teaches at California State University, Long Beach, where he first joined the faculty in 1980. Thales' measurement using shadows of the height of the Great Pyramid at Giza is sometimes mentioned in texts on the history of mathematics. Teaching about this feat in a math history class led the authors to wonder whether this method really works in practice.

Daniel L. Stock received his M.S. in computer engineering and his B.S. in mathematics from Case Western Reserve University in 1980. He spent the rest of the 1980's starting studies toward a Ph.D. in computer science, being lured away to write an Ada compiler, and starting the article that finally appears here. He spent most of the 1990's being a househusband, studying *Scrabble*, and finding another Ada expert and tilings fan to co-author the article. **Brian Wichmann** has recently retired from the National Physical Laboratory in the UK. He is a software engineer who designed part of the Ada programming language and worked on the revision of the language to ensure its suitability for high integrity systems.

Vol. 73, No. 5, December 2000

MATHEMATICS MAGAZINE

The *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 the MATHEMATICS MAGAZINE to an individual member of the Association is $16 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $64. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount; and new members receive a 40% dues discount for the first two years of membership.) The nonmember/library subscription price is $68 per year.
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 Ms. Elaine Pedreira, Advertising Manager, the Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.

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

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

Printed in the United States of America

ARTICLES

Odd Spiral Tilings

DANIEL L. STOCK
3614 Kings Mill Run
Rocky River, OH 44116

BRIAN A. WICHMANN
5 Ellis Farm Close
Mayford, Woking, Surrey, GU22 9QN
United Kingdom

Introduction

Despite their visual appeal, spiral tilings have attracted only a small number of mathematical papers since they were first mentioned by Voderberg [**1**]. Of interest in Figure 1 are the five spiral arms emanating from the center. Indeed, in their seminal work on tilings and patterns [**2**], Grünbaum and Shephard leave open the existence of spiral tilings with any odd number of arms beyond three. We shall demonstrate techniques for constructing spiral tilings with *any* odd number of arms. Together with Goldberg's technique to construct even-armed spiral tilings [**3**], this means that we can produce spiral tilings with any number of arms.

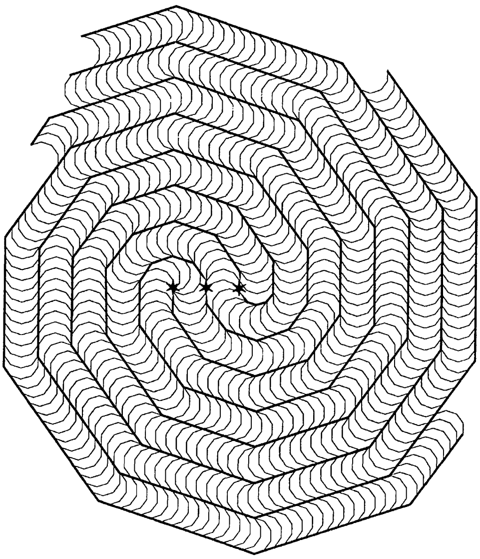


FIGURE 1
A five-armed spiral tiling with the arm separators highlighted and their end-points shown as stars.

Loosely speaking, a *spiral tiling* is a tiling by congruent polygons that has a spiral appearance. As Grünbaum and Shephard point out, however, it can be tricky to define a spiral tiling more precisely. Previous writings on the subject either give no definition

[**1**, **3**, **4**, **5**, **6**, **7**], or mention a definition so broad that it includes even the checkerboard tiling by congruent squares [**2**].

Preliminaries

Rather than attempt to provide a general definition of spiral tilings, we shall define one large class of tilings by congruent non-convex polygons whose spiral nature follows from the definition. This class includes Voderberg's tiling, Grünbaum and Shephard's remarkable tiling [**2**, frontispiece], most of the spiral tilings produced by Goldberg's method (including many samples with any even number of arms), and, we believe, most of the other spiral tilings in the literature.

First, we clear up a matter of terminology. The terms *vertex* and *edge* can each refer to two different things in discussions of tilings by polygons [**2**]. In particular, *vertex* could refer either to a vertex of the tiling (i.e., a point where three or more tiles meet) or to a vertex of a polygon (i.e., a corner). The term *edge* can refer to the curves that connect either type of vertex. Following Grünbaum and Shephard, we will use the term "vertex" for a point where three or more tiles meet, and "edge" for a curve in the tiling that connects such vertices. Each edge of a tiling is common to exactly two tiles.

To motivate the definition of our class of tilings, note that each tile in the spiral tiling in Figure 1 fits neatly into a cavity provided by neighbors in the same arm. Hence, concavity appears to be essential to the spiral appearance. Specifically, arms of the tiling are separated by edges that are not inside the convex hull of tiles that border them. Such edges are important: we say that an edge E of a tiling T by congruent simple polygons is a *separating edge* of T if the interiors of the convex hulls of the two tiles that have E in common are disjoint. Figure 1 shows the separating edges for the tiling in bold. The separating edges form simple curves that spiral out and separate the arms of the tiling.

We shall use this idea to define our class of spiral tilings. Specifically, for any natural number n, we define a tiling T by congruent simple polygons to be a *well-separated spiral tiling* with *n arms* if there exist n semi-infinite simple curves (topological rays), called the *arm separators* of T, such that:

1. the union of the arm separators of T is equal to the union of T's separating edges;
2. each pair of arm separators is disjoint, except possibly for a common endpoint;
3. each arm separator winds infinitely often around its endpoint.

There are five arm separators in the spiral tiling of Figure 1, so this is a well-separated spiral tiling with five arms. The stars show where we might place the endpoints of the arm separators. Two of the stars have two arm separators emanating from them. Either of these two stars could just as well have been placed at any point along the arm separators that emanate from it.

Construction

Goldberg [**3**] provided a method to create spiral tilings with any even number of arms. He does this by taking an appropriate radially symmetric tiling, sliding one half-plane of the tiling with respect to the other half-plane, and modifying the result to give an aesthetically pleasing spiral tiling. The further we slide the half-planes, the more arms we get in the result—but there is always an even number of them, due to symmetry considerations. We shall employ a similar sliding technique, but with an additional trick that yields an extra arm.

To illustrate our construction we shall use the *reflexed decagon* tile given by Simonds **[6]**. FIGURE 2 shows this tile. Note that the chord about which the decagon is reflected has length d, which is the diameter of the incircle of the original decagon; s denotes the length of one of its sides. In general, for any integer m greater than 2, we can construct Simonds' *reflexed* $2m$*-gon* from a regular $2m$-gon as follows. Consecutively label the corners of the regular $2m$-gon as $A_1, \dots, A_{2m}$, then reflect the sides $A_1A_2, \dots, A_{m-1}A_m$ across the line A_1A_m. The $m-1$ reflected sides and the unreflected $m+1$ sides of the original $2m$-gon make up the reflexed $2m$-gon.

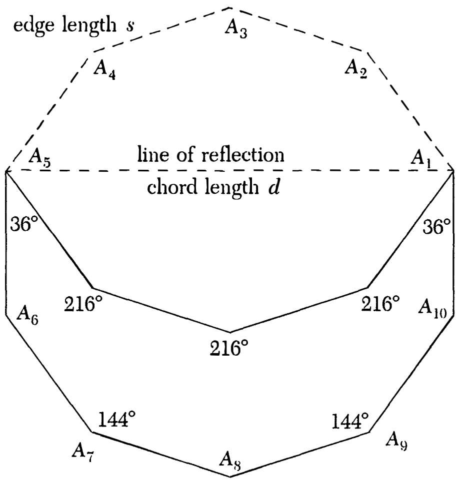


FIGURE 2
A reflexed decagon, constructed by reflecting four consecutive sides of a regular decagon.

We shall begin our construction with Simond's attractive, radially symmetric tiling, part of which forms the top half of FIGURE 3. We can consider this tiling to be the union of an infinite number of concentric decagonal annuli that surround a central tiled decagon. We will say that one of these annuli has size i if i tiles in the annulus meet each outer side of the decagon boundary of the annulus. In the top of FIGURE 3, the ith annulus out from the center (marked with a star) of the tiling has size $2i$. The annuli of sizes 4, 8, 12, and 18 are shaded to make it easier to see what happens as our construction progresses.

To get our spiral tilings, we will use a cut, shift, and paste method to build spiral arms out of pieces of annuli of different sizes. If we only used annuli of even sizes, it turns out we would end up with an even number of arms. So we need some annuli of odd sizes. For i greater than 2, these annuli can be constructed from rows of decagons as in the previous case, and can be assembled to form the tiling part of which is shown in the bottom of FIGURE 3. This tiling has a singular regular decagon in the middle; the ith annulus out from this central decagon has size $2i+1$. We shall consider the singular decagon to be a degenerate annulus of size 1. In this bottom half, the annuli of sizes 5, 11, 15, and 19 are shaded.

The trick is to use the central decagonal "hole" of FIGURE 3 to start an "odd" arm. We note that any individual annulus can be replaced by its reflected image so that each of the two tilings partially in FIGURE 3 is just one of an infinite set of comparable tilings.

We will combine halves of FIGURE 3 to make a tiling with an odd number of arms. Each half-plane has some *teeth* sticking out beyond the half-plane and some *toothholes* inside the half-plane. Each tooth or toothhole has as its outer border four

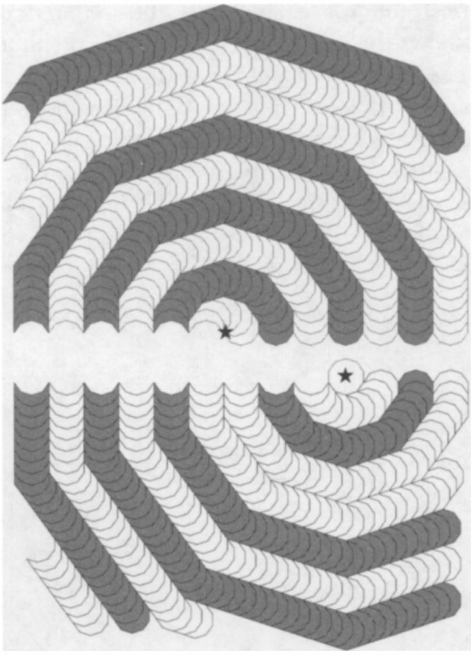


FIGURE 3
Constructing a seven-arm spiral from half-annuli.

contiguous sides of a regular decagon. Each half-annulus of size greater than one has a tooth added at one end and a toothhole indented into the other, forming an *arch*. If we reflect such an arch across a vertical line through the centers, then the tooth and the toothhole are interchanged.

Now, as in Goldberg's construction, we have shifted the lower part of FIGURE 3 to one side (here, to the right) by $nd/2$, where n is the (odd) desired number of arms in the spiral tiling ($n = 7$ in this case). In order to shove the two parts together to get a seven-armed spiral, we reflect some of the arches in each part. We do not reflect any of the arches of size less than or equal to n, as they already mesh wherever they meet another arch of size less than or equal to n; this occurs in the portion of the tiling between the two former centers of symmetry. (It is now safe to admit that we carefully planned our original orientations of the arches in one of the several ways that allow for this to happen!) For i larger than n, we reflect the arch of size i if doing so will make it mesh with the arch of size $i - n$ in its final position; the fact that we shifted by $nd/2$ makes sure that these two arches have some combination of adjacent teeth and toothholes. Now we can push the halves together, and get the seven-armed spiral shown in FIGURE 4.

It remains to show that the tiling we have constructed is a well-separated spiral tiling. To do so, we trace the separating edges through the construction. The separating edges are the decagons that form the boundaries of each annulus. These become corresponding half-decagons in FIGURE 3, if we ignore the edges along the teeth and toothholes for the moment. When we push the two halves of the tiling together, the edges along which the teeth and toothholes mesh are generally not separating edges, since they fall inside the convex hull of the tile that forms the toothhole. The one exception is the toothhole that was originally formed from the starred decagon in the bottom half of FIGURE 3.

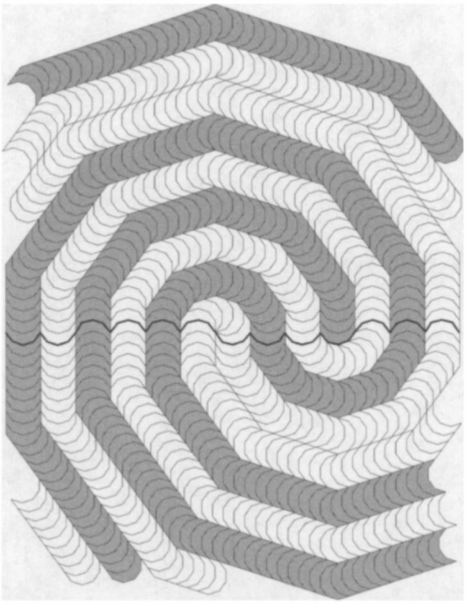


FIGURE 4
The completed seven-arm spiral showing the join of the two halves.

Thus, the separating edges of the completed FIGURE are precisely those edges that were separating edges in the appropriate halves, including the four edges that are the vestiges of the central starred decagon of the bottom half of FIGURE 3. In other words, the separating edges are precisely the half-decagons that form the outsides of the arches. Therefore, by our construction, we can construct one arm separator for each j in the range 1 to n using the halves of decagons with side lengths $js, (j+n)s, (j+2n)s, \dots$. The arm separator corresponding to $j=n$ does not share an endpoint with any other arm separator—it "dead ends" into the center point of the degenerate arch of size 2. Each other arm separator does share an endpoint (the one corresponding to $j=k$ shares an endpoint with the one corresponding to $j=n-k$, for k in the range 1 to $n-1$; these endpoints lie on the line between the center of the odd arches and the center of the even arches). Since each arm separator is a topological ray that winds infinitely often around its endpoint, and since the arm separators meet only at endpoints, the tiling is indeed a well-separated spiral tiling with n arms.

Of course, similar constructions could be made with tiles other than reflexed decagons; any reflexed $2m$-gon, with m at least 3, will do. Larger values of m tend to give rather more "convincing" spirals. Empirically, factors that appear to contribute to making a spiral tiling more "convincing" include having "more concave" tiles (i.e., each tile taking up a smaller portion of the area of its convex hull) and having less sharp corners on the arm separators. Increasing m improves both of these factors.

Extensions

Many other odd-armed spiral tilings are possible, using similar construction methods but starting with different radially symmetric tilings. For example, we could use reflexed m-gons for odd m as in **[5, 6, 7]**, yielding a rather different effect. Or we could use the "versatile" of Grünbaum and Shephard **[4]**, as in FIGURE 5.

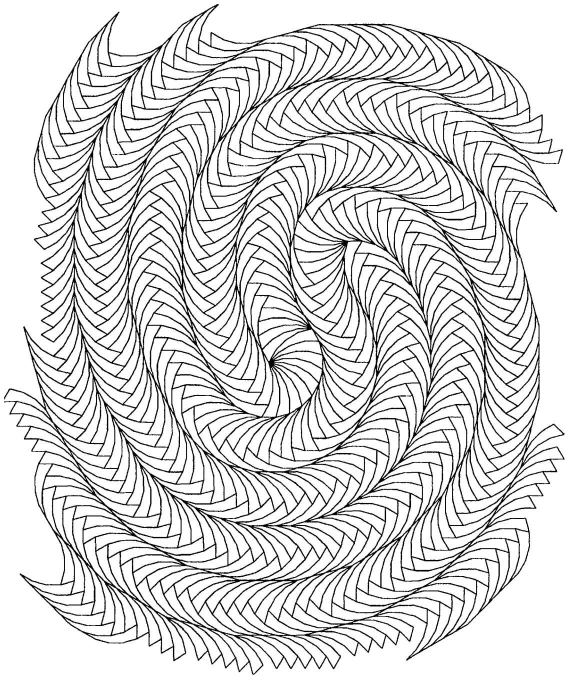


FIGURE 5
A five-armed spiral with versatiles.

In all of these cases (except possibly for FIGURE 5), the corresponding one-armed spiral tiling was previously known. Thus, as one anonymous referee suggested, we could alternatively have constructed our odd-armed spiral tilings by starting from any of these known one-armed spiral tilings, pulling it into two halves, and shifting and reflecting as in our current construction.

It is clear that some well-separated tilings can be altered to produce other tilings with a spiral appearance. Some examples appear elsewhere [**2**]; another is the tiling of FIGURE 6. The figure exemplifies two common techniques of altering a well-separated spiral tiling while maintaining much of the visual effect: removing some or all of the separating edges, and subdividing tiles. By considering these techniques, we can arrive at a definition that appears to fit nearly all of the published spiral tilings to date. Specifically, we say that a tiling U by congruent simple polygons is a *derived spiral tiling* if for some well-separated spiral tiling T, the non-separating edges of U cover all the non-separating edges of T.

One referee suggested another interesting line of analysis: to classify the well-separated tilings that can be constructed from reflexed regular polygons. The referee provided two striking tilings which are constructed from a different reflexed decagon, as shown in FIGURES 7 and 8. Here, the polygon has a smaller indentation and therefore the visual spiral property is perhaps less marked. Defining a reflexed (n, k)-gon as a regular n-gon in which k consecutive edges are reflected, then some interesting questions to ask are:

1. Which reflexed (n, k)-gons tile the plane, and which of those admit well-separated spiral tilings?
2. For the odd armed-tilings constructed here, there is just one unbounded arm in one direction, while the others are unbounded in both directions; in general, how many arms of each type can be produced?

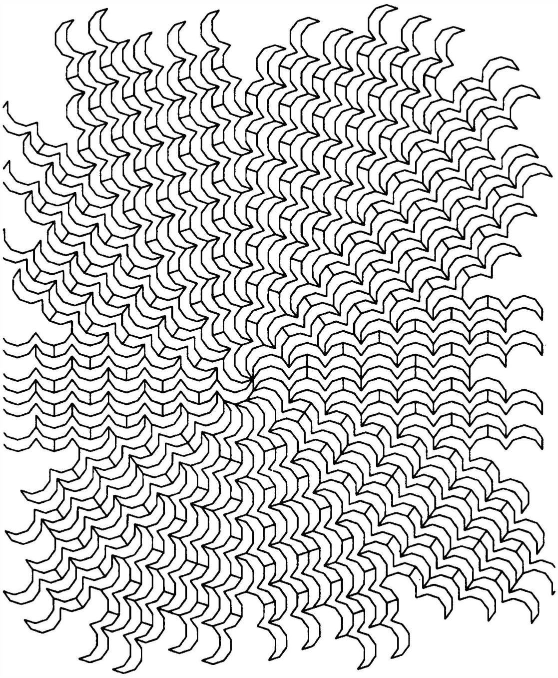


FIGURE 6
A spiral tiling with $1\frac{1}{2}$ reflexed decagons.

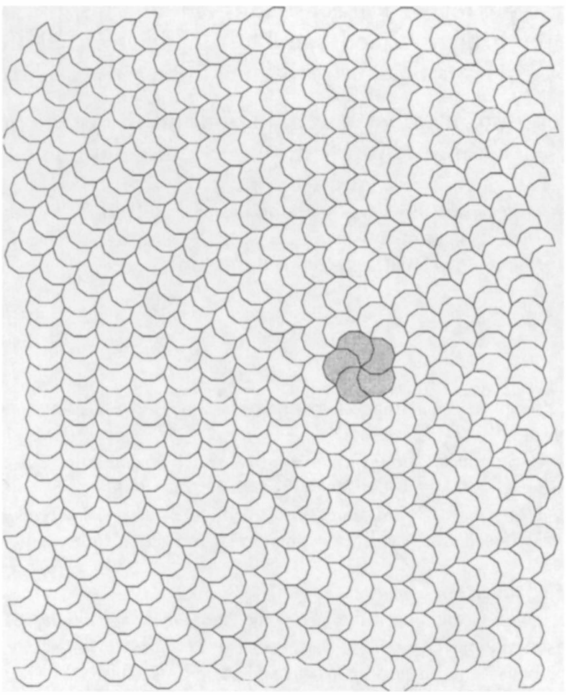


FIGURE 7
Five-armed spiral with a reflexed (10, 3)-gon.

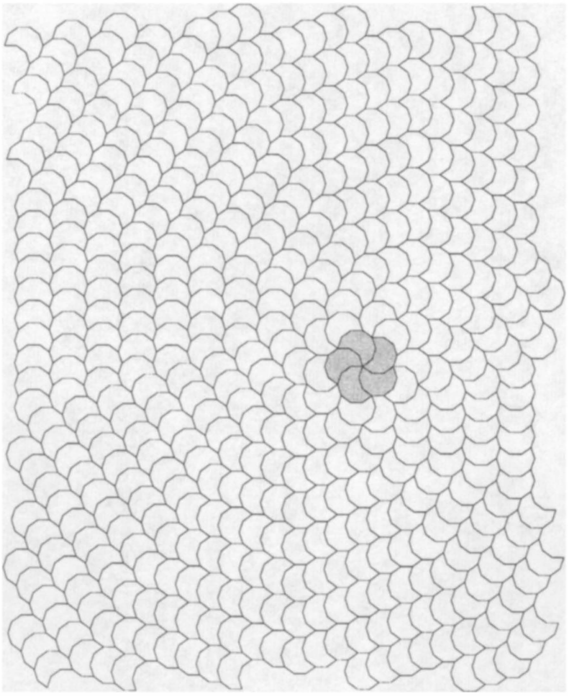


FIGURE 8
Ten-armed spiral with a reflexed (10, 3)-gon.

3. Is it possible to have a spiral tiling with an infinite number of arms? More generally, what are the topological constraints on the arms?
4. For what n is there an n-armed spiral tiling in which all the separators are congruent? FIGURE 7 shows that such a 5-armed tiling is possible.

The Inn at Honey Run in Millersburg, Ohio, displays a quilt, produced by Mary Miller and Ruth Schabach in the Amish style, but to the design of FIGURE 1.

Acknowledgment. The work of Mr. Stock was supported in part by R. R. Software, Inc. Thanks to three anonymous referees of earlier versions of this paper for many helpful suggestions. Thanks also to William Korbelak, John R. Stock, Marge Stock, and Ruth Stock for their assistance.

REFERENCES

1. H. Voderberg, Zur Zelegung der Umgebung eines ebenen Bereiches in kongrunente, *Jahresber. Deutsch. Math.-Verein* 46 (1936), 229–231.
2. B. Grünbaum and G. C. Shephard, *Tilings and Patterns*, W. H. Freeman & Co., New York, NY, 1987.
3. M. Goldberg, Central tessellations, *Scripta Math.* 21 (1955), 253–260.
4. B. Grünbaum and G. C. Shephard, Spiral tilings and versatiles, *Mathematics Teaching* 88 (1979), 50–51.
5. D. R. Simonds, Central tessellations with an equilateral pentagon, *Mathematics Teaching* 81 (1977), 36–37.
6. D. R. Simonds, Untitled note, *Mathematics Teaching* 84 (1978), 33.
7. G. Hatch, Tessellations with equilateral reflex polygons, *Mathematics Teaching* 84 (1978), 32.
8. B. A. Wichmann, An encyclopaedia of tiling patterns, *The Mathematical Gazette* 78 (1994), 265–273. (The collection is shortly to be published by World Scientific on a CD-ROM.)

Thales' Shadow

LOTHAR REDLIN
The Pennsylvania State University
Abington College
Abington, PA 19001

NGO VIET
SALEEM WATSON
California State University, Long Beach
Long Beach, CA 90840

Introduction

History records that in the 6th century B.C. Thales of Miletus measured the height of the great pyramid at Giza by comparing its shadow to the shadow of his staff [**3**]. But there are differing versions of how he may actually have done this [**5**]. We will do some mathematical detective work to explore which version is more likely.

The earliest version is attributed to Hieronymus (4th century B.C.) by Diogenes (2nd century A.D.), who writes:

> "[Thales]... succeeded in measuring the height of the pyramids by observing the length of the shadow at the moment when a man's shadow is equal to his own height."

Thales here observes that when one object casts a shadow equal to its height, then all objects cast shadows equal to their own heights. In FIGURE 1, this means that when $H = S$, then $h = s$. This "equal shadow" phenomenon allows one to measure the height of a tall object by measuring the length of its shadow along the ground. But Thales may have used another method. Plutarch (2nd century A.D.) writes: "Although the king of Egypt admired Thales for many things, he particularly liked the way in which he measured the height of the pyramid without any trouble or instrument." Plutarch continues:

> "[Thales] set up a stick at the tip of the shadow cast by the pyramid, and thus having made two triangles by the sun's rays, he showed that the ratio of the pyramid to the stick is the same as the ratio of the respective shadows."

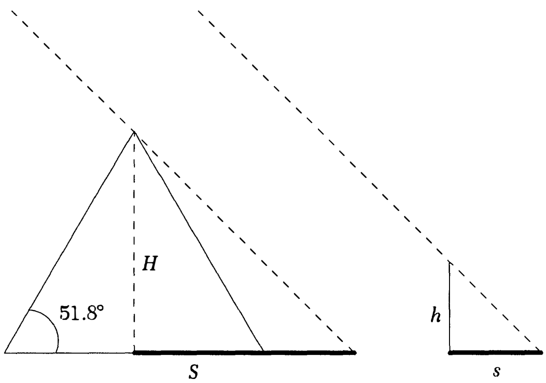


FIGURE 1
Shadows of pyramid and stick.

This more general "ratio" method does not require the shadow of the stick to be equal to its length. In Figure 1 Thales computes H from the ratio $H/h = S/s$. It is likely that Thales knew about such ratios for legs of right triangles, since the Egyptians also had used such techniques in calculations involving pyramids, calling the method *seked* **[5]**. But neither Thales nor the Egyptians had a general theory of similar triangles. The innovation that Thales makes here is his observation that the sun's rays form a right triangle with the stick and its shadow, so that *seked* can be applied to this abstract right triangle.

Both the equal shadow method and the ratio method are simple and elegant. The equal shadow method is particularly elegant because it requires no calculations—the length of the shadow is the height of the object being measured. But the shape of the pyramid presents several difficulties, the simplest being that at certain times Thales' staff casts a shadow but the pyramid casts no shadow at all! However, even when the pyramid does cast a shadow, there are still some practical problems to implementing these methods. Before we examine these problems we consider some recorded information about Thales himself, which may give us a hint about his view of theory and applications.

Thales is reputed to have been the first to put geometry on a logical demonstrative basis **[2]**; textbooks on the history of mathematics refer to him as the first mathematician. Evidently, Thales was more interested in logic and proof than in practical matters. According to legend, he was once walking, intently gazing up at the stars, when he fell into a well. A woman with him exclaimed, "How can you tell what is going on in the sky when you can't even see what is lying at your feet?" So it appears that Thales fit the stereotype of a pure mathematician. But this image must be balanced with his association with the very practical Egyptians. In fact, Thales was one of the first Greeks to travel to Egypt, where it is said he learned geometry and also discovered many propositions himself. Egyptian geometry was a tool to serve practical needs which often required extreme precision. Egyptian monuments, still standing millennia later, are witnesses to the accuracy of their builders. Indeed the great pyramid is thought to have been aligned so perfectly north that its minute deviation from true north is attributed by some scientists to continental drift! So, did Thales actually measure the height of the pyramid or did he merely perform a beautiful thought experiment?

Implementing the ratio method

To use any shadow method one needs to know the length of the shadow as measured from the center of the pyramid. This cannot be done directly since the mass of the pyramid lies between the tip of the shadow and the center. It can be readily done, however, if the shadow is perpendicular to one side of the pyramid as shown in Figure 2a. In that case the length of the shadow to the center of the pyramid is simply the length of the shadow along the ground plus the length of half the side of the pyramid. If the shadow is skew, as in Figure 2b, then calculating the length of the shadow requires the use of the law of cosines, an idea not available to Thales.

So all Thales needed to do was to visit the pyramid one day and wait for the moment when the shadow of the pyramid is perpendicular to one of the sides. Let us see how this could happen. The pyramid is located at 30° N latitude and since the axis of the earth is tilted 23.5° from the celestial pole, it follows that the pyramid is always located above the plane of the ecliptic. Thus the shadow of the pyramid can never lie on its south side. Let's first consider the case when the shadow lies north of the pyramid. Figure 3 shows the situation *at noon*.

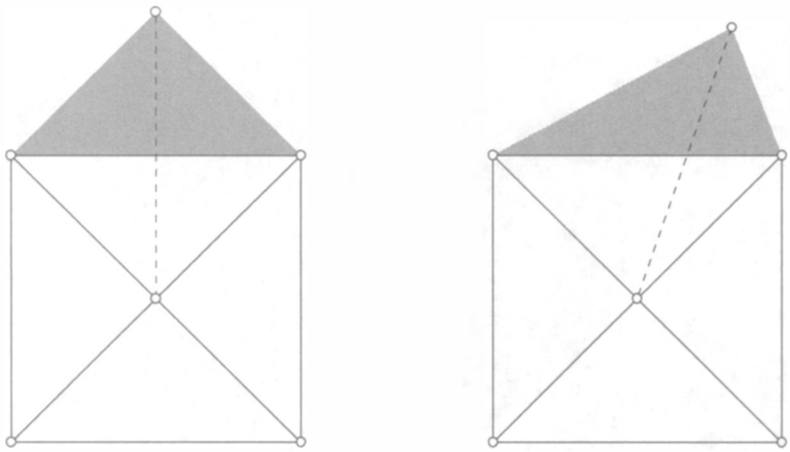

(a)
(b)

FIGURE 2
Top views of pyramid and its shadow.

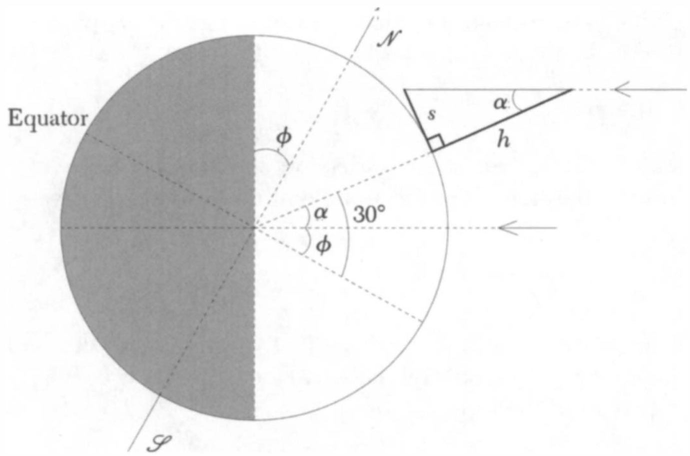


FIGURE 3
Shadow of the pyramid at noon.

During the course of a year, as the earth moves around the sun, the angle ϕ in FIGURE 3 varies between $-23.5°$ and $23.5°$, so the angle $\alpha = 30° - \phi$ varies between $\alpha = 30° + 23.5° = 53.5°$ at the winter solstice, and $\alpha = 30° - 23.5° = 6.5°$ at the summer solstice (see **[4]**). The angle α is called the *zenith distance* of the sun because it is the angle formed by the sun and the zenith. From early March to early October, the zenith distance of the sun is small enough at noon so that the pyramid casts no shadow at all—the sun is so high in the sky that all four faces are illuminated. This is because the faces of the pyramid rise at an angle of 51.8° to the horizontal (or $90° - 51.8° = 38.2°$ from the vertical), so whenever the zenith distance of the sun is less than 38.2°, it will shine on all four faces. For the rest of the year, the shadow is perpendicular to the north side once each day. As FIGURE 3 indicates, this occurs at noon—when the line joining the centers of the sun and earth, the polar axis, and the axis of the pyramid all lie in the same plane.

During the spring and summer months, the shadow is also perpendicular to the west face once in the morning, and to the east face once in the afternoon. To see this we need a three-dimensional view of the earth. In FIGURE 4, we have placed

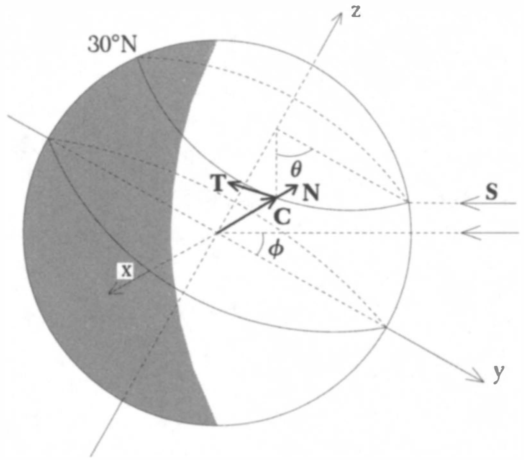


FIGURE 4
Pyramid and its shadow—a 3-dimensional view.

coordinate axes with the origin at the center of the earth, the z-axis oriented along the polar axis, and the y-axis oriented so that the sun lies in the yz-plane.

The sun's rays strike the xy-plane at the angle ϕ. Thus the rays of the sun point in the direction

$$\mathbf{S} = -\cos\phi\mathbf{j} - \sin\phi\mathbf{k}.$$

The circle of latitude at 30°N has radius $R\cos 30° = \sqrt{3}R/2$ (where R is the radius of the earth). Thus the position vector of each point on this circle is given by the vector function

$$\mathbf{C}(\theta) = \frac{\sqrt{3}}{2}R\sin\theta\mathbf{i} + \frac{\sqrt{3}}{2}R\cos\theta\mathbf{j} + \frac{1}{2}R\mathbf{k}$$

where θ is the angle shown in FIGURE 4. Note that $\mathbf{C}(\theta)$ is perpendicular to the sphere of the earth for every θ. The shadow falls east or west when it falls tangent to the latitude circle, that is, in the direction of

$$\mathbf{T}(\theta) = \mathbf{C}'(\theta) = \frac{\sqrt{3}}{2}R\cos\theta\mathbf{i} - \frac{\sqrt{3}}{2}R\sin\theta\mathbf{j}.$$

The axis of the pyramid points in a direction $\mathbf{N}$ normal to the surface of the earth, so at any point on the 30°N latitude circle, we can use $\mathbf{N}(\theta) = \mathbf{C}(\theta)$. The shadow will fall along the tangent to the latitude circle when the tangent $\mathbf{T}$, the normal $\mathbf{N}$, and the rays of the sun $\mathbf{S}$ all lie in the same plane; that is, when $\mathbf{T} \times \mathbf{N} \cdot \mathbf{S} = 0$. Calculating the triple product and simplifying we obtain the condition

$$\cos\theta = \sqrt{3}\tan\phi.$$

Thus the pyramid will cast a shadow perpendicular to its east or west face when it is located at a position corresponding to a value of θ that satisfies this condition. For instance, at the summer solstice, when $\phi = 23.5°$, we get $\theta \approx \pm 41.14°$. This corresponds to times of about 9:15 AM and 2:45 PM. As the summer progresses, these times will fall earlier in the morning and later in the afternoon, with increasingly longer shadows. At the equinoxes, when $\phi = 0°$, we have $\theta = \pm 90°$, so the shadow will be perpendicular to the west and east faces at 6:00 AM and 6:00 PM respectively—sunrise and sunset on these dates when day and night are equal in duration. It is easy to see that in the fall and winter, when $\phi < 0°$, at no time during daylight hours will the shadow be perpendicular to the east or west face.

Can the "equal shadow" method be used?

Now suppose Thales was to use the "equal shadow" method. Having found the times when the shadow is perpendicular to the pyramid, he must now find *among these times* "the moment when a man's shadow is equal to his own height." This moment of "equal shadow" can actually happen on at most four days in any given year, as we now show. First let us consider the north face. If the shadow is perpendicular to the north face, then the length of the shadow of an object of height h is $s = h \tan \alpha$ where α is the zenith distance of the sun at noon as in FIGURE 3. Thus we have the "equal shadow" phenomenon when $\alpha = 45°$. During the course of a year α ranges from 6.5° to 53.5° and back to 6.5°, so it appears that α can equal 45° twice. But α changes in increments of approximately $\frac{1}{4}$ degree per day—actually about 94°/365.25 days $\approx 0.26°$ per day. So α may be equal to 45° at most twice. However it is unlikely that α would ever exactly equal 45°, and the error could be as much as $\pm 0.13°$. Since the pyramids are approximately 480 ft high, this would result in an error of approximately 2.3 ft. Given the precision of the Egyptians, this error seems rather high.

Next, for the equal shadow phenomenon to occur at the east or west faces of the pyramid, the angle between **T** and **S** must be 45°. But then

$$\cos 45° = \frac{\mathbf{T} \cdot \mathbf{S}}{|\mathbf{T}||\mathbf{S}|} = \sin \theta \cos \phi,$$

so $\sin \theta = 1/(\sqrt{2} \cos \phi)$. Since the shadow must also be perpendicular to the east or west faces we also have $\cos \theta = \sqrt{3} \tan \phi$, so

$$1 = \sin^2 \theta + \cos^2 \theta = \frac{1}{2\cos^2 \phi} + 3\tan^2 \phi.$$

Solving we get $\cos \phi = \pm\sqrt{7/8}$, so $\phi \approx 20.7°$. (The other possible solutions for ϕ do not apply here, since ϕ takes on values between $-23.5°$ and 23.5°, and negative ϕ corresponds to fall and winter when the pyramid does not cast a shadow perpendicular to its east or west face.) Thus there are two days in the year when the "equal shadow" phenomenon occurs on the east or west face—the early spring day and the late fall day when $\phi \approx 20.7°$. Again, just as in the case of the north face, it is unlikely that the "equal shadow" and the "perpendicular shadow" phenomena will coincide precisely on these days, so some imprecision is unavoidable here too.

In any case it seems like a lot of trouble for Thales to hang around the pyramid, possibly for months, waiting for the propitious moment when the equal shadow is also perpendicular to a side. But the King said that Thales measured the height of the pyramids without any trouble. Could Thales have possibly used a different equal-shadow method? One that could work on any day?

Can the "equal shadow" method be salvaged?

The main problem with implementing the "equal shadow" method is that the measurement of the shadow is obstructed by the mass of the pyramid. Another method Thales could have tried, which also involves "equal shadow," is suggested in the Project Mathematics video **[1]**. The idea is to wait for the shadow of a man to *lengthen* by an amount equal to his height; at the same time the shadow of the pyramid will lengthen by an amount equal to the height of the pyramid (see FIGURE 5). This method is "dynamic" in the sense that it requires observation of the shadow over a period of time, whereas the first method is "static" in that it relies on observing the shadow at just one instant.

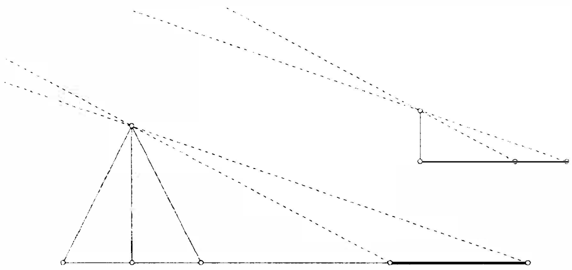


FIGURE 5
Lengthening of the shadow.

The important thing here is that this new method appears to avoid the problem of having to measure the shadow to the center of the pyramid. Indeed all measurements take place well away from the pyramid. So let us again see how Thales could practically implement this new method. Thales must visit the pyramid at a time when the tip of the shadow is away from the base of the pyramid and mark the location of the tip of the shadow. He must then wait for the shadow to lengthen. But Thales discovers, to his dismay, that the shadow doesn't simply lengthen as suggested in Figure 5, it also moves as shown in Figure 6. This presents a new problem, to measure the length of the new longer shadow one must use the center of the pyramid as a reference point. But now we are faced with the same situation as before: the mass of the pyramid obstructs our measurement. So this method isn't really going to work either.

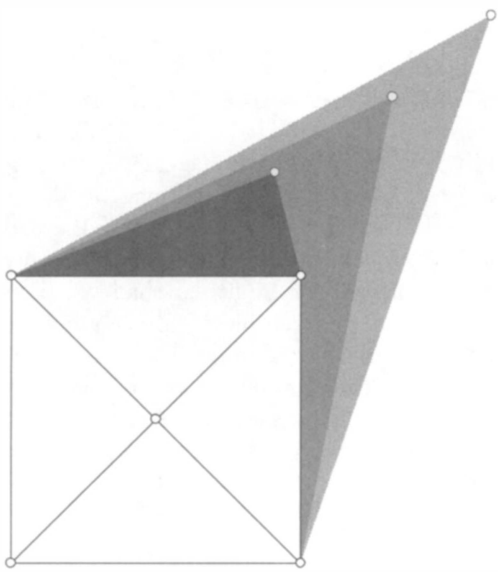


FIGURE 6
Lengthening of the shadow—top view.

What if Thales were to simply connect the tip of the original shadow with the tip of the longer shadow and wait for that distance to equal the height? He would now have a distance on the ground equal to the height. This situation is illustrated in Figure 7. Triangles ABC and abc are similar, as are triangles DAB and dab. From this it is easy to see that $h/H = ab/AB = bc/BC$; since we chose $h = bc$ it follows that $H = BC$. Thus the height of the pyramid can be determined by measuring BC along the ground. This method works, but Thales could not possibly have used it, since he had no knowledge of the proportionality of general similar triangles.

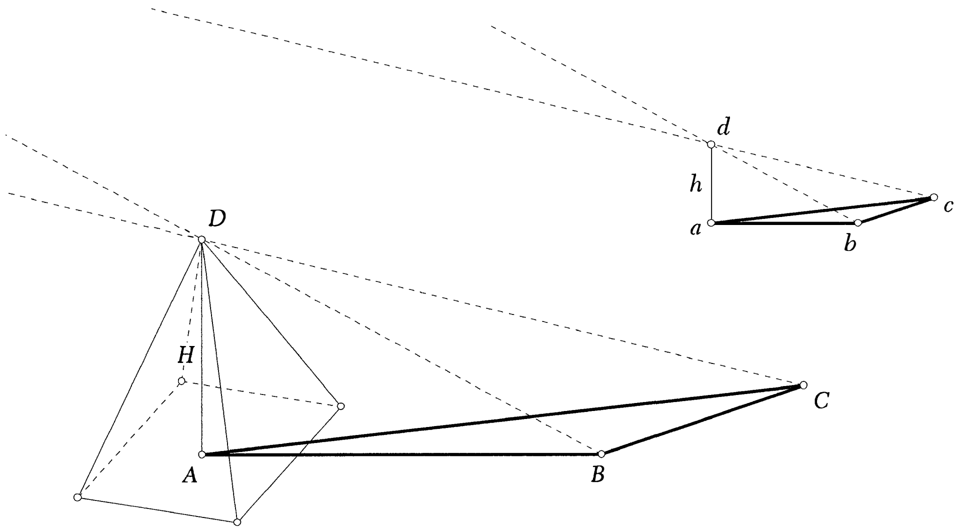


FIGURE 7
A method using similar triangles.

Conclusion

If Thales had used the "equal shadow" method, he would have had to do so on one of only four days in a year, and then only obtained a rough approximation. It seems more likely that he used a "ratio" method. This he would have had an opportunity to do at least once a day for most of the year, and as often as twice on a good day.

Did Thales actually measure the height of the pyramid at all? It is impossible to say for sure, but the *idea* of measuring the height of such a tall object using only its shadow is so beautiful and striking that it overshadows any of its practical applications. This anecdote survives because it encapsulates a great idea that continues to delight and inspire.

REFERENCES

1. Tom Apostol, *Similarity*, Project Mathematics (videotape), California Institute of Technology, Pasadena, CA, 1990.
2. David Burton, *The History of Mathematics: An Introduction*, 3rd edition, McGraw-Hill, New York, NY, 1997.
3. Morris Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, Oxford, UK, 1972.
4. Lucia Carolyn Harrison, *Daylight, Twilight, Darkness, and Time*, Silver Burdett and Company, New York, NY, 1935.
5. Sir Thomas Heath, *A History of Greek Mathematics*, vol. I, Oxford University Press, Oxford, UK, 1921.

The Solitaire Army Reinspected

BÉLA CSÁKÁNY
ROZÁLIA JUHÁSZ
University of Szeged
Aradi vértanúk tere 1
6720 Szeged
Hungary

The solitaire army in new circumstances

Peg Solitaire soldiers—*pegs* for short—move on a plane square lattice. A peg P can jump over a horizontally or vertically neighboring comrade Q onto a free square, removing Q at the same time. The game starts with a configuration of pegs—a solitaire army—and the aim of the moves is usually to obtain another configuration with a prescribed property, e.g., one with a unique peg on a fixed square, or with a peg on a given remote square. For the essentials on Peg Solitaire see the definitive book *Winning Ways* (**[2]**, Chapter 23), where problems of both types are treated in detail. Concerning a problem of the second kind, we have the following basic result of J. H. Conway (see **[6]**, pp. 23–28; **[2]**, pp. 715–717, 728; and **[1]**): *No solitaire army stationed in the southern half-plane can send a scout into the fifth row of the northern half-plane, but an army of* 20 *pegs can send a scout into the fourth row.*

For the proof, to every square s of the plane assign a value $p(s)$ as follows. Let σ be the golden section, i.e., $\sigma = (\sqrt{5} - 1)/2$, so $\sigma + \sigma^2 = 1$. Fix a square s_0 in the fifth row of the northern half-plane. For any square s, let $p(s) = \sigma^k$, where k is the *Manhattan distance* between s_0 and s (this means that s_0 and s are exactly k horizontal or vertical one-square steps apart). Define the *potential of a set of squares* as the sum of values of all squares in this set, and the *potential of an army* as the potential of the set occupied by that army. The potential of any army with a peg standing on s_0 is at least 1. On the other hand, the potential of the infinite army occupying all squares of the (southern) half-plane is exactly 1; we can compute it by observing that values in every column form geometric progressions with quotient σ. The rule of moves implies that no move can increase the potential; it follows that a finite army garrisoned in the southern half-plane cannot reach s_0. This kind of reasoning will occur several times in the sequel; we call it *the Conway argument*. The remaining part of Conway's result can be shown simply by displaying how the army of 20 should be deployed, and how the pegs should move (see below).

In fact, this means that if the front line of a solitaire army looks to the north, then it can advance four rows and no more, just four units of distance both in the Euclidean and the Manhattan sense. Armies, of course, do not always fight under such plain circumstances. Their front line may look to the southwest, for example, in which case the target may be the corner square of the first quadrant. Or, the territory to be scouted may be the "half-encircled" first quadrant; then the army has two perpendicular front lines, one facing north and one facing east. Or we may have two perpendicular fronts, one facing northeast and the other one northwest. FIGURE 1a shows an original northbound army; FIGURES 1b–d show the other possibilities just mentioned.

How far can the scouts be sent in cases (b), (c), and (d)? In what follows we answer these questions. The Conway argument provides upper estimates; we show that they are sharp in every case. We also prove a fact (stated in **[2]** without proof) concerning armies with a single "mounted man." We conclude with two problems about sending scouts into an "encircled ground" (cf. **[9]**).

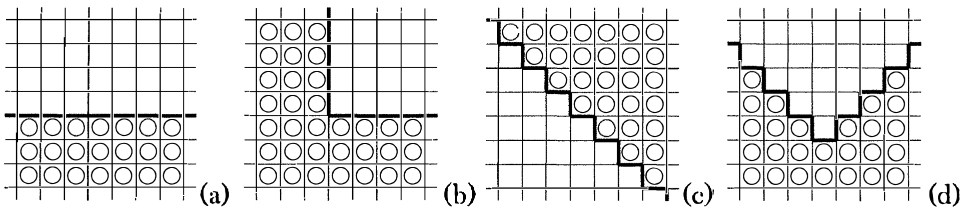


FIGURE 1
Solitaire armies in various circumstances.

Troops and advances

The aim of traditional Peg Solitaire is to evacuate a special battlefield except for one square. For this aim, the method of *packages and purges* is recommended in **[2]**. A *package* is a configuration of pegs which, by an appropriate sequence of moves (a *purge*), can be removed to the last peg. A simple purge is displayed in FIGURE 2a. The numbers indicate the order of moves: peg 2 jumps (onto the square 5) in the second move and returns to its starting place in the fifth move. Exponents 2, 3, etc., denote double, triple, etc., jumps.

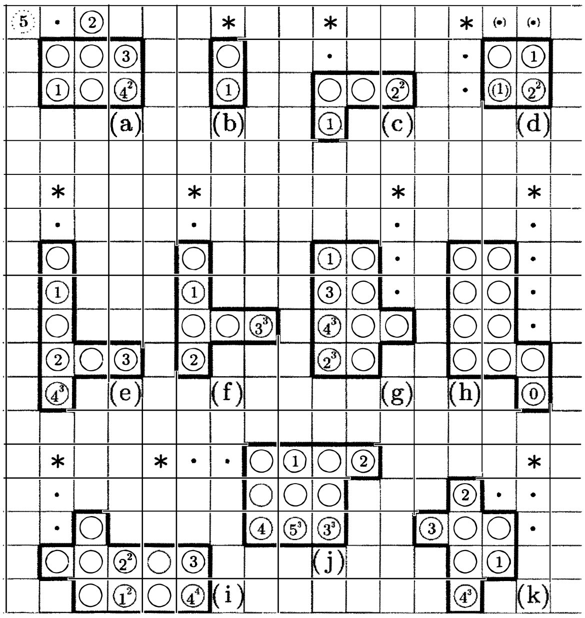


FIGURE 2
Small troops, small advances.

If the goal is to reach some remote square, instead of packages and purges we can apply the similar method of *troops and advances*. A *troop* is, in fact, a package, a member of which can reach a given distant square by a suitable sequence of moves (an *advance*). Here we list and display the troops to be deployed in order to send scouts as far as possible in cases (b), (c), and (d) of FIGURE 1. The smallest troop and advance are a pair of adjoining pegs (a *patrol*) and a single move. The patrol and other small troops are shown in FIGURE 2b, 2c, and 2d, where asterisks denote the square to be reached. We shall also operate with stronger, elite troops. They are given fitting names: *laser guns* (of length 5 and 4 in FIGURE 2e and 2f; their name comes from [**5**]), and *heavy guns* (of length 4 in FIGURE 2g and 2h). Guns of arbitrary length are possible and sometimes even necessary; note that the troop on FIGURE 2c is a laser gun of length 2 and a heavy gun of length 1 at the same time.

Squares marked with dots also must be free of pegs "when the guns begin to shoot." Furthermore, we call troops in FIGURE 2i, 2j, and 2k a *mustang*, a *tomahawk*, and a *halberd*, respectively.

We can compose bigger troops from smaller ones. FIGURE 3a shows that a tomahawk consists, in fact, of units 1 and 2, a laser gun, and a *block*. The squares s_1 and s_2 can be respectively occupied by the advances of unit 1 and unit 2, then a single jump of the peg on s_2 completes the operation. We denote this combined movement by the sequence $12J$. FIGURE 3b shows a *mortar* of length 6, composed from three blocks and a tomahawk. The record of its advance is $1234J^3$, where the exponent indicates that the peg on s_3 concludes the operation by a triple jump. Mortars also may be as long as needed.

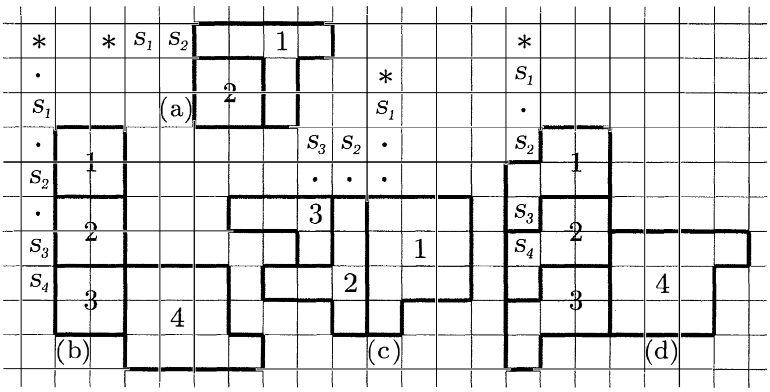


FIGURE 3
Organizing big troops from small ones.

The troop in FIGURE 3c sends a scout into the fourth row (of northern latitude), proving thereby the second part of Conway's result. Now the advance is $123J^2$. FIGURE 3d represents a *centaur*, a special troop we shall deploy soon. Its action is $1234J^3$. This notation may seem ambiguous, but the aim of the whole operation usually resolves what is left undetermined. For example, in the course of the Tomahawk Action (FIGURE 3a) block 2 moves only toward the northwest.

Pebbling and the skew front

Pebbling, a game introduced by M. Kontsevich ([**3**], [**4**], [**7**], [**8**]) is played on a square lattice in the first quadrant. Starting with one pebble in the corner square, each move consists of replacing a pebble by two, one on the north and one on the east neighboring square (FIGURE 4a). Kontsevich's problem was whether it is possible to clear the southwest triangle of 10 squares by pebbling moves (FIGURE 4b). This problem admits an "inverse" formulation as follows: Change the rule of moves so that every move undoes a possible move of ordinary Pebbling: one of the two pebbles lying as in FIGURE 4c advances south or west, and the other is removed from the field. Suppose that all squares of the first quadrant are occupied by a pebble army except for the 10 in the corner. Now the problem is whether this army is able to send a scout to the corner square. Defining the value function q by $q(s_0) = 1$ for s_0 the corner square, and $q(s) = 2^{-k}$ if the Manhattan distance between s and s_0 equals k, we can apply the Conway argument to prove that the pebble army cannot reach s_0. Indeed, the potential of the entire quadrant is 4, that of the 10-square triangle is 13/4, and hence the potential of the given pebble army is 3/4, which is less than $q(s_0)$.

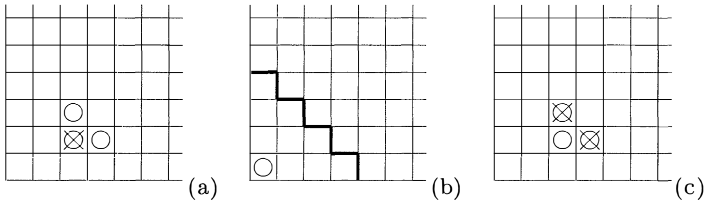


FIGURE 4
Pebbling and unpebbling.

What if, instead of pebbles, our army is recruited from tough solitaire pegs? Then, in order that the potential should not increase by (solitaire) moves, we must return to the old value function $p(s) = \sigma^k$. In this case, the Conway argument shows that an army whose skew front is at Manhattan distance 7 from s_0 cannot reach s_0. Indeed, as $\sigma^k + \sigma^{k+1} + \cdots = \sigma^{k-2}$, the potential (counted by columns) of such an army is $8\sigma^5 + \sigma^6 + \sigma^7 + \cdots = .867\cdots < 1$. If we reinforce the army with an additional (skew) line of pegs, reducing to 6 the distance between the front line and s_0, then its potential grows to $7\sigma^4 + \sigma^3 = 1.257\cdots > 1$ (FIGURE 5a).

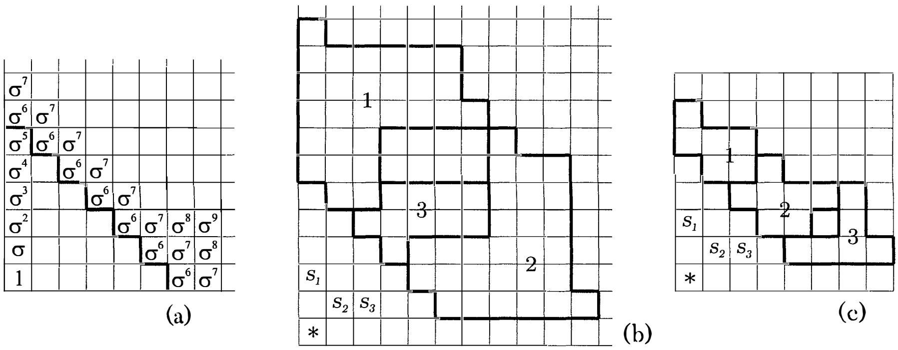


FIGURE 5
Scouting the southwestern half-plane.

Thus, sending a scout at distance 6 from the skew front-line is not prohibited by the Conway argument. It actually can be done by a joint endeavor $123JJ$ of two centaurs and a mustang; see FIGURE 5b.

We can extend the battlefield to the whole square plane without really changing the situation. If s_0 is at distance 7 from the skew front then the potential of the infinite army occupying all the squares on and behind this front equals

$$8\sigma^5 + \sigma^4 + 2(\sigma^7 + \sigma^9 + \cdots) = 8\sigma^5 + \sigma^4 + 2\sigma^6 = .978\cdots < 1.$$

This means that pegs stationed outside the first quadrant cannot provide essential help to their comrades within that quadrant.

We can also restrict the battlefield, for instance to the 8×8 chessboard. Let s_0 be the lower left corner square ("a1" in chess notation). Setting the skew front at distance 6 from s_0, the Conway argument now proves that no scout can reach s_0. However, if the distance is 5, that is, the front-line is the "a6–f1 diagonal," then s_0 is accessible for a troop of 19 pegs, namely the one consisting of a laser gun, a heavy gun, and a halberd, displayed on FIGURE 5c, by the advance $123JJ$. The number 19 cannot be decreased. We omit the simple but tedious proof we know of the latter fact, based on the Reiss theory of Peg Solitaire ([**2**], pp. 708–710).

Two fronts, horizontal and vertical

Now suppose we have the first quadrant to be reconnoitered; all other squares may be held by the solitaire army. The square at distance 5 from both fronts—square $(5,5)$ in short—can be reached by troops we introduced earlier as indicated on FIGURE 6a (cf. FIGURE 3). The advance is $1234'4''56J^5$; here $4'$ converts $4''$ into a tomahawk.

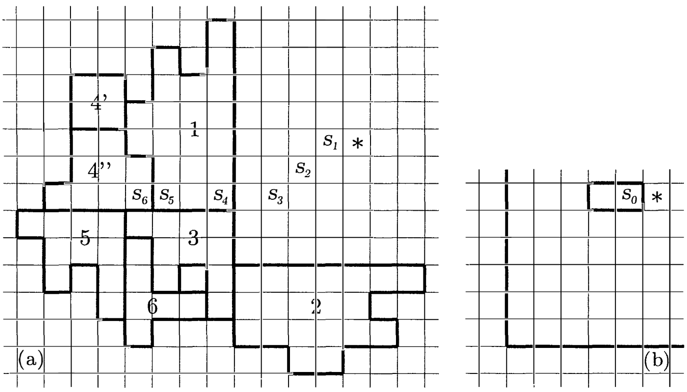


FIGURE 6
Scouting the first quadrant: possible and impossible.

Concerning square (6, 6), we must slightly modify the Conway argument. First let $s_0 = (6,6)$. Then the potential of our army equals $2\sigma - \sigma^8$ (the sum of potentials of the occupied half-planes minus the potential of the third quadrant) which exceeds 1, proving thus nothing. Observe, however, that in order to reach $(6,6)$, we have to send a patrol onto squares $(4,6)$ and $(5,6)$ to perform the final jump (FIGURE 6b). Letting $s_0 = (5,6)$, the potential of the patrol is $1+\sigma$, while that of the all army equals $1+\sigma-\sigma^7$, showing that $(6,6)$ is inaccessible.

What about sending a scout onto $(n,5)$, where $n > 5$? For this aim, tricks like the one applied for the case $(5,5)$ can be devised for $n = 6,7,8,9$. Fortunately, an ingenious observation in [**5**], which we call *the Eriksson–Lindström lemma*, makes them unnecessary. This lemma enables us to show: *In the case of two fronts, one horizontal and one vertical, a scout can be sent onto* $(n,5)$, *for any positive integer n.*

Partly following [**5**], we call an army, holding a finite part of some quadrant plus a single square adjacent to the border of this quadrant, a *quasi-quadrant with outpost.* Now the Eriksson–Lindström lemma says: *For any positive n there exists a quasi-quadrant with outpost at distance n from the corner square of the quasi-quadrant, which can send a scout onto the square marked by an asterisk in* FIGURE 7. The same figure also illustrates the proof for the case $n = 7$.

Here the heavy guns $1,2,\ldots,6$ complete the laser guns $7,9,11,\ldots,17$ by an additional square each. Then the echelon of all laser guns $7,8,9,\ldots,18$ produces the staircase $s_7, s_8, s_9, \ldots, s_{18}$. The heavy gun 19 sends a scout onto s_{19}, which finishes the action by a twelvefold jump. This method works for $n \geq 3$; the cases $n = 1$ and $n = 2$ are very simple.

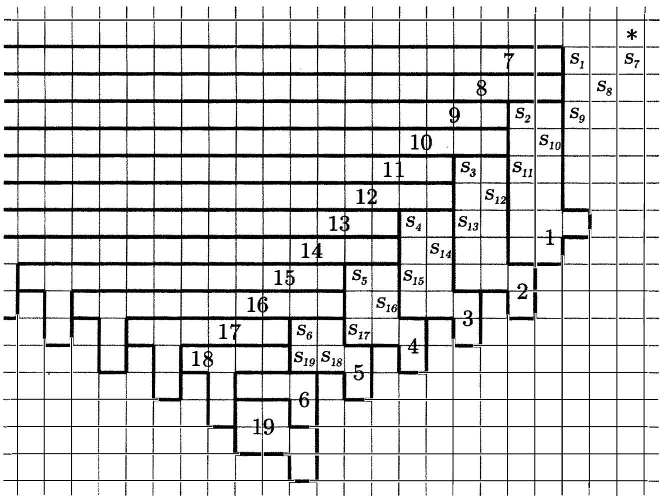


FIGURE 7
A special quasi-quadrant with outpost at distance 7 from the corner square.

Substituting "half-plane" for "quadrant" in the above definition, we get the notion of a *quasi-half-plane with outpost.* FIGURE 8 demonstrates the surprising fact that *a sufficiently large quasi-half-plane with outpost on an arbitrary square can send a scout onto any square at distance* 5 *from the border line.* Here 1 and 4 are quasi-quadrants with outposts, the Eriksson–Lindström lemma guarantees that they can send pegs to s_1, resp. s_4, and the protocol of the advance is $12J34J^2$. As a consequence, we obtain the promised result on the accessibility of $(n, 5)$ for arbitrary n.

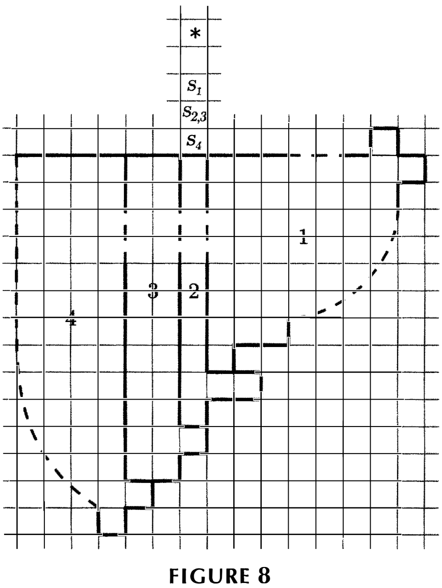


FIGURE 8
A quasi-half-plane with outpost sends a scout into an arbitrarily remote square of the northern fifth row.

Combining Figure 7 and 8 we see that the size of the quasi-half-plane with outpost we need for sending a scout onto the fifth row grows very fast with the distance between the outpost and the target square. An easy calculation shows that, for a distance of 67, we need an army of more than one million pegs!

A solitaire army with a solitary mounted man

In **[2]**, p. 717, one reads: "... we once showed that if any man of our army is allowed to carry a comrade on his shoulders at the start, then no matter how far away the extra man is, the problem [of sending a scout from the southern half-plane to a place in the fifth row of the northern half-plane] can now be solved." As we could not find any proof in the literature, we include one here.

Suppose there is a mounted man (i.e., two pegs stationed in the same square) in the southern half-plane. While both of them are there, this square cannot be jumped over, but the two pegs, one by one, can make legal jumps from there. Notice that a column of even length containing a mounted man one square apart from its end can be used as a laser gun: the extra man serves as the trigger. Figure 9a and 9b are slightly different; they show how to reach the fifth row when the distance of the mounted man from the border of the two half-planes is even or odd, respectively. The place of the mounted man is marked by a double square. The corresponding suitable advances are $12J3J45J6J^2$ and $123J45J6J^2$. Troops 4 and 6 are appropriate quasi-quadrants with outposts, sending one man each into the *first* northern square of the column, containing the mounted man.

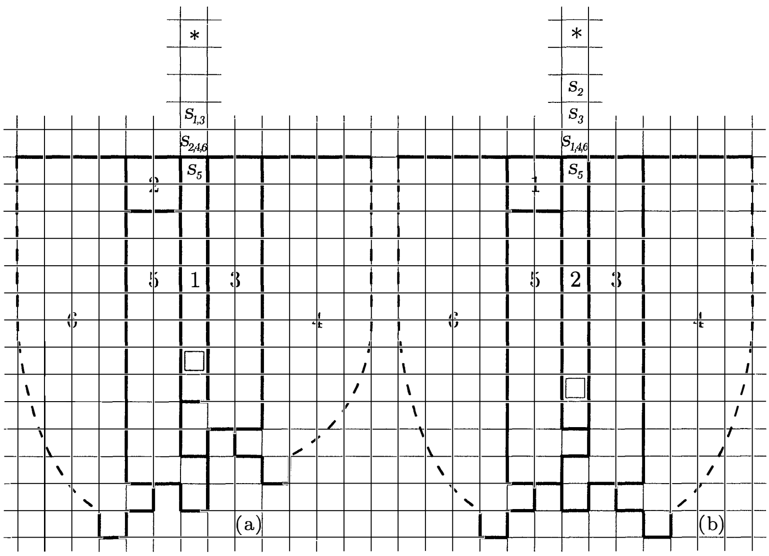


FIGURE 9
Scouting with a single mounted man.

Do not think, however, that the only possibility of reaching the fifth row is in the column of the mounted man. On the contrary: *For every pair C_1, C_2 of columns there exists a square $s_1 \in C_1$ in the southern half-plane such that a properly deployed troop in the southern half-plane with one mounted man in s_1 can send a scout onto the northern fifth square of C_2*. Instead of the lengthy full proof, we illustrate this fact in Figure 10 through a typical case.

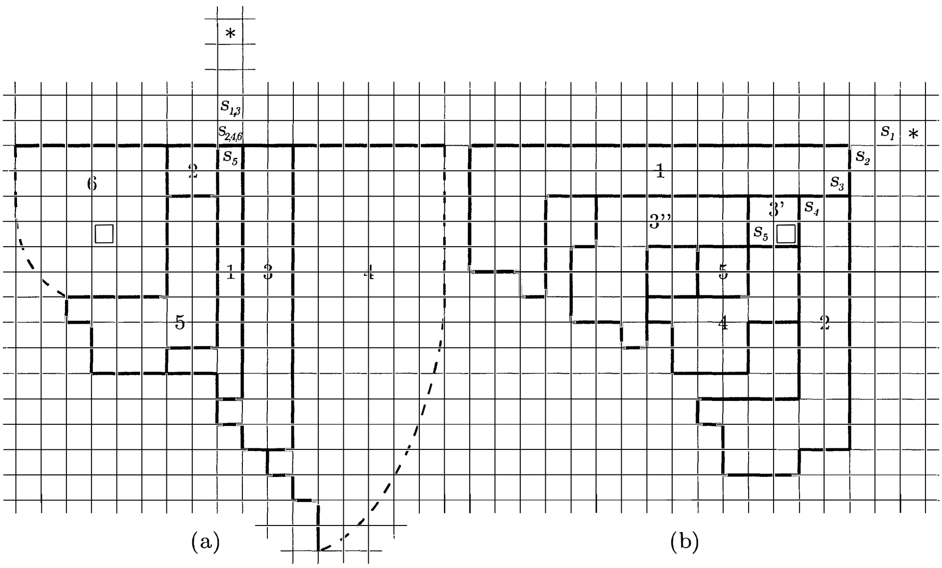


FIGURE 10
How a mounted man in a given column promotes scouting in another column.

As Figure 10a shows, it is enough to see that a suitable troop in the area 6 can send a peg onto the square s_2; then a scout reaches the target square by the advance $12J3J45J6J^2$. Figure 10b displays the suitable troop (involving several mortars) which sends a peg onto the desired square by $123'3''J45J^5$. One might worry that there is not enough space to deploy the troop in Figure 10b in the area marked by 6 in Figure 10a. However, we can guarantee the needed space in Figure 10a by replacing the guns 1, 3, and 5 by longer ones (of length 16, 18, and 12, respectively), and the quasi-quadrant with outpost (i.e., the troop 4) by a similar one whose outpost is at distance 19 from the corner.

The case of two skew fronts

Suppose that the solitaire forces are in the position of Figure 1d. Again, denote by (i, j) the square whose Manhattan distance from the left and right front lines equals i and j. Figure 11a shows that the square $(7, 7)$ marked by an asterisk is accessible: the two troops are exactly those of Figure 5b and 5c.

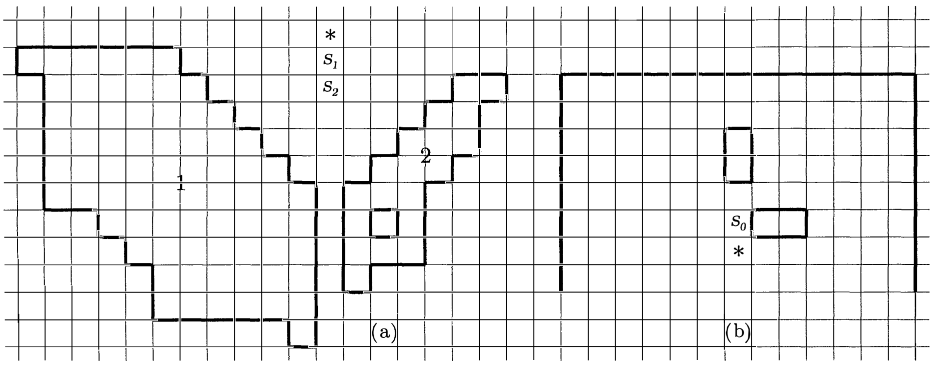


FIGURE 11
The maximal achievement in the case of two skew fronts. Patrols in encircled ground.

We prove that $(8,8)$ is inaccessible. In order to reach $(8,8)$, our army must send a patrol either to squares $(7,7)$ and $(6,6)$, or to squares $(9,7)$ and $(10,6)$. In the first case, let $s_0=(7,7)$ and $p(s_0)=1$; then the potential of the whole occupied area is $15\sigma^5+2(\sigma^7+\sigma^9+\dots)=15\sigma^5+2\sigma^6=1.464\dots$, while the potential of the patrol equals $1+\sigma=1.618\dots$ showing that the patrol cannot be sent to the desired place. The second case is even worse: for $s_0=(9,7)$, the potential of the patrol is the same, while that of the army will be $8\sigma^5+\sigma^6+9\sigma^7+\sigma^8=1.108\dots$. The squares $(n,7)$ are also inaccessible if $n>7$: for $(9,7)$ the preceding trick works, and the original Conway argument is applicable for $n>9$. In summary: *The squares that can be reached from two skew fronts are those at Manhattan distance no more than* 6 *from at least one front line, and the square* $(7,7)$.

Scouts in encircled ground

Finally, suppose that a quadrilateral area of size $n\times n$ with horizontal and vertical sides is fully encircled by a solitaire army, where n is odd. What is the maximal n such that a scout can be sent to the central square of this quadrilateral? Write $n_{\max}$ for this n. We already know that $n_{\max}\ge 9$. The Conway argument provides the upper limit $n_{\max}<15$. Suppose $n_{\max}=13$, and consider the last patrol, a member of which jumps into the centre. The members of this patrol are produced by two other patrols, i.e., by four pegs. They can occupy five essentially different positions; a sample is displayed on FIGURE 11b. Placing s_0 suitably, the summary potential of the army turns out to be less than that of the four pegs, contradicting the hypothesis. In our example the potential of all squares out of the 13×13 quadrilateral (i.e., of four half-planes minus four quadrants) equals

$$\sigma+\sigma^3+2\sigma^2-2\sigma^9-2\sigma^{11}=1.581\dots,$$

less than the potential of the two patrols: $1+\sigma=1.618\dots$. Hence $9\le n_{\max}\le 11$, so $n_{\max}$ is either 9 or 11. For quadrilaterals with skew sides, a similar question may be raised, and it can be treated in a similar manner. Let $n_{\max}$ be the number of squares constituting a diagonal of the maximal quadrilateral in this case. Then we obtain that $n_{\max}$ equals either 13 or 15. In both cases, the exact value of $n_{\max}$ remains unknown.

Acknowledgment. The helpful interest of our colleagues, especially of Sándor Csörgő, Jeremy Parrott, and Vilmos Totik, and the support of the Hungarian Ministry of Education under Grant No. FKFP 0877/1997 are gratefully acknowledged.

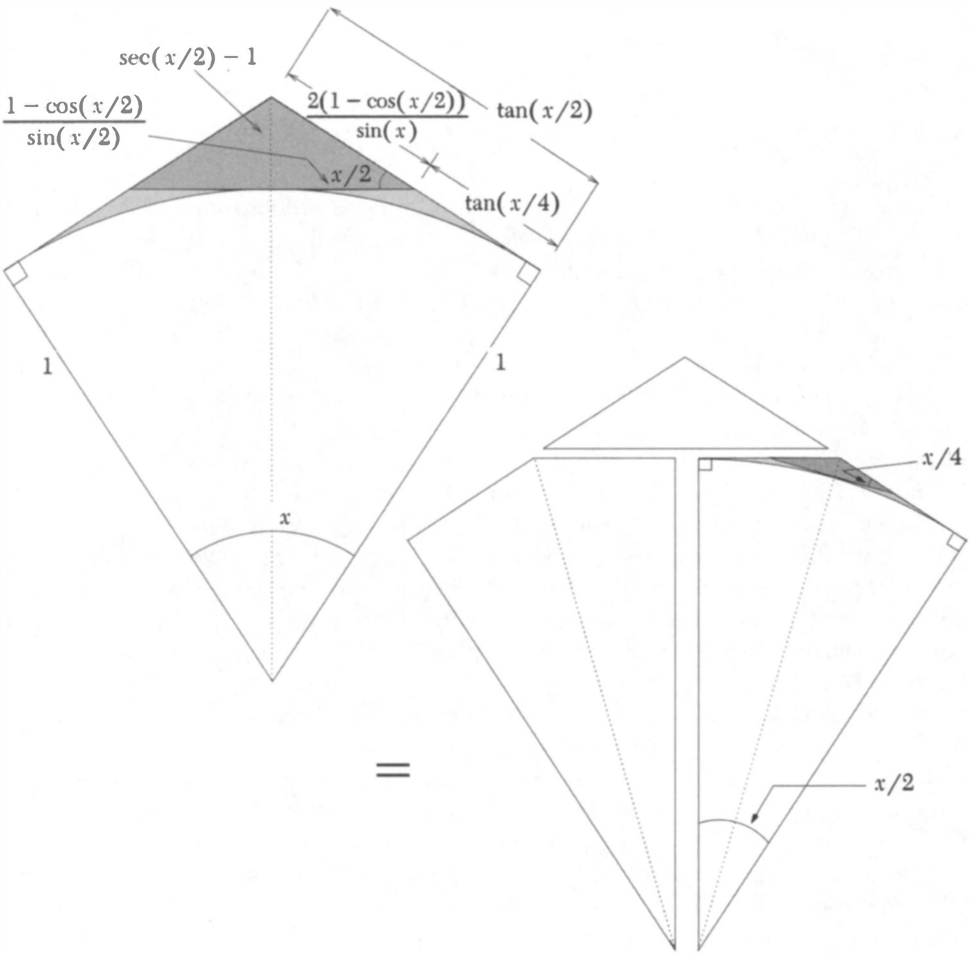
REFERENCES

1. M. Aigner, Moving into the desert with Fibonacci, this MAGAZINE 70 (1997), 11–21.
2. E. R. Berlekamp, J. H. Conway, and R. K. Guy, *Winning Ways*, Academic Press, London, UK, 1982.
3. F. Chung, R. Graham, J. Morrison, and A. Odlyzko, Pebbling a chessboard, *Amer. Math. Monthly* 102 (1995), 113–123.
4. H. Eriksson, Pebblings, *Electronic J. Combinatorics* 2 (1997), Research Paper No. 7, 18 pp.
5. H. Eriksson and B. Lindström, Twin jumping checkers in $\mathbb{Z}^d$, *European J. Combinatorics* 16 (1995), 153–157.
6. R. Honsberger, *Mathematical Gems, Vol.* 2, Math. Assoc. of Amer., Washington, DC, 1976. (Chapter 3).
7. A. Khodulev, Pebble spreading, *Kvant*, 1982, No. 7, pp. 28–31, 55. (In Russian)
8. M. Kontsevich, Problem M715, *Kvant*, 1981, No. 11, p. 21. (In Russian)
9. Sun Tzu, *The Art of War*, Oxford Univ. Press, Oxford, UK, 1971.

Proof Without Words: Slicing Kites into Circular Sectors

$$\text{Areas:} \quad \sum_{n=1}^{\infty} \frac{2^n[1-\cos(x/2^n)]^2}{\sin(x/2^{n-1})} = \tan\left(\frac{x}{2}\right) - \frac{x}{2}, \qquad |x| < \pi$$

$$\text{Side-lengths:} \quad 2\sum_{n=1}^{\infty} \frac{1-\cos(x/2^n)}{\sin(x/2^{n-1})} = \tan\left(\frac{x}{2}\right), \qquad |x| < \pi$$



—Marc Chamberland
Grinnell College
Grinnell, IA 50112

NOTES

Invisible Points in a House of Mirrors

G.W. TOKARSKY
University of Alberta
Edmonton, Alberta
Canada T6G 2G1

F.G. BOESE
Max-Planck-Institut für Extraterrestrische Physik
Postfach 1312
D-85741 Garching
Germany

Introduction

On entering a house of mirrors at a circus, you find yourself surrounded by reflected images of yourself. If the house is well lit and has walls completely covered with mirrors, some intriguing questions arise: Is there a location from which you cannot see any reflections of yourself? Or, better yet, is there a location at which you disappear from everyone's view?

We will represent people and lights mathematically by points; the house of mirrors will be represented by a two-dimensional polygonal room whose sides all act as mirrors. Light rays travel along straight lines and reflect off the walls so that the angle of incidence equals the angle of reflection. Any light ray that strikes a vertex (i.e., a boundary point without a tangent line) is considered to end there.

We will call a point *invisible* if it has the property that any light ray that passes through it in any direction never returns to that point. All other points are said to be *visible*. Physically, an invisible point acts like the opposite of a black hole. If an invisible point exists in our house of mirrors, then the answer to our first question is yes, since for you to see a reflected image of yourself, light must travel *from* yourself back *to* yourself.

Invisible points can be found at vertices of triangles and squares. Then the triangles or squares can be "tiled" together to form polygonal rooms with *interior* invisible points (see **[1]** and **[2]**). Ian Stewart in **[3]** and **[4]** gives an entertaining account of these results. For further illumination methods and problems, see **[5]** and **[6]**.

In this article we will prove the existence of invisible vertices for a general polygonal region, and then use the tiling technique in **[2]** to construct more examples of polygonal rooms with interior invisible points. We will also partially classify invisible points and pose some new questions.

We start with the main result.

THEOREM 1. *Let a polygonal room P have vertices at* $A, A_1, A_2, \ldots, A_m$ *taken counterclockwise, with respective interior angle sizes* x, $x_1 = n_1 x$, $x_2 = n_2 x, \ldots, x_m = n_m x$, *measured in degrees, where* $n_1, n_2, \ldots, n_m$ *are positive integers. If* x *divides* 90, *then the vertex* A *is invisible.*

Proof. Let P be a polygonal room as in the statement of the theorem. We will measure all angles mod $2x$. Since $2x$ divides 180 and all the interior angles of P sum to a multiple of 180, there is an even number of angles that are congruent to x mod $2x$; all the remaining angles are congruent to 0 mod $2x$.

Now assume that a light ray leaves vertex A at an angle θ, where $0 < \theta < x$ is measured counterclockwise from AA_1 as shown in FIGURE 1. Let side AA_1 be assigned the symbol $\pm\theta$. Inductively assign either the symbol $\pm\theta$ or the symbol $x \pm \theta$ to each side successively, according to the following rule: Let side $A_{i-1}A_i$ be assigned either $\pm\theta$ or $x \pm \theta$. Then side A_iA_{i+1} receives the same symbol if $x_i \equiv 0 \mod 2x$; otherwise, A_iA_{i+1} receives the other symbol. (See FIGURE 2.) Notice that, since there is an even number of interior angles congruent to x mod $2x$, the last side A_mA must be labeled $x \pm \theta$.

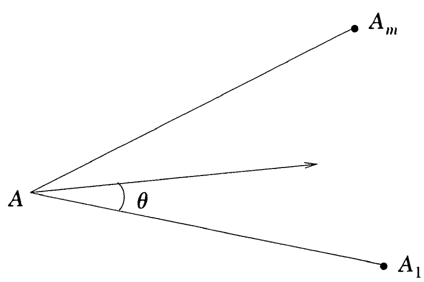


FIGURE 1
An exiting light ray.

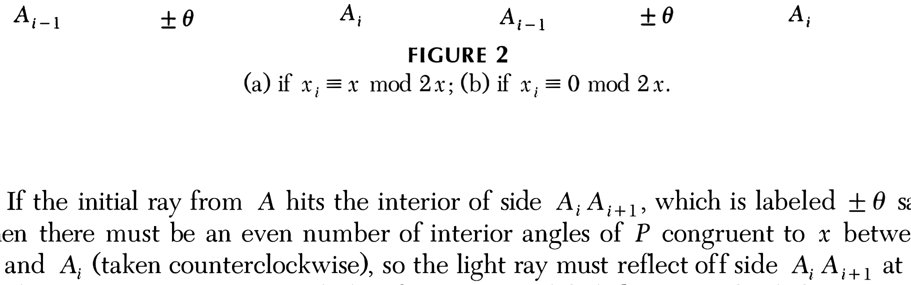


FIGURE 2
(a) if $x_i \equiv x \mod 2x$; (b) if $x_i \equiv 0 \mod 2x$.

If the initial ray from A hits the interior of side A_iA_{i+1}, which is labeled $\pm\theta$ say, then there must be an even number of interior angles of P congruent to x between A and A_i (taken counterclockwise), so the light ray must reflect off side A_iA_{i+1} at an angle congruent to $\pm\theta$. Similarly, if A_iA_{i+1} is labeled $x \pm \theta$, the light ray must reflect at an angle congruent to one of these angles, since then there is an odd number of interior angles of P congruent to x in between.

Now, by induction on the number of segments in the path of the light ray, we show that the light ray reflects off each side according to its label. This is true for $s = 1$ by the preceding argument. Suppose it is true for $s = k$ and take the $(k+1)^{\text{st}}$ segment of the path, which might appear as in FIGURE 3. There must be then an odd number of interior angles of P congruent to x between A_j and A_l and hence the light ray must reflect off side A_lA_{l+1} at an angle congruent to $x \pm \theta$. The other cases are similar, and the induction is complete.

We conclude that, if the light ray returns to A, then it does so at one of the angles $\pm\theta$ measured from AA_1. Since $0 < \theta < x$, the ray must return to A at the same angle θ that it left. The light ray can do this if it hits some side at 90°. Thus either $\pm\theta \equiv 90 \mod 2x$ or $x \pm \theta \equiv 90 \mod 2x$. In either case, this forces x to divide θ, which is impossible. ■

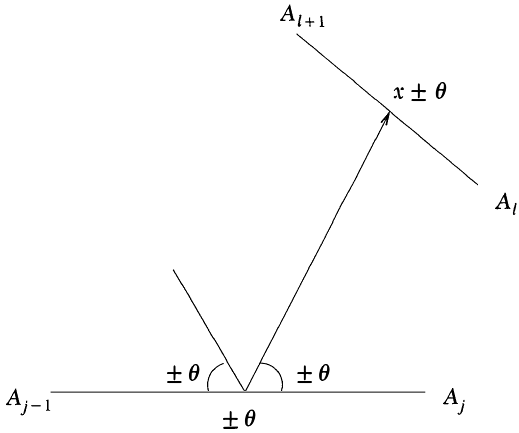


FIGURE 3
A portion of a light ray path.

If the conditions of the theorem are satisfied, then the only allowable integer values for x are the divisors of 90. For example, each of the following polygons has an invisible vertex at A:

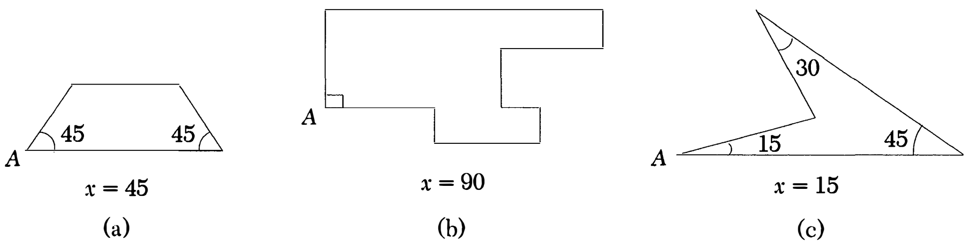


FIGURE 4
Invisible vertices.

It is also worth noting that, in a polygonal room whose sides are all either horizontal or vertical, every vertex of interior angle size 90° is an invisible point.

Building larger rooms If P is a polygonal room as in Theorem 1, with vertex A a specified invisible point, then we can build a larger polygonal room Q with interior invisible points by using copies of P as tiles, such that the interiors of common edges become interior to Q. (See, e.g., FIGURE 5.) In addition, any two tiles sharing a common edge must be mirror images of each other in that edge. The tiles must not overlap and all vertices labeled differently from A in P must remain vertices in Q.

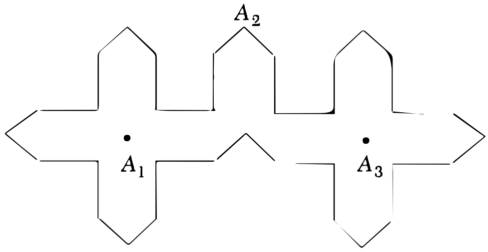


FIGURE 5
Room with invisible points.

Because the tiles are mirror images of each other, a light ray path between any two points labeled A in Q would fold up to a corresponding light ray path from A to itself in a single copy of P. The fact that the vertices in P different from A remain vertices in Q guarantees that the first light ray path must avoid these vertices and hence that the (folded up) light ray path in P avoids all vertices of P different from A. But this then produces a light ray path from A to A in P, which is impossible. We have proved the following corollary to our first theorem:

COROLLARY. *If we build a bigger polygonal room Q using copies of P as tiles as described above, and such that all vertices of P different from A remain vertices in Q, then a light ray cannot travel between any two points labeled A in Q. In particular, all points labeled A in Q are invisible points in Q.*

For example, using the polygon in FIGURE 4(a) as a tile, we can construct the polygonal room in FIGURE 5, with three invisible points labeled A_1, A_2, and A_3.

Hiding places We can now answer our second question, about points invisible from anywhere in the room. The answer obviously depends on the placement of the lights in the polygonal room. Referring to FIGURE 5 again, if lights are placed at A_1 and A_2, then, by the corollary, light rays from these points never reach A_3. Thus, if you stand at A_3 you and your reflections disappear completely from view, as if you had stepped into a closet.

More examples appear in FIGURE 6, in which an isosceles right triangle is used as a tile. If lights are placed at A_1, A_2, A_3, and A_4, then at A_5, you would disappear from view.

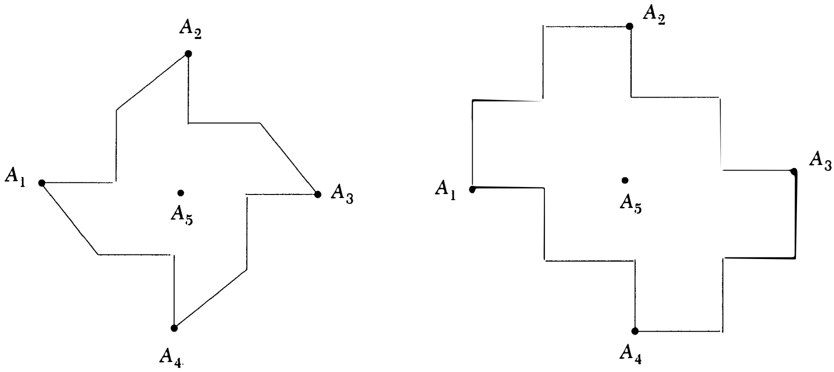


FIGURE 6
Rooms with invisible points.

Triangular rooms We do not know whether the condition of Theorem 1 is a necessary condition for a vertex to be invisible, even for triangles. Observe, too, that the theorem fails if 90 is replaced by 180. For example, if A is any vertex of an equilateral triangle, then A is a visible point since a light ray along an altitude from A reflects back to A. For certain special cases, however, the condition of Theorem 1 is necessary and sufficient. In what follows, $m(\angle Y)$ denotes the degree measure of the angle at the vertex Y.

THEOREM 2. *If P is a right triangle ABC, with right angle at B and $m(\angle A) = x$, then A is invisible if and only if x divides* 90.

Proof. If $90 = nx$, then $m(\angle A) = x$, $m(\angle B) = 90 = nx$, and $m(\angle C) = (n-1)x$ and A is an invisible point by Theorem 1. If x does not divide 90 $\left(\text{say } \frac{90}{k+1} < x < \frac{90}{k} \text{ with } k \geq 1\right)$ then A is visible using a light ray leaving A at an angle $\theta = 90 - kx$, as shown in FIGURE 7.

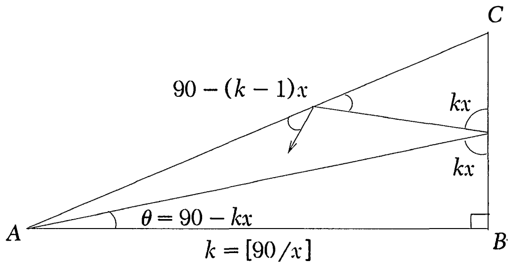


FIGURE 7
A light ray path.

As the reader can easily check, such a light ray path is possible and alternates between AC and AB with the reflecting angle successively increasing in size by x. After exactly $k+1$ reflections, the ray hits at 90° and reflects back to A. ■

THEOREM 3. *If $\triangle ABC$ is isosceles with $BA = BC$ and $m(\angle A) = m(\angle C) = x$, then A is invisible if and only if x divides* 90.

Proof. If x divides 90 (say $90 = nx$) and $m(\angle A) = m(\angle C) = x$, then $m(\angle B) = 180 - 2x = (2n-2)x$ with $n \geq 2$ and A is an invisible point by Theorem 1.

If x does not divide 90, then as before let $k = \left\lfloor \frac{90}{x} \right\rfloor$ and let a light ray leave A at an angle $\theta = 90 - kx$ from AC. After k reflections, the ray will hit either AC or BC at 90° and return to A. ■

The following result holds for arbitrary triangles:

THEOREM 4. *If O is an invisible point of $\triangle ABC$, then O must be a vertex. In addition, $\triangle ABC$ must be either right-angled or obtuse, and O must be the vertex of smallest angle.*

Proof. Let O be an invisible point of triangle ABC. If O is not on the longest side, say AB, then we can drop a perpendicular from O to the interior of side AB. If O is in the interior of side AB, then we can drop a perpendicular from O to the interior of the nearest side. In either case, O is visible.

This means that O must be one of A or B. If all vertex angles of $\triangle ABC$ are acute, then we can drop a perpendicular from O to the interior of the opposite side. This means that angle C, the largest angle, must be either a right angle or obtuse.

Let $m(\angle A) = x < y = m(\angle B)$ with $m(\angle C) \geq 90$. Then any light ray leaving B at an angle (measured from BA) of $\theta = 90 - kx$, with k a positive integer and $0 < \theta < y$, will after k reflections hit either AB or AC at 90°, and reflect back to B. Such a light ray exists because $x + y \leq 90$. Thus $O = A$. ■

Combining Theorems 2 and 4, we get a complete classification of the invisible points of a right triangle, namely as the acute vertices of size x, where x divides 90. Theorems 3 and 4 also classify the invisible points of an isosceles triangle. Whether a similar result holds for general triangles is an open question, and we make the following conjecture:

CONJECTURE. *A point O in a triangle is invisible if and only if it is a vertex of size x where x divides 90 and some other vertex has size px, with p a positive integer.*

Certainly if this condition holds, then the third vertex has size $(2n - p - 1)x$ where $180 = 2nx$, and P is invisible by Theorem 1. Thus the conjecture reduces to showing that any invisible point (which must be a vertex by Theorem 4) satisfies the given condition.

REFERENCES

1. G. W. Tokarsky, Problem 94-3, *SIAM Review* 37 (1995), 107–109.
2. G. W. Tokarsky, Polygonal rooms not illuminable from every point, *Amer. Math. Monthly* 102 (1995), 867–879.
3. Ian Stewart, Mathematical recreations, *Scientific American*, August 1996, 100–103.
4. Ian Stewart, Feedback, *Scientific American*, February 1997, 98.
5. V. Klee and S. Wagon, *Old and New Unsolved Problems in Plane Geometry and Number Theory*, Mathematical Association of America, Washington, DC, 1991.
6. H. T. Croft, K. J. Falconer, and R. K. Guy, *Unsolved Problems in Geometry*, Springer-Verlag, New York, NY, 1991.

Bijections for the Schröder Numbers

LOUIS W. SHAPIRO
Howard University
Washington, DC 20059

ROBERT A. SULANKE
Boise State University
Boise, ID 83725

1. Introduction

The *small Schröder numbers*,

$$(s_n)_{n\geq 0} = (1, 1, 3, 11, 45, 197, \ldots),$$

count many combinatorial configurations including the ways to place properly parentheses in a string of letters. The *large Schröder numbers*,

$$(r_n)_{n\geq 0} = (1, 2, 6, 22, 90, 394, \ldots),$$

also count many configurations, in particular, sets of upper diagonal lattice paths using the steps (1, 0), (1, 1), and (0, 1). Using "pictorial proofs," we will relate the sequences, $(s_n)_{n\geq 0}$ and $(r_n)_{n\geq 0}$, by a sequence of bijections (one-to-one onto functions) between various sets of configurations. Consequently, we will see why the terms of $(r_n)_{n \geq 1}$ are twice those of $(s_n)_{n \geq 1}$.

In his 1870 paper, "Vier Kombinatorische Probleme" **[9]** (see also **[11]**), Ernst Schröder considered counting the ways to place parentheses on a string of letters. We recursively define a *bracketing* on a string of letters so that each letter is itself a bracketing and so that any consecutive sequence of two or more bracketings becomes a bracketing when enclosed by a pair of parentheses. However, we omit the parentheses enclosing any single letter and the outer parentheses enclosing the bracketings covering all the letters of the string. For example, the strings a and ab have only a and ab as the respective bracketings, while the string $abcd$ has eleven bracketings:

$$abcd, (ab)cd, (abc)d, a(bc)d, a(bcd), ab(cd),$$
$$((ab)c)d, (a(bc))d, a((bc)d), a(b(cd)), (ab)(cd).$$

We define the *small Schröder numbers*, $(s_n)_{n \geq 0}$, so that s_n is the number of bracketings for a string of $n+1$ letters. Schröder's most concise formulation for these numbers was given by the generating function

$$\sum_{n \geq 0} s_n x^n = \left(1 + x - \sqrt{1 - 6x + x^2}\,\right)/4x.$$

For perspective we note that in 1838, Eugène Catalan (see **[11**, p. 212**]**) considered counting the *binary* placement of parentheses on a string of letters. For example, the string $abcd$ has five *binary bracketings*, namely, $((ab)c)d$, $(a(bc))d$, $a((bc)d)$, $a(b(cd))$, and $(ab)(cd)$. What we now call the Catalan numbers, namely

$$(c_n)_{n \geq 0} = \left(\frac{1}{n+1}\binom{2n}{n}\right)_{n \geq 0} = (1, 1, 2, 5, 14, 42, \ldots),$$

count the binary bracketings for each string of $n+1$ letters. One can reduce any of the sets of configurations we encounter to one enumerated by the Catalan numbers.

From the bijections implicitly defined in FIGURE 1, we see that bracketings are immediately related to two other configurations, namely (1) planted trees with internal nodes without degrees at least two; and (2) dissections of convex polygons. (A *plane tree* is a rooted unlabeled tree, where two trees are the same if one can be continuously transformed into the other in the plane while the non-root nodes remain above the root. A plane tree is called a *planted tree* if its root has degree one. Given a convex polygon with a designated base, a *dissection of the polygon* is a set of non-crossing line segments joining some of the non–adjacent vertices.) In FIGURE 1, we see that for each dissection of an $(n+2)$-gon there is a naturally determined planted tree with $n+1$ leaves, which in turn determines a bracketing on $n+1$ letters. Since the existence of a bijection between two sets implies they have equal cardinality, we have the following.

THEOREM 1. *For $n \geq 0$, the number of bracketings on a string of $n+1$ letters is equal to the number of dissections of a convex polygon with $n+2$ sides, one of which is designated as the base.*

For $n \geq 0$, let R_n denote the set of lattice paths in the xy-plane that run from $(0, 0)$ to (n, n), that never pass below the line $y = x$, and that use horizontal steps, $(1, 0)$, diagonal steps, $(1, 1)$, and vertical steps, $(0, 1)$. The bottom row of FIGURE 4 shows the six paths of R_2. We call the paths of R_n, the *Schröder paths of length n*, and we define the *large Schröder number*, r_n, to be the cardinality of R_n. One of the concise

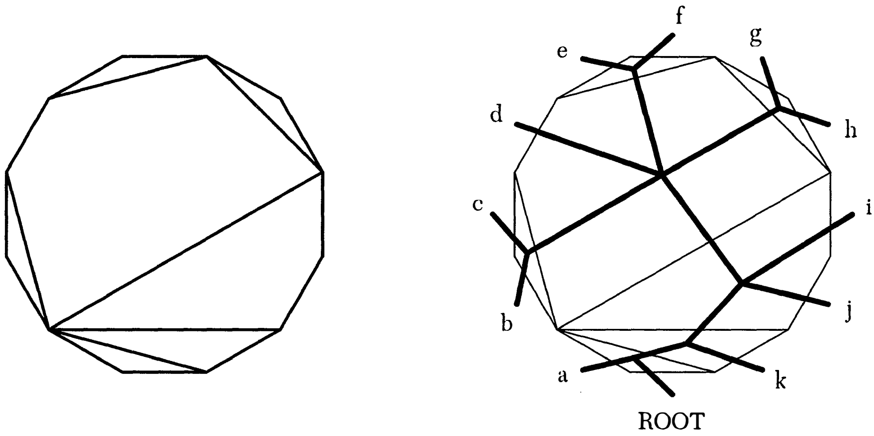


FIGURE 1
The dissected polygon and the planted tree corresponding to the bracketing $a(((((bc)d(ef)(gh))ij)k)$.

formulas for these numbers is in terms of the Catalan numbers (see **[1]**):

$$r_n = \sum_{k=0}^{n} \binom{2n-k}{k} c_{n-k}, \quad \text{for } n \geq 0.$$

Using generating function methods (as in **[7]**) one can find the generating function for $(r_n)_{n \geq 0}$ and then show that $r_n = 2s_n$ for $n \geq 1$. Here however, we will continue the pictorial bijective approach to explain this "doubling."

THEOREM 2. *For $n \geq 1$, the number of Schröder paths of length n is twice the number of dissections of a convex polygon with $n+2$ sides, one of which is designated as the base.*

2. Notes for the picture proof

Our proof of Theorem 2 is indicated in FIGURES 2 and 3, and again in FIGURE 4. By previewing these figures one should find many of our words redundant. Using two colors, G and B, we double the set of all dissections of an $(n+2)$-gon by coloring each base either G or B. We will prove Theorem 2 by establishing a bijection from the dissections with a bicolored base to the path set R_n as a composition of five bijections.

Bijection 1 (FIGURE 2): Here we consider *planted full binary trees with bicolored right (non-leaf) branches*, that is, planted trees having internal nodes with a left and a right out edge, such that each non-leaf-adjacent right out edge is colored by G or B. FIGURE 2 illustrates our algorithm for growing such a tree in a dissected polygon. Here the base and the root are colored G. As we grow such a tree, we add the dashed diagonal segments so that they completely triangulate the polygon and are always crossed by the right branches of the tree. A right branch that crosses a dashed segment is drawn as a thick B edge, while a right branch crossing an existing interior segment is drawn as a thin G edge. Technically, our tree growing algorithm creates just two new edges at each of n steps; our FIGURE 2 has combined some steps.

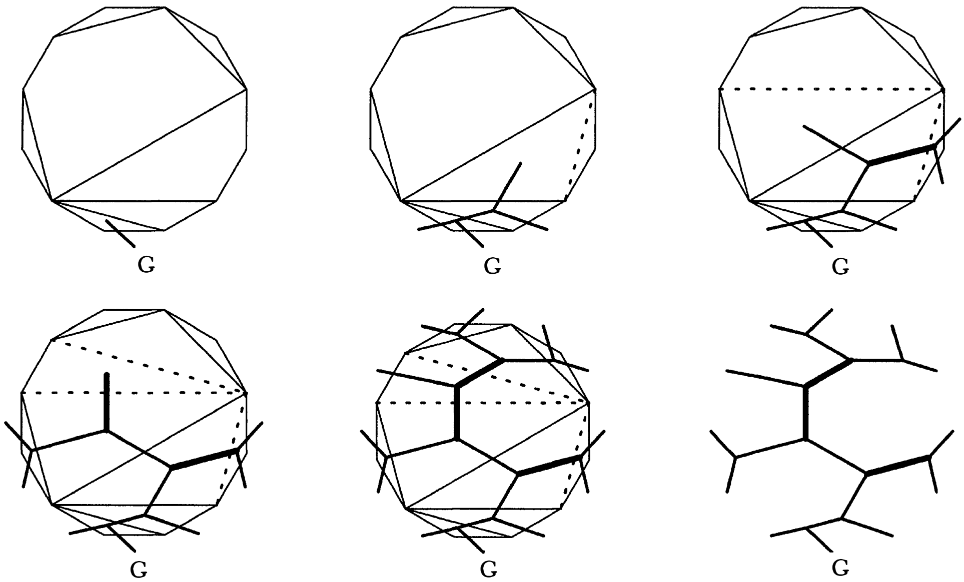


FIGURE 2
Constructing the planted full binary tree with bicolored right (non-leaf) branches that is associated with the dissected polygon of FIGURE 1a.

Bijection 2 (FIGURE 3ab): FIGURE 3a shows the tree from FIGURE 2. In FIGURE 3b we consider *planted trimmed binary trees with bicolored right branches*, that is, a planted tree having internal nodes with either a left out edge, a right out edge, or both, where each right out edge is either G or B. Such a tree can be obtained from a full binary tree with bicolored right (non-leaf) branches by removing the leaves and their adjacent edges. This process yields a one-to-one correspondence since edges can be added to any trimmed binary tree to obtain a full binary tree in a unique well-defined manner.

Bijection 3 (FIGURE 3bcd): In each such trimmed tree we will label the branches above each node as follows: If the node has only a left branch, place an N to the left of the branch. If the node has only a right branch, place an E to the left of branch. If the node has two branches, place an A to the left of the left branch and a T to the left of the right branch. We also place an A to the left and a T to the right of the root edge. From each trimmed tree we can obtain a word on the letters N, E, A, and T as follows: We traverse each tree in a depth-first search manner from the left as in FIGURE 3c. We record each label when the corresponding edge is first encountered. Tracing around the tree in FIGURE 3b yields the sequence AENAATEATTT.

A *parallelogram polyomino* is an array of unit squares that is bounded by two lattice paths using the steps, $(0,1)$ and $(1,0)$, and intersecting only initially and finally. Two steps, one from each of the boundary paths, are called "diagonally opposing" if they have the same position in the consecutive ordering of the steps in their respective paths (i.e., the end points of the two steps lie on the same line of slope -1).

From each word on N, E, A, and T, we then form two paths enclosing a parallelogram polyomino by consecutively mapping each letter to a pair of diagonally

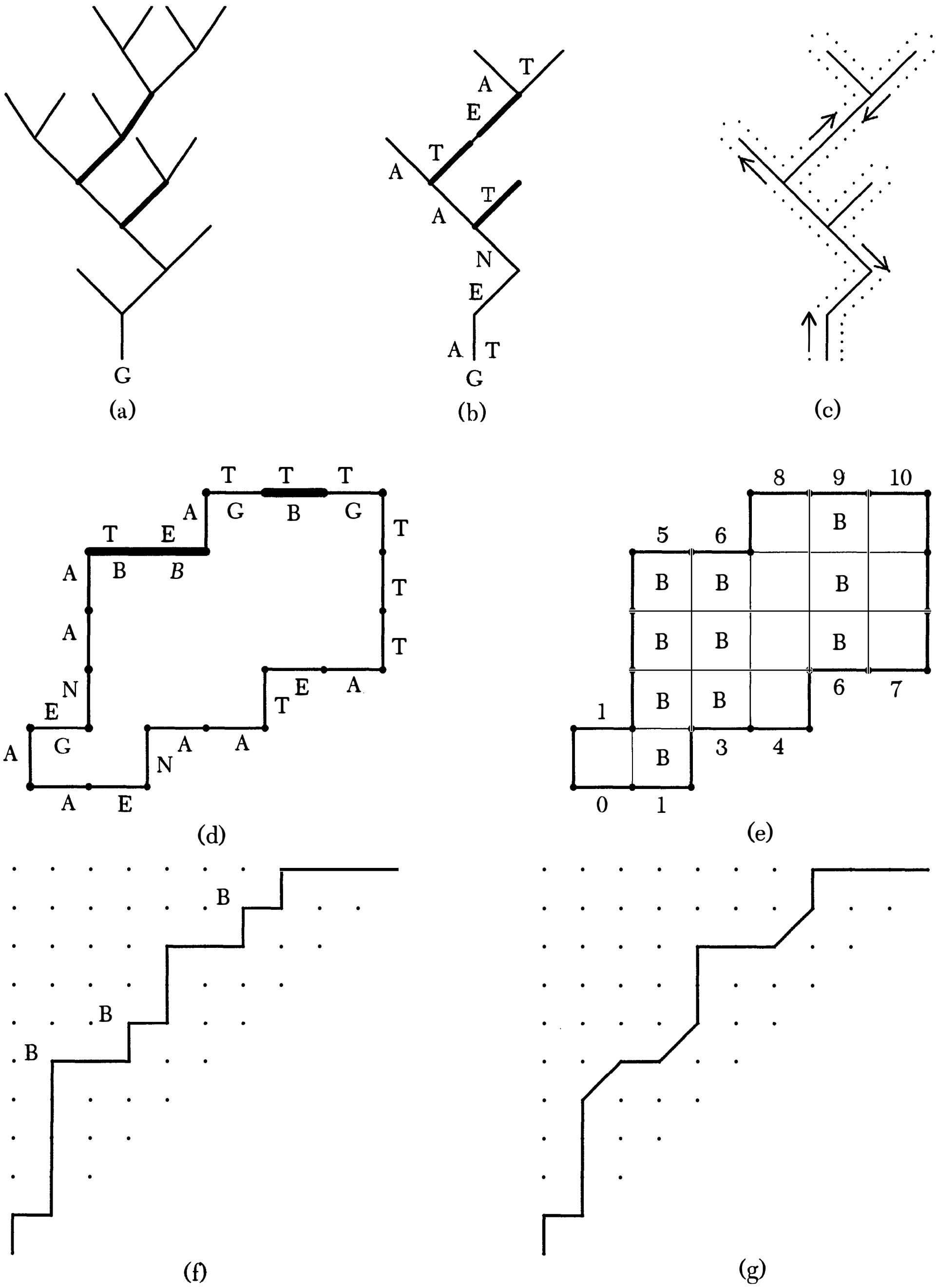


FIGURE 3

Constructing the lattice path associated with the planted full binary tree with bicolored right (non-leaf) branches of FIGURE 2. Thick edges have color B.

opposing steps according to the scheme of Table 1 where each pair of arrows indicates a diagonal opposition of steps. This word AENAATEATTT determines the polyomino in FIGURE 3d. Moreover, as we traverse the tree of FIGURE 3b we record the color of each right edge traversed when each E or T is recorded, assigning this color to the corresponding horizontal step on the upper-left boundary of the polyomino. We assign the root color to the last horizontal step on the upper-left boundary.

TABLE 1

A (“apart”) yields	↑		E (“east”) yields	→	
		→			→
N (“north”) yields	↑		T (“together”) yields	→	
		↑			↑

Bijection 4 (Figure 3def): We first define a *Catalan path* from $(0, 0)$ to (n, n) as a lattice path that uses the steps $(1, 0)$ and $(0, 1)$, runs from origin to (n, n), and never passes below the line $y = x$.

For Figure 3e, assign the indices $0, 1, 2, \ldots, n$ to the steps of both boundary paths of the polyomino in Figure 3d. In the figure we have recorded only the indices for the horizontal steps. Assign the color of each top horizontal step to the column beneath. Then map each polyomino to a Catalan path so each peak (right-hand turn) of the path has coordinates (x, y), where x is the lower index and y is the upper index of a column of the polyomino. In Figure 3f, the coordinates of the peaks are $(0, 1)$, $(1, 5)$, $(3, 6)$, $(4, 8)$, $(6, 9)$, and $(7, 10)$. Color each peak the same as the corresponding column.

One referee suggested an alternative description for Bijection 4: In Figure 3d cut the upper path at the *right* endpoints of its horizontal steps, and then straighten the resulting segments to get vertical segments of steps. The top step of each segment is colored B or G. Also cut the lower path at the *left* endpoints of its horizontal steps, and then straighten the resulting segments to get horizontal segments of steps. Concatenate these segments by starting with the first vertical segment and alternately appending the segments of each type as originally ordered. The nonintersecting property implies that the sequence of vertical segments lengths strictly majorizes that for the horizontal ones and so the result is a Catalan path that only returns to the diagonal at its endpoint. Color the peaks according to the color of the topmost step of each vertical segment and delete the first and last steps to obtain a Catalan path as in Figure 3e.

Bijection 5 (Figure 3fg): Finally, we obtain a Schröder path by converting the B peaks on the Catalan path to diagonal steps.

3. Background remarks

The Schröder numbers are ancient and arguably misnamed. A few years ago, as noted by Stanley **[10]** in his interesting survey of the Schröder numbers, David Hough discovered that the small Schröder numbers were apparently known to Hipparchus over 2100 years ago. Specifically, Plutarch (see **[10]**) records that Hipparchus knew the tenth number, $s_9 = 103049$, in the context of counting certain logical propositional forms. See also **[11**, p. 213]. Problem 6.39 of Stanley's recent book **[11]** lists 19 different configurations enumerated by the Schröder numbers and gives additional references for these numbers.

Motivating this note was a preprint version of Problem 6.17c of **[11]**, which asked for a bijective proof relating dissections to Schröder paths. Moser and Zayachkowski **[5]** studied such paths in 1961. We note that Rogers and Shapiro **[8]** and later Bonin, Shapiro, and Simion **[1]** have also defined maps that bijectively relate bracketings to Schröder lattice paths.

Counting dissections of polygons dates from about 1758, when Euler and Segner considered the enumeration of complete dissections of a polygon, or *triangulations*, which resulted in the European introduction of the Catalan numbers. These numbers were known to the Mongolian mathematician, Ming An-tu, who obtained recurrences for them in the 1730's. (See [**11**, p. 212].) A bijection between the set of triangulations of a polygon and the set of planted binary trees was perhaps first given by Forder [**4**]. Restrictions of our Bijections 1 to 4 will map the two triangulations of an $(n + 2)$-gon to the set of Catalan paths running to (n, n). In FIGURE 4, notice that the triangulations of

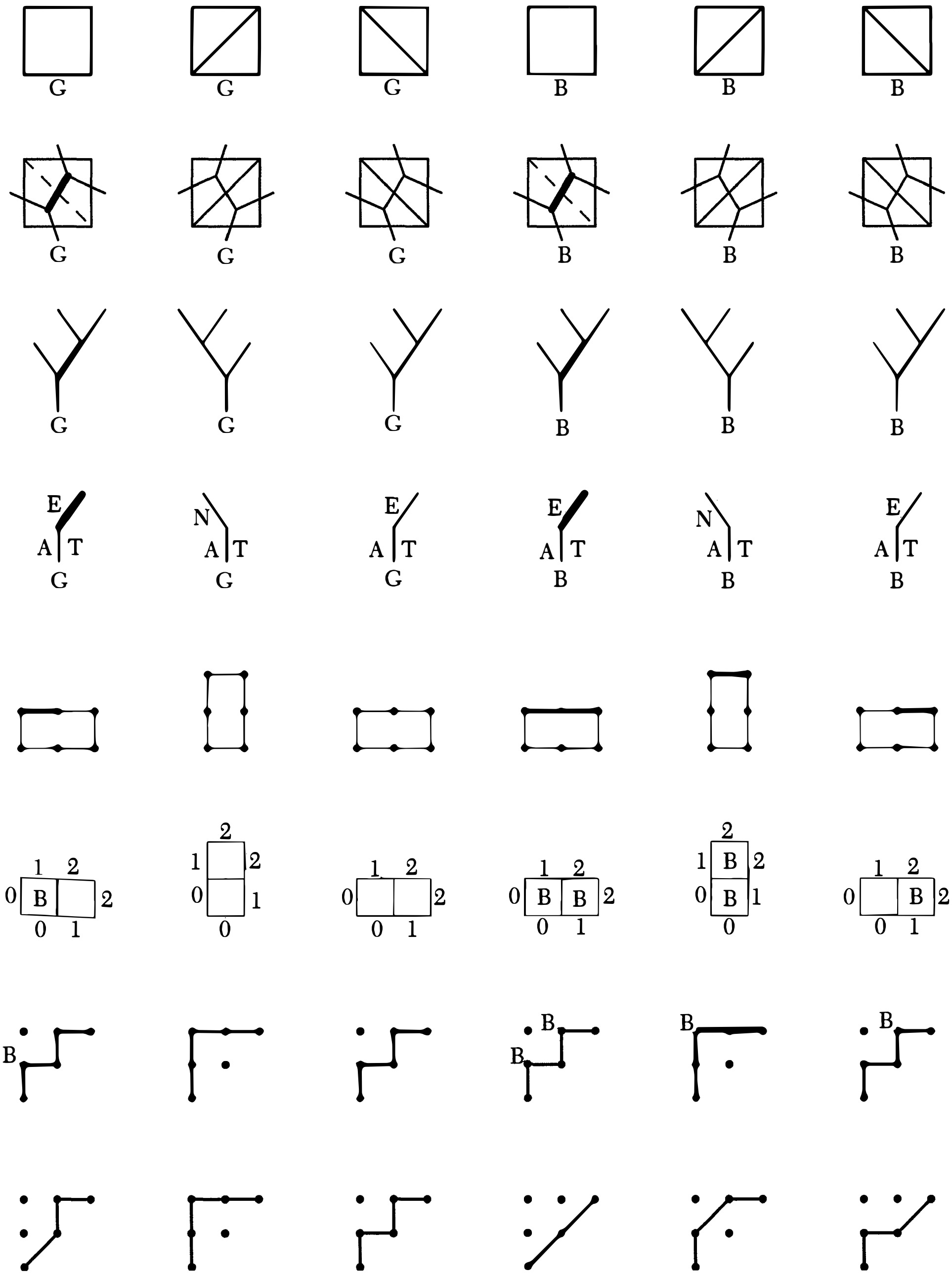


FIGURE 4
The sequence of bijections from the dissections of a square to the Schröder paths in R_2.

the square with G base are mapped to Catalan paths. Our use of trimmed binary trees with bicolored right branches agrees with the recent use by Foata and Zeilberger **[3]** in their bijective proof showing the small Schröder numbers to satisfy the recurrence $(n+1)s_n = 3(2n-1)s_{n-1} - (n-2)s_{n-2}$, for $n \geq 2$.

The four correspondences defined in Table 1 can be gleaned from Delest and Viennot **[2]**. Recently, Barcucci, Del Lungo, Fezzi, and Pinzani (see **[7]**) considered parallelogram polyominoes consisting of black and white columns and were the first to find that r_n counts all such polyominoes with semiperimeter $n+1$. In the mid 1950's, Narayana **[6]** gave a bijection between non-colored parallelogram polyominoes and Catalan paths which is equivalent to ours of FIGURE 3ef.

Acknowledgment. We appreciate the referees' suggestions for improving this note.

REFERENCES

1. J. Bonin, L. Shapiro, R. Simion, Some q-analogues of the Schröder numbers arising from combinatorial statistics on lattice paths, *J. Stat. Plann. and Infer.* 34 (1993), 35–55.
2. M. P. Delest and G. X. Viennot, Algebraic languages and polyominoes enumeration, *Theor. Computer Science* 34 (1984), 169–206.
3. D. Foata and D. Zeilberger, A classic proof of a recurrence for a very classical sequence, *J. Comb. Theor., Ser. A.*, 80 (1997), 380–384.
4. H. G. Forder, Some problems in combinatorics, *Math. Gazette* 45 (1961).
5. L. Moser and W. Zayachkowski, Lattice paths with diagonal steps, *Scripta Math.* 26 (1961), 223–229.
6. T. V. Narayana, *Lattice Path Combinatorics with Statistical Applications*, Univ. of Toronto Press, Toronto, Canada, 1979.
7. E. Pergola and R.A. Sulanke, Schröder triangles, paths, and parallelogram polyominoes, *Journal of Integer Sequences*, Vol. 1 (1998), Art. 98.1.7.
8. D. G. Rogers and L. W. Shapiro, Some correspondence involving the Schröder numbers and relations, *Comb. Math.*, Proc. of the Intern. Conf., Canberra 1977, *Lecture Notes in Math.* 686, Springer-Verlag, Berlin, Germany, (1978), 267–276.
9. E. Schröder, Vier Kombinatorische Probleme, *Z. Math. Phys.* 15 (1870), 361–376.
10. R. P. Stanley, Hipparchus, Plutarch, Schröder, and Hough, *Amer. Math. Monthly* 104 (1997), 344–350.
11. R. P. Stanley, *Enumerative Combinatorics*, Vol. II, Cambridge University Press, Cambridge, UK, 1999.

Spanning Trees: Let Me Count the Ways

DOUGLAS R. SHIER
Clemson University
Clemson, SC 29634-0975

1. Introduction An important pedagogical lesson is that there need not be a single, or even preferred, solution to a given problem. I learned this lesson well from students at the London School of Economics who were studying graph theory as part of their first year in a master's degree program in operations research. While I had in mind what was clearly the "ideal" way to solve a certain problem in applied graph theory, the variety of solutions they generated was both refreshing and inspiring. This article aims to present this seemingly innocent problem and several different solutions. More important than the specific answer are the various routes that lead to the final destination. As a by-product, we will be treated to a minitour of elementary combinatorial mathematics—with excursions into matrices, generating functions, inclusion–exclusion, and recursion.

2. The problem A telecommunications company has n ground stations, to be linked with its two orbiting satellites. In how many ways can we link the ground stations to the satellites so that all parts of the resulting system can communicate with one another efficiently? (FIGURE 1 shows one possibility.) Every pair of ground stations can communicate (via either one or two intermediate satellite connections); the two satellites A and B can also communicate with each other via a ground station. This is to be achieved, moreover, using the smallest possible number of links.

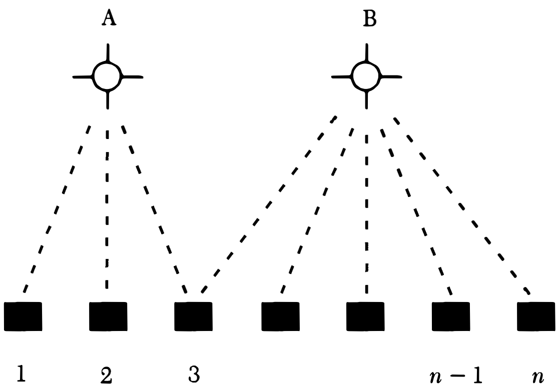


FIGURE 1
A possible configuration of two satellites and n ground stations.

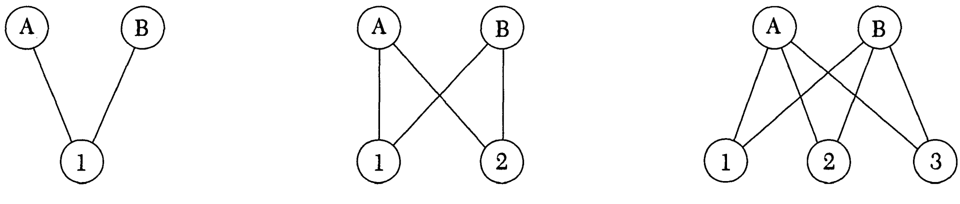
To abstract this problem, we see that a candidate solution is an undirected graph [4] on $n+2$ vertices, with the property that all vertices are accessible (either directly or indirectly) from one another using edges (proposed communication links). This is simply the requirement that the candidate graph G be *connected*: every pair of distinct vertices of G is joined by some path. For the sake of efficiency, no unnecessary edges should be present: that is, G should be without *cycles*. A connected graph without cycles is called a *tree*, and our problem requires a tree that includes all vertices: a *spanning tree*.

We denote by K_{n_1, n_2} the *complete bipartite graph* on n_1 and n_2 vertices. This graph consists of a set S_1 with n_1 vertices and a disjoint set S_2 with n_2 vertices; the edges of K_{n_1, n_2} are those joining every vertex of S_1 to every vertex of S_2. The original problem can now be restated:

Problem: Find the number t_n of spanning trees in $K_{2,n}$.

Notice that $K_{2,n}$ has $n+2$ vertices and $2n$ edges. It is straightforward to verify that any (spanning) tree has one fewer edge than the number of vertices. Thus a spanning tree T for our problem will have $n+2$ vertices and $n+1$ edges.

We look first at the simplest possible cases (see FIGURE 2):



(a) (b) (c)

FIGURE 2
The cases $n = 1, 2, 3$.

Case 1 ($n = 1$): Since $K_{2,1}$ is itself a tree, it contains precisely one spanning tree, so $t_1 = 1$.

Case 2 ($n = 2$): Removing any one of the four edges of $K_{2,2}$ leaves a spanning tree, giving $t_2 = 4$.

Case 3 ($n = 3$): Here T must have $n + 1 = 4$ edges selected from the $2n = 6$ edges of $K_{2,3}$. There are $\binom{6}{2} = 15$ ways of selecting the two edges to *remove* from $K_{2,3}$. So long as both such edges are not incident with the same ground station i, we obtain a spanning tree, giving $t_3 = 15 - 3 = 12$.

Further experimentation with small examples becomes increasingly complicated, so the following sections present seven different ways to solve this problem for arbitrary n. (Readers may wish first to guess the general solution.)

3.1 Direct counting Since $K_{2,n}$ has $n + 2$ vertices and $2n$ edges, any spanning tree T will end up selecting $n + 1$ out of these $2n$ edges. To aid in counting such trees, we note that the satellites A and B must be connected by a unique path in T, via some vertex $i \in \{1, 2, \ldots, n\}$. This leaves $n + 1 - 2 = n - 1$ edges of T to be used to join up the remaining $n - 1$ ground stations other than i. Consequently, each $j \neq i$ must be directly connected to precisely one of A or B, as shown in Figure 3.

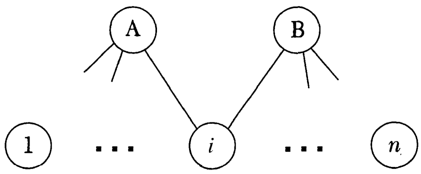


FIGURE 3
A unique vertex i is joined to both A and B in T.

We can now count the number of spanning trees directly. Exactly one of the n ground stations is selected as the common neighbor i of both A and B; there are thus n choices for i. Each $j \neq i$ can be chosen independently to connect to either A or B, giving two choices for each of these $n - 1$ stations. Thus $t_n = n \times 2 \times 2 \times \cdots \times 2 = n2^{n-1}$. In particular, $t_1 = 1(2^0) = 1$, $t_2 = 2(2^1) = 4$, and $t_3 = 3(2^2) = 12$, as previously found.

3.2 Conditioning Here we *condition* on the number k of edges in T that emanate from A. Note that $n + 1 - k$ edges then emanate from B; see Figure 4. As observed earlier, precisely one vertex i is joined directly to both A and B.

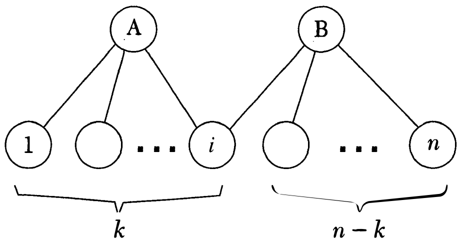


FIGURE 4
Conditioning on the k edges from A.

There are $\binom{n}{k}$ ways of selecting the k vertices that are joined to A. Since for each such selection there are k choices for the special vertex i, there are then $k\binom{n}{k}$ ways of constructing a spanning tree with k tree edges emanating from A. Since the situations for different k are mutually exclusive, we obtain

$$t_n = \sum_{k=1}^{n} k\binom{n}{k} = \sum_{k=0}^{n} k\binom{n}{k}. \tag{1}$$

The reader is invited to show that the sum (1) does in fact equal $n2^{n-1}$. [*Hint*: Start by differentiating the identity $(1+x)^n = \sum_{k=0}^{n}\binom{n}{k}x^k$ with respect to x.]

3.3 Deletion Instead of considering different ways to select the $n+1$ tree edges out of the $2n$ edges of $K_{2,n}$, we can focus on the $2n-(n+1)=n-1$ edges to be deleted. The important criterion is that no two deleted edges can be incident with the same ground station. The first deleted edge, say incident with ground station i_1, can be any of the $2n$ edges of $K_{2,n}$; the second edge, incident with ground station i_2, can be any of $2n-2$ edges, as it must avoid isolating vertex i_1. Continuing, the third deleted edge can be chosen from any of $2n-4$ edges, as it must avoid isolating either vertex i_1 or i_2. In general, the $(j+1)$st deleted edge can be chosen from any of $2n-2j$ edges. Multiplying these possibilities gives $\prod_{j=0}^{n-2}(2n-2j)$. However, we have seriously overcounted, since selecting in order edges $e_1, e_2, \ldots, e_{n-1}$ gives the same end result as selecting these edges in a different order. Indeed, every permutation of $e_1, e_2, \ldots, e_{n-1}$ gives the same result, so that this product needs to be divided by $(n-1)!$. Simplifying the result (thanks, reader!) again yields $t_n = n2^{n-1}$.

3.4 Inclusion–exclusion The *inclusion–exclusion* principle [2] lets us count items by developing successive overestimates and underestimates. Here we are interested in selecting $n+1$ edges to form a spanning tree. There are $N = \binom{2n}{n+1}$ selections in all, but many of them isolate a ground station. Define E_i to be the set of selections of $n+1$ edges in which ground station i is isolated, $i = 1, 2, \ldots, n$. Notice that in any selection of $n+1$ edges, vertices A and B cannot be isolated.

We are interested in $|\overline{E}_1 \cap \overline{E}_2 \cap \cdots \cap \overline{E}_n|$, the number of selections that do not isolate any ground station. This quantity can be found by invoking the inclusion–exclusion principle, which can be stated as follows:

$$\begin{aligned}&\left|\overline{E}_1 \cap \overline{E}_2 \cap \cdots \cap \overline{E}_n\right| \\ &\quad = N - \sum_i |E_i| + \sum_{i<j} |E_i \cap E_j| - \cdots + (-1)^n |E_1 \cap E_2 \cap \cdots \cap E_n|.\end{aligned}$$

Here $|E_i| = \binom{2n-2}{n+1}$ for each i since none of the $n+1$ selected edges can be incident with i. Similarly $|E_1 \cap E_2 \cap \cdots \cap E_k| = \binom{2n-2k}{n+1}$, and the above expansion becomes

$$t_n = \sum_{k=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} (-1)^k \binom{n}{k}\binom{2n-2k}{n+1}. \tag{2}$$

Since the left-hand side of (2) must equal $n2^{n-1}$, we have established a combinatorial identity, one that is not immediately obvious. The reader is invited to verify this identity directly for small values of n. (Reference [3] is an excellent source of combinatorial identities. In the present case, (2) was established by counting a set in two different ways, and then equating the results.)

3.5 Matrix algebra The *matrix-tree theorem* [1] counts the number of spanning trees of any undirected graph G on r vertices, by evaluating a certain determinant. This theorem states that the number of spanning trees of G is given by any cofactor of the matrix $M(G) = \mathrm{Deg}(G) - \mathrm{Adj}(G)$, where the $r \times r$ diagonal matrix $\mathrm{Deg}(G)$ has the degrees of the vertices (number of incident edges) along the diagonal and the $r \times r$ adjacency matrix $\mathrm{Adj}(G)$ has a 1 as its (i, j) entry if vertices i and j are adjacent (joined by an edge) and a 0 otherwise.

Here $G = K_{2,n}$, and if we order the vertices as $A, B, 1, 2, \ldots, n$ then

$$M(K_{2,n}) = \begin{array}{c} \\ A \\ B \\ 1 \\ 2 \\ \vdots \\ n \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & 1 & 2 & \cdots & n \end{array} \\ \begin{pmatrix} n & 0 & -1 & -1 & \cdots & -1 \\ 0 & n & -1 & -1 & \cdots & -1 \\ -1 & -1 & 2 & 0 & \cdots & 0 \\ -1 & -1 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & 0 & 0 & \cdots & 2 \end{pmatrix} \end{array}.$$

It is left to the (overworked) reader to delete the first row and column of $M(K_{2,n})$ and then evaluate the resulting determinant, say by expanding along the new first row. With a little determination, one obtains the familiar final answer $n2^{n-1}$.

3.6 Cryptomorphic approach We can sometimes solve a problem by recognizing it as a disguised ("cryptomorphic") version of a more familiar (or seemingly more tractable) problem. In the present case, we will set up a 1–1 correspondence between the spanning trees T of $K_{2,n}$ and the edges of the hypercube H_n. The hypercube is an n-dimensional generalization of the square ($n = 2$) and the cube ($n = 3$); H_n has 2^n vertices, each of degree n.

To provide this 1–1 correspondence, consider a spanning tree T of $K_{2,n}$. We construct an n-vector $\vec{x} = (x_1, x_2, \ldots, x_n)$ from T in the following way. For each ground station i, we assign $x_i = 0$ if i is joined only to A, $x_i = 1$ if i is joined only to B, and $x_i = *$ if i is joined to both. Since there is a unique ground station joined to both A and B, $\vec{x}$ contains exactly one $*$. Figure 5 illustrates this construction for a particular spanning tree T of $K_{2,5}$, giving $\vec{x} = (0, 1, *, 0, 1)$. Now we observe that $\vec{x}$ corresponds to a unique *edge* of H_5: namely, the edge connecting $(0, 1, 0, 0, 1)$ and $(0, 1, 1, 0, 1)$. In general, the constructed vectors $\vec{x}$, and thus the spanning trees of $K_{2,n}$, uniquely correspond to the edges of H_n. Our equivalent problem, now that the disguise has been revealed, is to count the edges in H_n. Since each of the 2^n vertices of H_n has degree n, the total number of edges is $\frac{1}{2}n2^n = n2^{n-1}$ [Why?].

3.7 Recursion Many counting problems can be reduced to similar problems of smaller size. This leads to a recurrence relation that can be solved to produce the general answer.

Specifically, a spanning tree involving stations $2, \ldots, n$ can be extended to include station 1 in two ways. First, we can add a single edge from 1 to either A or B; so the t_{n-1} spanning trees for $2, \ldots, n$ give rise to $2t_{n-1}$ spanning trees for $1, 2, \ldots, n$. Alternatively, there are 2^{n-1} spanning trees in which station 1 is joined to *both* A and B. The following recurrence relation, with initial condition $t_0 = 0$, results:

$$t_n = 2t_{n-1} + 2^{n-1}, \quad \text{for } n \geq 1. \tag{3}$$

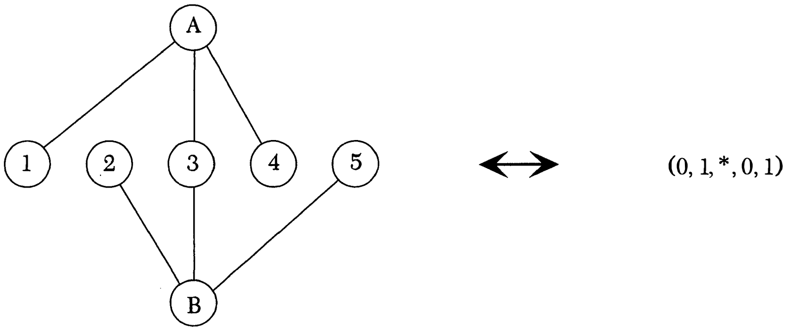


FIGURE 5
A spanning tree T and its corresponding vector $\vec{x}$.

One way to solve the recurrence (3) is with *generating functions* (see, e.g., [**2**]). [*Hint*: Define $T(x) = \sum_{n=0}^{\infty} t_n x^n$ and exploit the recurrence relation to obtain the closed form $T(x) = x(1-2x)^{-2}$. Using binomial coefficients, expand this expression in terms of powers of x to obtain $T(x) = \sum_{k=0}^{\infty}(k+1)2^k x^{k+1}$ and thus $t_n = n2^{n-1}$.] Yet again we obtain the familiar answer, but by a much more circuitous route.

Acknowledgment. We thank the referees for helpful suggestions that greatly improved the exposition.

REFERENCES

1. Ronald Gould, *Graph Theory*, Benjamin/Cummings, Menlo Park, CA, 1988.
2. Ralph P. Grimaldi, *Discrete and Combinatorial Mathematics*, 4th edition, Addison-Wesley, Reading, MA, 1999.
3. John Riordan, *Combinatorial Identities*, R. E. Krieger, Huntington, NY, 1979.
4. Robin J. Wilson and John J. Watkins, *Graphs: An Introductory Approach*, Wiley, New York, NY, 1990.

Illegal Aliens and Waiting Times

EUGENE F. KRAUSE
University of Michigan
Ann Arbor, MI 48109-1109

Introduction In this paper we apply elementary probability to a problem suggested by the current political debate over a vexing social issue. An interesting dividend is the unexpected appearance of Euler's number e.

The population of the United States consists of about 280 million legal residents and about 5 million illegal aliens. A two-part get-tough policy has been proposed: (1) legal residents would be issued non-counterfeitable identification cards; (2) police would be allowed to stop people at random to check for identification. Legal residents would, of course, be let go; illegal aliens would be deported. Opponents of the proposed policy describe it as unacceptably intrusive. Proponents counter that it would be only a minor nuisance. Leaving aside all political questions, can we use mathematics to quantify just *how* inconvenient such a policy would be? Specifically, can we answer the following questions?

QUESTION 1. *How many times should a legal resident expect to be stopped for identification if the policy remains in effect until every illegal is deported? Until half of the illegals are deported?*

QUESTION 2. *If each legal resident is willing to be stopped just once (on average), what fraction of the population of illegals will escape deportation? (We shall see that the answer to this question is* $1/e$.)

A mathematical model We make three broad assumptions to keep our model simple. (1) No legal residents are added to or subtracted from the population while the program is in progress. (2) No illegal aliens are added to the population while the program is in progress, and none are subtracted except those deported. (3) The police check for identification in a truly random fashion.

One can object that these assumptions are unrealistic, and thus any mathematical conclusions derived from them will not fit the real-world situation. That is, of course, the fundamental difficulty of "applied mathematics." One way to mitigate the difficulty is to make more realistic assumptions—and pay for the added realism with more complicated mathematics. Another way out, and the one we will take, is to agree to interpret the predictions of the (simple) mathematical model as only rough approximations to what might happen in the real world.

On the basis of our assumptions we now view the United States as a large urn containing 280 million black balls (legal residents) and 5 million red balls (illegals). Balls are drawn from this urn, at random, one at a time. When a black ball is drawn, it is returned to the urn before the next draw. When a red ball is drawn, it is permanently removed. Thus our questions take the following forms.

QUESTION 1. *On average, how many times is each black ball drawn before all of the red balls are removed? Before half of the red balls are removed?*

QUESTION 2. *What fraction of the original number of red balls can be expected to remain in the urn after each black ball has been drawn once (on average)?*

Posing and solving the mathematical problems Ultimately we are interested in questions concerning the drawing of balls from an urn with partial replacement: blacks go back, reds do not. To answer those questions, though, it turns out (as one might expect) that we need to answer analogous but simpler questions about two "pure" situations: drawing in which both blacks and reds are replaced, and drawing in which neither blacks nor reds are replaced.

To distinguish among the various modes of drawing, a rather elaborate notation seems to be necessary. Since our questions have to do with "waiting times" (the number of draws one should expect to make to achieve some end), our notation is built around the letter W. An unadorned W denotes drawing with replacement. $\overline{W}$ denotes drawing without replacement, and $\hat{W}$ denotes drawing with black balls replaced and red balls not replaced. Subscripts indicate the numbers of red (left subscript) and black (right subscript) balls originally in the urn; superscripts (left for red, right for black) indicate the number of balls we are waiting for. Here are the formal definitions.

DEFINITIONS. An urn contains r red and b black balls. Balls are drawn randomly one at a time.

(1) ${}_r^kW_b$ = waiting time for the k^{th} drawing of a red ball, where drawing is done with replacement.

(1′) ${}_rW_b^k$ = waiting time for the k^{th} drawing of a black ball, where drawing is done with replacement.

(2) ${}_r^k\overline{W}_b$ = waiting time for the k^{th} drawing of a red ball, where drawing is done without replacement.

(2′) ${}_r\overline{W}_b^k$ = waiting time for the k^{th} drawing of a black ball, where drawing is done without replacement.

(3) ${}_r^k\hat{W}_b$ = waiting time for the k^{th} drawing of a red ball, where blacks are replaced but reds are not.

(4) ${}_r\hat{W}_b^k$ = waiting time for the k^{th} drawing of a black ball, where blacks are replaced but reds are not.

For example, ${}_5^3\hat{W}_9$ denotes the waiting time for the drawing of the third red ball from an urn containing 5 red and 9 black balls, where black balls are replaced but reds are not. Notice that ${}_x^kW_y = {}_yW_x^k$ and ${}_x^k\overline{W}_y = {}_y\overline{W}_x^k$, but ${}_x^k\hat{W}_y \neq {}_y\hat{W}_x^k$. Thus we seek four formulas. They are the substance of Theorems 1–4 which follow.

THEOREM 1.

$$ {}_r^kW_b = k\left(1 + \frac{b}{r}\right) \tag{5} $$

Proof. It is well known that for Bernoulli trials, which is the situation when drawing is done with replacement, the waiting time for an event E is the reciprocal of its probability. Thus

$$ {}_r^1W_b = \frac{1}{r/(r+b)} = \frac{r+b}{r} = 1 + b/r, $$

and obviously ${}_r^kW_b = k \cdot {}_r^1W_b$. ∎

By the symmetry that we observed following our six definitions,

$$ {}_rW_b^k = k\left(1 + \frac{r}{b}\right). \tag{5′} $$

THEOREM 2.

$$ {}_r^k\overline{W}_b = k\left(1 + \frac{b}{r+1}\right) \tag{6} $$

Proof. This result seems to be less well known than (5), although it apparently has a kind of obscure folklore status. An intuitive plausibility argument goes something like this. Imagine a record of draws of all of the red (R) and black (B) balls. Say it looks like this:

$$ \underbrace{\qquad}\ \mathrm{R}\ \underbrace{\mathrm{BBB}}\ \mathrm{R}\ \underbrace{\mathrm{B}}\ \mathrm{R}\ \underbrace{\quad}\ \mathrm{R}\ \underbrace{\mathrm{BBBB}}\ \mathrm{R}\ \cdots $$

Then the r reds partition the b blacks into $r+1$ strings of repeating blacks. In the example above those strings have lengths $0, 3, 1, 0, 4, \ldots$. The average length of each string of repeating blacks is thus $b/(r+1)$. Hence the waiting time for the k^{th} red is

$$ k\left(\frac{b}{r+1}\right) + k = k\left(1 + \frac{b}{r+1}\right). $$

A careful proof of (6) can be made by treating r as an arbitrary natural number, inducting on k, and, within that induction, inducting on b. ∎

Again, by symmetry,

$$ {}_r\overline{W}_b^k = k\left(1 + \frac{r}{b+1}\right). \tag{6′} $$

THEOREM 3.

$$ {}_r^k\hat{W}_b = k + b \sum_{i=0}^{k-1} \frac{1}{r-i}, \tag{7} $$

Proof. Theorem 3 follows easily from Theorem 1. Until the first red ball is drawn, we are drawing *with* replacement. Thus, by Theorem 1,

$$ {}_r^1\hat{W}_b = {}_r^1W_b = 1 + \frac{b}{r}. $$

After the first red ball has been drawn, we again draw with replacement until the second red ball is drawn. Thus

$$ {}_r^2\hat{W}_b = {}_r^1\hat{W}_b + {}_{r-1}^{\ \ 1}W_b = \left(1 + \frac{b}{r}\right) + \left(1 + \frac{b}{r-1}\right). $$

The extension of this argument to yield Theorem 3 is obvious. ∎

Computationally, Theorem 3 leaves much to be desired. In our motivating problem $r = 5$ million and $b = 280$ million. To use Theorem 3 to find the waiting time until the last red ball has been chosen, we would have to sum the first 5 million terms of the harmonic series. Fortunately, Euler was good enough to leave us an approximation formula (see, for example, [**1**, p. 538]) for such partial sums,

$$ \sum_{i=1}^{r} \frac{1}{i} \approx \ln r + \frac{1}{2r} + 0.57721 \tag{8} $$

(where 0.57721 is an approximation to Euler's constant, γ), which yields a computational friendly approximation to the formula in (7).

COROLLARY 1.

$$ {}_r^k\hat{W}_b \approx k + b\left[\ln\frac{r}{r-k} - \frac{k}{2r(r-k)}\right] \quad \text{if } k < r; \tag{9} $$

$$ {}_r^k\hat{W}_b \approx r + b\left[\ln r + \frac{1}{2r} + 0.57721\right] \quad \text{if } k = r. \tag{10} $$

Proof. The case $k = r$ in (10) is an immediate consequence of (8) and the observation that

$$ \sum_{i=0}^{r-1} \frac{1}{r-i} = \sum_{i=1}^{r} \frac{1}{i}. $$

The case $k < r$ in (9) follows from the observation that

$$ \sum_{i=0}^{k-1} \frac{1}{r-i} = \sum_{i=1}^{r} \frac{1}{i} - \sum_{i=1}^{r-k} \frac{1}{i} $$

and two substitutions from (8). ∎

COROLLARY 2. *Drawing is done with black balls replaced and red balls not. F is a rational number of the form m/r with $m = 0, 1, 2, \dots, r$. To reduce the population of red balls from r to Fr ("reduce it by a factor of F"), each black ball should expect to* *be drawn approximately*

$$-\ln F - \frac{1-F}{2rF} \quad \text{times if } F \neq 0; \tag{11}$$

$$\ln r + \frac{1}{2r} + 0.57721 \quad \text{times if } F = 0. \tag{12}$$

Proof. If the population of reds is to drop from r to Fr, then $(1-F)r$ reds must be drawn. The number of times, on average, that each black ball must be drawn then is

$$\frac{1}{b}\left[{}^{(1-F)r}_{\;\;r}\hat{W}_b - (1-F)r\right]. \tag{13}$$

If $F \neq 0$, then $(1-F)r < r$ and substituting from (9) into (13) yields (11). If $F = 0$, then substituting from (10) into (13) yields (12). ■

THEOREM 4.

$${}_r\hat{W}_b^k = k + r\left(1 - [b/(b+1)]^k\right) \tag{14}$$

Proof. As Theorem 3 followed from Theorem 1, so Theorem 4 follows from Theorem 2. But this time the argument is a bit more difficult. We prove (14) by induction on k. For the case $k = 1$ we note that until the first black ball is drawn we are drawing *without* replacement. Thus, by (6′),

$${}_r\hat{W}_b^1 = {}_r\overline{W}_b^1 = 1 + \frac{r}{b+1} = 1 + r\left(1 - \left[\frac{b}{b+1}\right]^1\right).$$

Now we assume (14) for k and consider the case $k+1$. After the k^{th} black has been drawn, the number of reds that have been drawn is ${}_r\hat{W}_b^k - k$ which, by the inductive hypothesis, is

$$r\left(1 - \left[\frac{b}{b+1}\right]^k\right).$$

Thus the number of reds remaining in the urn is

$$r_1 = r\left(\frac{b}{b+1}\right)^k$$

and, of course, the number of blacks in the urn is still b. Now drawing resumes, without replacement, until the next black is drawn. Hence

$$\begin{aligned} {}_r\hat{W}_b^{k+1} &= {}_r\hat{W}_b^k + {}_{r_1}\hat{W}_b^1 \\ &= k + r\left(1 - \left[\frac{b}{b+1}\right]^k\right) + \left(1 + r_1\left[1 - \frac{b}{b+1}\right]\right) \\ &= (k+1) + r\left(1 - \left[\frac{b}{b+1}\right]^{k+1}\right). \end{aligned}$$ ■

COROLLARY 3. *For large b, the number of red balls remaining in the urn after black balls have been drawn k times is approximately $re^{-k/b}$. That is, the red population has been shrunk by a factor of $e^{-k/b}$.*

Proof. By Theorem 4,

$$ {}_r\hat{W}_b^k = k + r\left(1 - [b/(b+1)]^k\right) = k + r\left(1 - \left[(1+1/b)^b\right]^{-k/b}\right). $$

Thus, for large b,

$$ {}_r\hat{W}_b^k \approx k + r(1 - e^{-k/b}). $$

The k in this sum represents the k draws of black balls. Thus $r(1 - e^{-k/b})$ red balls have been drawn, and thus $re^{-k/b}$ remain in the urn. ■

Note. To see that Corollaries 2 and 3 are consistent with each other, for large r and b and F not too near 0, let x be the number of draws per black ball to reduce the red population from r to Fr. By Corollary 2, $x \approx -\ln F$. By Corollary 3, $F \approx e^{-xb/b}$. Obviously the equations corresponding to these estimates are equivalent.

Answering the original questions Recall Question 1: On average, how many times is each black ball (legal resident) drawn (stopped for identification) before 100% of the red balls (illegal aliens) are removed (deported)? Ibid 50% removed (deported)? The answer to the 100% part of Question 1 is found by setting $r = 5{,}000{,}000$ in (12): each black ball (legal) should expect to be drawn (stopped) approximately 16.00216 times. The answer to the 50% part of Question 1 is found by setting $r = 5{,}000{,}000$ and $F = \frac{1}{2}$ in (11): each black ball (legal) should expect to be drawn (stopped) approximately 0.69315 times.

Question 2 asked: What fraction of the red balls (illegals) would remain in the urn (escape deportation) if, on average, each black ball were drawn once? One way to answer this question is to set (11) equal to 1 and solve for F:

$$ 1 = -\ln F - \frac{(1-F)}{10{,}000{,}000\,F} \approx -\ln F \Rightarrow \approx \frac{1}{e}. $$

A more direct way is to set $k = b$ in Corollary 3.

Three concluding remarks 1. The 100% and 50% values in Question 1 were chosen arbitrarily. To shed more light on that question we have produced a table (Table 1) showing a range of values for the fraction of illegal aliens to be deported—the quantity $(1 - F)$—and the corresponding approximate values of the number of times each legal resident should expect to be stopped—the quantities in (11) and (12). The values in this table are for the case of $r = 5$ million. Notice, however, that the expressions in (11) and (12), and hence the values in Table 1, do not depend on b, the number of legal residents. It does not matter if there are 280 million legal residents, or 1 billion, or 1000. In order to catch (say) 90% of the 5 million illegals, each legal should expect to be stopped about 2.3 times.

TABLE 1

Percent of illegals to deport	10	30	50	70	90	99.9	99.999	99.99998
Number of stops per legal resident	0.11	0.36	0.69	1.20	2.30	6.91	11.51	14.92

2. Theorem 4 was not needed to answer Questions 1 and 2. Its inclusion in this paper was an esthetic decision. We needed it to tie down a mathematical loose end and complete the theoretical picture.

3. One reason for moving from the task of answering a specific real-world question to the task of building a general theoretical model is that the theoretical model can be used to answer other specific questions in other contexts.

EXAMPLE 1. In some supermarkets strawberries are put out on a table and customers are allowed to choose their own berries one by one. Typically the customer will pick up a berry and look it over. If it is good, he will keep it. If it is rotten, he will put it back on the table. Suppose there are 200 good berries and 100 rotten ones on the table, and a customer arrives who is determined to get 60 good berries. How many berries should that customer expect to handle?

Solution. Think of the table as an urn. The rotten berries are the black balls (the ones that are replaced), so $b = 100$. The good berries are the red balls (the ones that are not replaced), so $r = 200$. The number of berries that the customer should expect to handle until he has 60 good ones is ${}_{200}^{60}\hat{W}_{100}$ which, by (9), is approximately 95.56.

EXAMPLE 2. Same context, but now the produce manager has supplied a garbage can next to the strawberry table into which customers throw their rotten berries. How many berries should the customer expect to handle to get 60 good ones?

Solution. The expected number to handle is ${}_{200}^{60}\overline{W}_{100}$ which, by (6), is approximately 89.85.

EXAMPLE 3. Original context (no garbage can), but now the customer will get disgusted and quit after he has handled his tenth rotten berry. How many good ones will he have?

Solution. He will have ${}_{200}\hat{W}_{100}^{10} - 10$ good berries which, by (14), is approximately 18.94.

REFERENCE

1. K. Knopp, *Theory and Application of Infinite Series*, Hafner Publishing Co., New York, NY, 1947.

Heron Triangles: An Incenter Perspective

K.R.S. SASTRY
Jeevan Sandhya
Raghuvana Halli
Bangalore 560 062
India

Introduction Heron lived in Alexandria in the first century AD. He gave the triangle area formula $\Delta = \sqrt{s(s-a)(s-b)(s-c)}$. The triple (a, b, c) describes the sides of triangle ABC and $s = \frac{1}{2}(a+b+c)$. We use the word "side" also to mean the length of a side of the triangle. Furthermore, the discovery of the $(13, 14, 15)$ triangle with area 84 is attributed to Heron. To honor Heron, a triangle with rational sides and area is called a *Heron triangle*. These rationals can always be made integers. Hence our discussion considers Heron triangles for which the sides and the area are integers. In **[5]** Dickson mentions that Hoppe studied a special class of Heron triangles in which the sides are in arithmetic progression. Hoppe describes such triangles as triples of the form

$$(a, b, c) = \left(3(\lambda^2 + \mu^2), 2(3\lambda^2 + \mu^2), 9\lambda^2 + \mu^2\right).$$

Notice that $c + a = 2b$ holds. Let's rewrite the preceding equation as

$$\frac{c+a}{b} = \frac{2}{1}. \tag{1}$$

Let us leave Hoppe for a while and visit Euclid. The well-known angle bisector theorem says: *The bisector of an angle of a triangle sections the opposite side into segments that are in the ratio of the other sides*. Another well-known theorem asserts: *The incenter of a triangle is the concurrence point of its angle bisectors.* These theorems can be found in **[4]**.

Consider triangle ABC together with the incenter I. Extend BI to meet AC at E as shown in FIGURE 1:

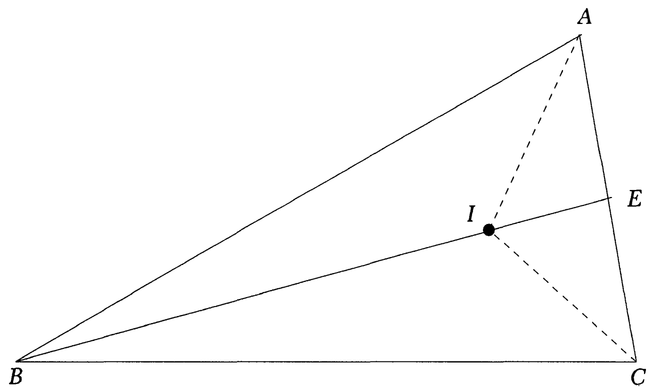


FIGURE 1
Viewing Hoppe geometrically.

Applying the angle bisector theorem to both triangles ABE and BCE gives

$$\frac{BI}{IE}=\frac{AB}{AE}=\frac{BC}{CE}=\frac{AB+BC}{AC}=\frac{c+a}{b}.$$

Compare the preceding equation to (1). Are you surprised? To put it another way, Hoppe described Heron triangles in which the incenter sections some angle bisector in the ratio 2:1. This geometric interpretation of Hoppe's work motivates the following problem:

> Describe Heron triangles ABC in which the incenter I sections the angle bisector BE in the ratio $(c+a):b=u:v$, $u>v$ and $\gcd(u,v)=1$. (2)

In this note we provide this description.

Background material An angle α is called a *Heron angle* if $\sin\alpha$ and $\cos\alpha$ are rational. This is equivalent to $\tan\left(\frac{\alpha}{2}\right)$ being rational. The equivalence holds because if $\tan(\alpha/2)=q/p$, a rational number with $p>q$ and $\gcd(p,q)=1$, then

$$\sin\alpha=\frac{2pq}{p^2+q^2}\quad\text{and}\quad\cos\alpha=\frac{p^2-q^2}{p^2+q^2}. \tag{3}$$

Consider a triangle ABC; let A, B, C denote the measures of the angles and a, b, c the measures of the sides opposite these angles. Expanding the equation $\frac{B}{2}+\frac{C}{2}=\frac{\pi}{2}-\frac{A}{2}$ by the tangents (or the cotangents) one finds:

$$\cot\left(\frac{A}{2}\right)+\cot\left(\frac{B}{2}\right)+\cot\left(\frac{C}{2}\right)=\cot\left(\frac{A}{2}\right)\cot\left(\frac{B}{2}\right)\cot\left(\frac{C}{2}\right). \tag{4}$$

The law of sines is

$$\frac{a}{\sin A}=\frac{b}{\sin B}=\frac{c}{\sin C}. \tag{5}$$

From the equation $\cot\left(\frac{A}{2}\right)=\sqrt{\frac{1+\cos A}{1-\cos A}}$ it is easy to deduce that

$$\cot\left(\frac{C}{2}\right)\cot\left(\frac{A}{2}\right)=\frac{s}{s-b}. \tag{6}$$

The main result For convenience let us write the rational number u/v as λ. From (2) we then have $c+a=\lambda b$, with $\lambda>1$. Therefore $s=\frac{1}{2}(\lambda+1)b$ and $s-b=\frac{1}{2}(\lambda-1)b$. Then (6) tells us that $\cot\left(\frac{C}{2}\right)\cot\left(\frac{A}{2}\right)=\frac{\lambda+1}{\lambda-1}$. We put this value into (4) and write the resulting equation as a quadratic in $\cot(\frac{A}{2})$:

$$(\lambda-1)\cot^2\left(\frac{A}{2}\right)-2\cot\left(\frac{B}{2}\right)\cot\left(\frac{A}{2}\right)+(\lambda+1)=0.$$

The quadratic formula now yields

$$\cot\left(\frac{A}{2}\right)=\frac{1}{\lambda-1}\left[\cot\left(\frac{B}{2}\right)+\sqrt{\cot^2\left(\frac{B}{2}\right)-(\lambda^2-1)}\right].$$

We need not consider the negative root: it simply interchanges the expressions for the sides a and c (we do not regard triangles different if the sides appear in a different order). Here $\cot\left(\frac{A}{2}\right)$, $\cot\left(\frac{B}{2}\right)$, and $\cot\left(\frac{C}{2}\right)$ are rational numbers. Therefore the expression under the radical sign must be y^2, a rational square. Let us put $\cot\left(\frac{B}{2}\right) = x$. Then $x^2 - (\lambda^2 - 1) = y^2$, or

$$\frac{x+y}{\lambda - 1} = \frac{\lambda + 1}{x - y} = \frac{m}{n}, \quad \gcd(m, n) = 1\,.$$

Here m/n represents this common rational number. When solved for x and y,

$$x = \frac{(\lambda - 1)m^2 + (\lambda + 1)n^2}{2mn}, \quad y = \frac{(\lambda - 1)m^2 - (\lambda + 1)n^2}{2mn}.$$

From $\lambda > 1$ we have $x > 0$. Also $x + y > 0$. Here y represents the positive square root for the pair (m, n), where $m > \sqrt{\frac{\lambda+1}{\lambda-1}}\,n$, and y represents the negative square root for the pair (m, n) if $0 < m < \sqrt{\frac{\lambda+1}{\lambda-1}}\,n$. Either pair (m, n) may be used to generate the same Heron triangle. (We show this in the example to come later.) The values x and y then determine

$$\cot\left(\frac{A}{2}\right) = \frac{m}{n}, \quad \cot\left(\frac{B}{2}\right) = \frac{(\lambda - 1)m^2 + (\lambda + 1)\,n^2}{2mn}, \quad \cot\left(\frac{C}{2}\right) = \frac{(\lambda + 1)n}{(\lambda - 1)m}. \tag{7}$$

From (7) we first determine $\tan(A/2)$, $\tan(B/2)$, and $\tan(C/2)$. Next we use these values into (3) to determine the values of $\sin A$, $\sin B$, and $\sin C$. Then we use the sine rule (5) and replace λ by u/v. A routine derivation leads to the side length expressions for the Heron triangle, namely

$$\begin{aligned}(a, b, c) = \Big(&(u - v)^2 m^2 + (u + v)^2 n^2,\\ &2v\big[(u - v)m^2 + (u + v)n^2\big], (u^2 - v^2)(m^2 + n^2)\Big).\end{aligned} \tag{8}$$

We dropped the constant of proportionality in (8) because we keep $\gcd(a, b, c) = 1$. The above triangle has the area $\Delta = 2mnv(u^2 - v^2)[(u - v)m^2 + (u + v)n^2]$. Moreover, it is easy to check that $(c + a) : b = u : v$. Thus the triples (8) generate Heron triangles as (2) indicates. Observe that setting $u = 2$ and $v = 1$ in (8) gives Hoppe's result. But we have more to say:

> Let us fix u and v and vary the parameters m and n over $\mathbb{N}$. The resulting triples in (8) describe a subset of Heron triangles. Each member triangle of this subset exhibits the property in (2). $(*)$

The assertion $(*)$ holds because the ratio $(c + a) : b = u : v$ is independent of m and n. But the story doesn't end there:

> Suppose now we vary $u, v : u > v$ over $\mathbb{N}$. Apply $(*)$ successively for each such pair (u, v). The triples in (8) describe the complete set of Heron triangles. $(**)$

An interesting consequence of $(**)$ is this: The same Heron triangle may be obtained from (8) more than once. This is because the incenter of a given triangle sections at least two angle bisectors in distinct ratios $u_1 : v_1$ and $u_2 : v_2$. These ratios define distinct infinite subsets of Heron triangles. But each of these subsets contains the given triangle. Notice that an equilateral triangle is *not* a Heron triangle (why?) and does not have this property. Following is a numerical illustration to describe the situation more clearly.

A numerical example Consider the Heron triangle $ABC:(a, b, c) = (25, 36, 29)$; one can check that the area is an integer. The incenter of this triangle sections the bisector of $\angle ABC$ in the ratio $u_1 : v_1 = (c+a) : b = 3:2$. These values for u and v in (8) give a subset of Heron triangles:

$$S_1 = \{(m^2 + 25n^2), 4(m^2 + 5n^2), 5(m^2 + n^2)\}.$$

Here $\lambda = 3/2$. The pair $(m, n) = (5, 2)$ is such that $m > \sqrt{5}n$. So S_1 determines the Heron triangle $(a, b, c) = (25, 36, 29)$ via the positive square root y (we should remember to make $\gcd(a, b, c) = 1$). However, the pair $(m, n) = (2, 1)$ is such that $0 < m < \sqrt{5}n$. In this case S_1 determines the same triangle $ABC:(a, b, c) = (29, 36, 25)$ via the negative square root y. Notice the interchange of values for a and c as we mentioned earlier.

Next, the incenter of the starting triangle ABC sections the bisector of $\angle BCA$ in the ratio $u_2 : v_2 = (a+b) : c = 13:5$. With these values (8) generates an infinite subset of Heron triangles:

$$S_2 = \{(16m^2 + 81n^2), 5(4m^2 + 9n^2), 36(m^2 + n^2)\}.$$

The values $m = 9$, $n = 8$ in S_2 gives us the Heron triangle $(a, b, c) = (36, 25, 29)$. Of course we should remember to reduce $\gcd(a, b, c)$ to 1. Can you explain why the sides appeared in a different order? In general, from the triples (8) it is possible to recover a triangle six times. We leave this determination to the reader.

Conclusion The determination of Heron triangles continues to be of interest. Many popular problems involve Heron triangles (see e.g., [**1**, **2**, **5**, **6**, **7**, **8**, **9**]). But the beauty of the subject is that you can find new ways to determine them. The present discussion suggests another direction: *Consider the excenter I_B opposite the vertex B of triangle ABC. Find the ratio in which I_B sections the bisector of angle ABC. Use this ratio to describe Heron triangles.*

A referee points out that, when c or a is the hypotenuse, the calculation $(c+a) : b = u : v$ produces the parameters u, v for a *Pythagorean triangle* (a Pythagorean triangle is a right triangle with integer sides, so it is a Heron triangle; see [**3**]. This fact is available from our discussion too: Set $m = n$. From (7) we have $\cot\left(\frac{A}{2}\right) = \frac{m}{n} = 1$ so $\angle BAC = \frac{\pi}{2}$. Then (8) yields the Pythagorean triples $(a, b, c) = (u^2 + v^2, 2uv, u^2 - v^2)$.

We recently became aware of the preprint [**2**], in which the authors extend their work in [**1**]. However, that extension is limited to the special case $u = 2\alpha$, $v = 1$, (where α is a positive integer) of the present discussion.

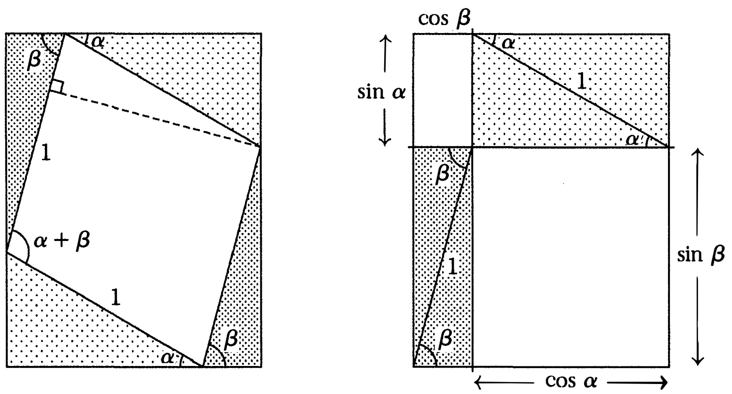
Acknowledgment. The author thanks the referees for their suggestions, which improved the readability of this note.

REFERENCES

1. R. A. Beauregard and E. R. Suryanarayan, Arithmetic triangles, this MAGAZINE 70 (1997), 105–115.
2. R. A. Beauregard and E. R. Suryanarayan, General arithmetic triangles and Bhaskara's equation (preprint).
3. R. A. Beauregard and E. R. Suryanarayan, Proof Without Words: Parametric representation of Pythagorean triples, this MAGAZINE 69 (1996), 189.
4. H. S. M. Coxeter, *Introduction to Geometry*, Wiley, New York, NY, 1961, 88–89.
5. L. E. Dickson, *History of the Theory of Numbers*, Vol. II, Chelsea, New York, NY, 1971, 171–201.
6. K. R. S. Sastry, Heron strikes again!, *Math. Comput. Ed.* 25 (1991), 242–250.
7. K. R. S. Sastry, Heron quadratics, *Function* 16 (Part 2, 1992), 35.
8. K. R. S. Sastry, Pythagorean triple problems, *Math. Comput. Ed.* 28 (1994), 320–331.
9. K. R. S. Sastry, Heron problems, *Math. Comput. Ed.* 29 (1995), 192–202.

Proof Without Words: The Sine of a Sum

The area of the white parallelogram on the left is $\sin(\alpha + \beta)$.



$$\sin(\alpha + \beta) = \sin\alpha \cdot \cos\beta + \cos\alpha \cdot \sin\beta.$$

—VOLKER PRIEBE AND EDGAR A. RAMOS
MAX-PLANCK-INSTITUT FÜR INFORMATIK
STUHLSATZENHAUSWEG 85
66123 SAARBRÜCKEN, GERMANY

Square Waves from a Black Box

TEMPLE H. FAY
Technikon Pretoria
and
University of Southern Mississippi
Hattiesburg, MS 39406-5045

STEPHAN V. JOUBERT
Technikon Pretoria
Pretoria 0001
South Africa

When we teach ordinary differential equations, or indeed any other course that uses software to facilitate instruction, we tell the students that we have one axiom:

The computer is not to be used as a black box!

We are sure many other instructors share our axiom. We recently ran across what we think is a fine example to make the point.

Spring models are useful examples in beginning DE courses, and we like to use the nonlinear hard/soft spring model

$$\dot{x} + ax + bx^3 = 0,$$

where $a > 0$. Here the mass on the spring has been normalized to 1 and the restoring force is taken to be the odd function $-ax - bx^3$. If $b > 0$, then the spring is called hard and all solutions are oscillatory and periodic; if $b < 0$, then the spring is called soft and for sufficiently small initial values, the solutions are oscillatory and periodic, but for other initial values the solutions grow without bound. For more details on this example and similar ones see **[1]** and **[2]**.

Consider the soft spring equation

$$\ddot{x} + x - x^3 = 0$$

with initial values $x(0) = 0$ and $\dot{x}(0) = 1/\sqrt{2}$. Using *Mathematica*'s ODE numerical routine **NDSolve**, our students produced solutions and phase plane trajectories much like those shown in FIGURE 1.

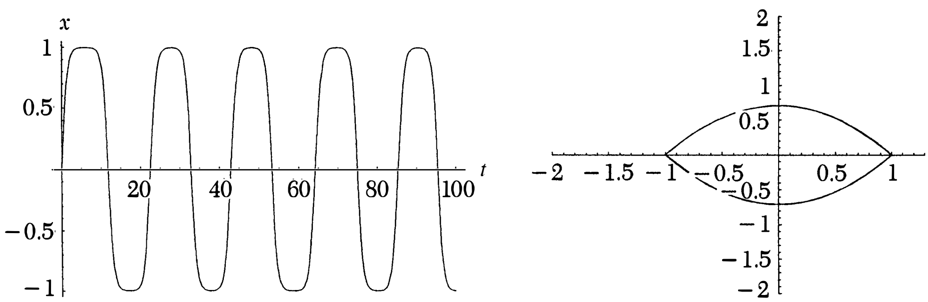


FIGURE 1

Based on this evidence, one would think that the solution to this initial value problem is periodic and that the solution is very nearly a square wave. This solution and trajectory were produced using the default precision set at 16. **NDSolve**, like many numerical schemes, proceeds step-by-step in an adaptive way, attempting to satisfy a certain truncation error tolerance. With precision 16, **NDSolve** tries for 6-digit precision at each step. The solution and trajectory shown in FIGURE 2 are for the same initial value problem, but with the precision set to 20, so that **NDSolve** tries for 10-digit precision.

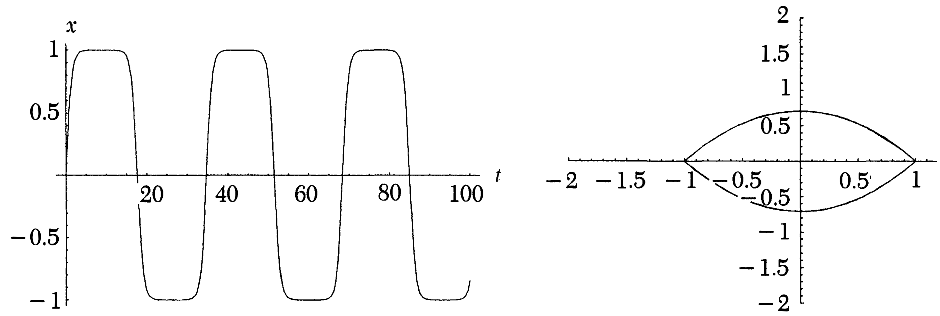


FIGURE 2

Certainly, the two solutions shown in FIGURES 1 and 2 are markedly different, but the trajectories appear to be identical. What is going on? With higher precision, we shouldn't see a change in the solution. The answer is that *both solutions are correct—for a while*.

The initial values for this problem lie on a trajectory in the phase plane called the *separatrix*. As time increases, the points $(x(t), \dot{x}(t))$ on the trajectory travel toward a saddle point at $(1,0)$. This saddle point is a stationary point for this equation. Hence once the trajectory reaches $(1,0)$ it should remain there forever since a stationary point represents zero velocity and zero acceleration. Unfortunately, the numerical algorithm being implemented by **NDSolve** eventually falls prey to truncation error (and perhaps roundoff error and propagation error). The x-value eventually becomes either greater or less than 1 and the $\dot{x}$-value becomes greater or less than 0. This puts the trajectory away from the stationary point and off it goes toward another saddle point at $(-1,0)$. Once the trajectory reaches this new stationary point, it remains there until the accumulated error "bumps it off" and away it goes again towards $(1,0)$. This is illustrated in FIGURE 3.

In FIGURE 3, the first frame shows the phase plane at time $t = 0$. The starting points for eight different trajectories

$$\{(0,0),(0,1/\sqrt{2}),(.25,.25),(.5,.5),(.7,-.7),(.8,-.8),(-.7,.7),(-.8,.8)\}$$

are shown as small dots. The two dashed intersecting parabolas form the separatrix for this equation. It divides the plane into separate regions of distinct motion. For example, the motion represented by a starting point within the bounded region containing the origin is oscillatory and periodic, and hence all the trajectories are closed curves within this region. The starting point $(0, 1/\sqrt{2})$ lies on the separatrix; the starting point $(0.5, 0.5)$ lies very close to but not on the separatrix.

The second frame in FIGURE 3 shows the trajectories for all starting points up to time $t = 1.5$. The point $(0,0)$ is a stationary point (a center) and the trajectory never leaves this point. All the trajectories for the other starting points have begun to trace

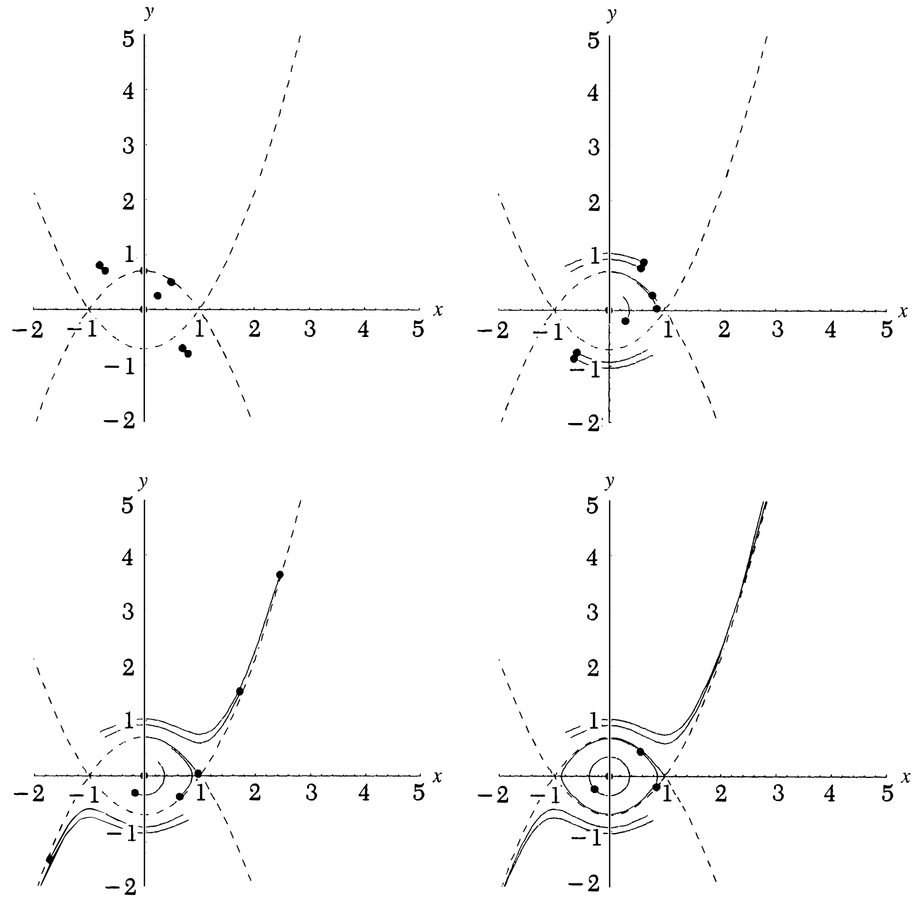


FIGURE 3

out curves moving in a clockwise direction. In the third frame, we see that the trajectory for the starting point $(0.5, 0.5)$ has passed through the x-axis and the trajectory for $(0, 1/\sqrt{2})$ has just about reached the saddle point $(1, 0)$ where it will remain "stuck" for a while and the other trajectories continue to move. The trajectory for $(0.25, 0.25)$ has curled around into the third quadrant and all the other trajectories are proceeding out of the graphics windows.

The last frame shows the trajectories for time $t = 20$. The point representing the endpoint of the trajectory which started at $(0, 1/\sqrt{2})$ remained stuck at $(1, 0)$ until approximately time $t = 13.2$ when it gets "bumped off" by accumulated truncation error and now begins to move along the separatrix toward the other saddle point at $(-1, 0)$. It will eventually reach $(-1, 0)$ and become "stuck" there for a while before it again gets "bumped off" at which time it travels along the separatrix toward $(1, 0)$ to repeat the cycle again, thus giving the appearance of a square wave solution.

This is all easier seen and appreciated in animation. A *Mathematica* 4.0 notebook, `blackbox.nb`, is available at `http://pax.st.usm.edu/~fay/project1/blackbox..nb`. Users may modify the parameters to investigate this and other spring equations for various starting points. *Mathematica* generates a number of frames similar to those in FIGURE 3, which can then be animated to show the movement of the trajectories and the "sticking."

It is interesting to repeat these observations, using higher and higher precisions in the **NDSolve** routine. The higher the precision, the longer the trajectory remains "stuck," as it should. It is also instructive to experiment with initial values different from but close to $(0, 1/\sqrt{2})$. If one takes the initial values $x(0) = 0$ and $\dot{x}(0) = 1/\sqrt{2} - 10^{-6}$, the solution is oscillatory and a plot of it very closely resembles a square wave. It is a good student exercise to predict the motion for starting points on other portions of the separatrix. The separatrix actually consists of six separate trajectories and two stationary (saddle) points.

We recommend the interesting paper [3] for a more in-depth discussion of precision problems and how to predict over what time interval one can expect a numerical solution to be accurate.

Acknowledgment. The authors thank the South African National Research Foundation and the Department of Mathematical Technology of the Technikon Pretoria for support. They also thank an anonymous referee for supplying the *Mathematica* module that we named `blackbox.nb`.

REFERENCES

1. T. H. Fay and S. V. Joubert, Energy and contour plots for the analysis of nonlinear differential equations, *Mathematics and Computer Education* 33 (1999), 67–77.
2. T. H. Fay and S. V. Joubert, Energy and the nonsymmetric nonlinear spring, *International Journal for Mathematics Education in Science and Technology* 30 (1999), 889–902.
3. R. Knapp and S. Wagon, Orbits worth betting on, *C·ODE·E*, Winter, 1996, 8–13.

A Bubble Theorem

OSCAR BOLINA
University of California
Davis, CA 95616-8633

J. RODRIGO PARREIRA
Cluster Consulting
Torre Mapfre pl 38
Barcelona, 080050
Spain

Introduction It is always a good practice to provide the physical content of an analytical result. The following algebraic inequality lends itself well to this purpose: For any finite sequence of real numbers $r_1, r_2, \dots, r_N$, we have

$$\left(r_1^3 + r_2^3 + \cdots + r_N^3\right)^2 \le \left(r_1^2 + r_2^2 + \cdots + r_N^2\right)^3. \tag{1}$$

A standard proof is given in [1]. An alternative proof follows from the isoperimetric inequality

$$A^3 \ge 36\pi V^2,$$

where A is the surface area and V the volume of any three-dimensional body. Setting the area $A = \sum_{i=1}^N 4\pi r^2$ and the volume $V = \sum_{i=1}^N (4/3)\pi r^3$ yields (1).

A bubble proof We give yet another proof, now using elements of surface tension theory and ideal gas laws to the formation and coalescence of bubbles. This proof, found in [2], runs as follows.

According to a well-known result in surface tension theory, when a spherical bubble of radius R is formed in the air, there is a difference of pressure between the inside and the outside of the surface film given by

$$p = p_0 + \frac{2T}{R}, \tag{2}$$

where p_0 is the (external) atmospheric pressure on the surface film of the bubble, p is the internal pressure, and T is the surface tension that maintains the bubble **[3]**.

Suppose initially that N spherical bubbles of radii $R_1, R_2, \dots, R_N$ float in the air under the same surface tension T and internal pressures $p_1, p_2, \dots, p_N$. According to (2),

$$p_k = p_0 + \frac{2T}{R_k}, \quad k = 1, 2, \dots N. \tag{3}$$

Now suppose that all N bubbles come close enough to be drawn together by surface tension and combine to form a single spherical bubble of radius R and internal pressure p, also obeying equation (2). When this happens, the product of the internal pressure p and the volume v of the resulting bubble formed by the coalescence of the initial bubbles is, according to the ideal gas law **[3]**, given by

$$pv = p_1 v_1 + \cdots + p_N v_N, \tag{4}$$

where v_k $(k = 1, 2, \dots, N)$ are the volumes of the individual bubbles before the coalescence took place. For spherical bubbles, (4) becomes

$$pR^3 = p_1 R_1^3 + \cdots + p_N R_N^3. \tag{5}$$

Substituting the values of p and p_k given in (2) and (3) into (5), we obtain

$$R^3 - R_1^3 - R_2^3 - \cdots - R_N^3 = \frac{2T}{p_0}\left(R_1^2 + R_2^2 + \cdots + R_N^2 - R^2\right). \tag{6}$$

Now, if the total amount of air in the bubbles does not change, the surface area of the resulting bubble formed by the coalescence of the bubbles is always smaller than the sum of the surface area of the individual bubbles before coalescence. Thus,

$$R_1^2 + R_2^2 + \cdots + R_N^2 \geq R^2. \tag{7}$$

Since the potential energy of a bubble is proportional to its surface area, (7) is a physical condition that the surface energy of the system is minimal after the coalescence.

It follows from (7) and the fact that p_0 and T are positive constants that the left hand side of equation (6) satisfies

$$R_1^3 + R_2^3 + \cdots + R_N^3 \leq R^3. \tag{8}$$

The equality, which implies conservation of volumes, holds when the excess pressure in the bubble film is much less the atmospheric pressure. Combining (7) and (8) yields the inequality (1), which is also valid for negative numbers.

Acknowledgment. O. B. would like to thank Dr. Joel Hass for pointing out the isoperimetric proof of (1), and FAPESP for support under grant 97/14430-2.

REFERENCES

1. G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Second Edition, Cambridge Mathematical Library, Cambridge, UK, 1988, p. 4.
2. H. Bouasse, *Capillarité et Phénomènes Superficiels*, Librairie Delagrave, Paris, France, 1924, p. 48.
3. A. Hudson and R. Nelson, *University Physics*, Harcourt Brace Jovanovich, Inc., New York, NY, 1982, p. 371 and p. 418.

A Periodic Property in $\mathbb{Z}_m$

SCOTT J. BESLIN
BRIAN K. HECK
Nicholls State University
Thibodaux, LA 70310

The following results arose serendipitously from correcting a student's errors on a test in graph theory. A circulant graph problem required the student to compute x^3 modulo 7 for each $x = 1, 2, \dots, 6$. Instead, she computed these powers modulo 6. Interestingly, such exponentiation yields $x^3 = x$ modulo 6 for every $x \in \{0, 1, 2, \dots, 5\}$. Equivalently, $x^3 = x$ for every element x in the ring $\mathbb{Z}_6$.

This result was somewhat unexpected. For suppose we seek an integer $m > 1$ such that, for some fixed integer $n > 1$, $x^n = x$ for all x in $\mathbb{Z}_m$. Remembering that $\mathbb{Z}_p$ is a field for prime p, we might think of Fermat's little theorem (see e.g., **[2]**): $x^p = x$ for every x in $\mathbb{Z}_p$. But primes are not the only possibilities for m. In this note we characterize those pairs (m, n) with the following property P:

The pair (m, n) has property P (the periodic property) if $x^n = x$ for all x in $\mathbb{Z}_m$.

We reserve the letter p to represent primes, and we will represent the elements of $\mathbb{Z}_m$ by $\{0, 1, 2, \dots, (m-1)\}$. We know from above that $(6, 3)$ has property P and that (p, p) has property P for all p. But Fermat's little theorem yields a deeper result: in the field $\mathbb{Z}_p$, $x^{p-1} = 1$ for all nonzero x. Furthermore, the multiplicative group of $\mathbb{Z}_p$ is cyclic, so, for odd primes, $p - 1$ is the *least* integer greater than 1 such that $x^{p-1} = 1$ for all nonzero x. It follows that in $\mathbb{Z}_p$, $x^{1+k(p-1)} = x$ for $k = 0, 1, 2, \dots$. Thus:

The pairs $(p, 1 + k(p-1))$ have property P, and these are all the pairs with property P when $m = p$. $(*)$

The pair (p, p) is obtained when $k = 1$.

Now suppose that some $x \in \mathbb{Z}_m$ has the property that $x^r = 0$ for some $r > 1$, and that $x^n = x$. Clearly $x^t = x$ for all $t = n^2, n^3, n^4, \dots$. Thus we can find $t > r$ such that $x^t = x$. But then $x^r = 0$ implies that $x^t = 0$. Therefore, if (m, n) has property P, $\mathbb{Z}_m$ cannot have a nonzero element x with one of its powers equal to 0. For example, $2^2 = 0$ in $\mathbb{Z}_4$ and $6^3 = 0$ in $\mathbb{Z}_{72}$. So neither $m = 4$ nor $m = 72$ can occur in a pair with property P.

More generally, suppose $m = p_1^{a_1} p_2^{a_2} \cdots p_i^{a_i}$ with $p_1 < p_2 < \cdots < p_i$ and at least one $a_j > 1$. Then the element $q = p_1 p_2 \cdots p_i$ in $\mathbb{Z}_m$ is nonzero, but $q^A = 0$ where $A = \max\{a_1, a_2, \dots, a_i\}$. Therefore:

If (m, n) has property P, then m is square-free.

Now consider a square-free integer $m = p_1 p_2 \cdots p_i$, with $p_1 < p_2 < \cdots < p_i$. In this case, $\mathbb{Z}_m$ is naturally isomorphic to the ring-theoretic direct sum $R = \mathbb{Z}_{p_1} \oplus \mathbb{Z}_{p_2} \oplus \cdots \oplus \mathbb{Z}_{p_i}$ (see, e.g., **[1]**). Clearly if $x = (x_1, x_2, \dots, x_i) \in R$, then $x^n = x$ if and only if $x_j^n = x_j$ in $\mathbb{Z}_{p_j}$ for all $j = 1, 2, \dots, i$. We know from $(*)$ that (p_j, n) has property P if and only if n is an element of the sequence $P_j = \{1 + k(p_j - 1)\}_{j=0}^{\infty}$. Let ℓ be the least common multiple (lcm) of the set $\{p_1 - 1, p_2 - 1, \dots, p_i - 1\}$, and let L be the sequence $\{1+k\ell\}_{k=0}^{\infty}$. It is a straightforward set-theoretic exercise to show that $\bigcap_{j=1}^{i} P_j = L$.

MAIN RESULT. *The pair (m, n) has property P if and only if m is square-free and n is an element of the sequence L defined above.*

Notice that when $m = p$, L gives the sequence of pairs $(p, 1+k(p-1))$ from $(*)$.

Example. Consider $m = 385 = 5 \cdot 7 \cdot 11$. Then $\ell = \mathrm{lcm}(4, 6, 10) = 60$. So $L = \{1, 61, 121, 181, \dots\}$. For instance, $(385, 121)$ has property P.

Remarks

1. If $n \le m$ is desired, we can let the sequence L run from $k = 0$ to $\left\lfloor \frac{m-1}{\ell} \right\rfloor$, the integer floor of $\frac{m-1}{\ell}$. In the example above, with $m = 385$, the largest value of n $(\le m)$ is

$$1 + \left\lfloor \frac{384}{60} \right\rfloor 60 = 1 + 6(60) = 361.$$

2. The problem can be solved more generally in the class of finite rings. Suppose R is a finite ring having the property that for some fixed $n > 1$, $x^n = x$ for all x in R. Then R must be a direct sum of finite fields, each of which has prime power order. An L-sequence similar to the one above can then be defined.
3. Other generalizations are possible. Suppose we want to find all *triples* (m, n, k) that have the property P': For $n > k > 0$, $x^n = x^k$ for all x in the ring $\mathbb{Z}_m$. (The case $k = 1$ is discussed above.) Because $\mathbb{Z}_m$ is decomposable into a ring-theoretic direct sum of rings with prime-power order, it suffices to consider triples for $\mathbb{Z}_{p^t}$.
   The multiplicative group of units of $\mathbb{Z}_{p^t}$ has order $q = p^{t-1}(p-1)$ (and in fact is cyclic for odd p), so $x^q = 1$ for every unit x. If x is not a unit, then p is a factor of x; therefore $x^t = 0$. Thus $x^{q+t} = x^t$ for all x in $\mathbb{Z}_{p^t}$, i.e. $(p^t, q+t, t)$ has property P'. The *pair* (p, p) is obtained when $t = 1$.
   For $R = \mathbb{Z}_{p_1^{t_1}} \oplus \mathbb{Z}_{p_2^{t_2}} \oplus \cdots \oplus \mathbb{Z}_{p_i^{t_i}}$, let $q_j = p_j^{t_j - 1}(p_j - 1)$, $\ell = \mathrm{lcm}(q_1, q_2, \dots, q_i)$, and $T = \max\{t_1, t_2, \dots, t_i\}$. Then $x^{k\ell + T} = x^T$ for all x in R and $k = 0, 1, 2, \dots$. For example, let $m = 3^2 \cdot 5^3 = 1125$, $R = \mathbb{Z}_{3^2} \oplus \mathbb{Z}_{5^3}$, $q_1 = 6$, $q_2 = 100$, $\ell = 300$, and $T = 3$. Then $x^{300k+3} = x^3$ for all x in $\mathbb{Z}_m$, $k = 0, 1, 2, \dots$. In particular, $(1125, 303, 3)$ is a triple with property P'.

Acknowledgment. The authors wish to thank the referee for helpful suggestions.

REFERENCES

1. John B. Fraleigh, *A First Course in Abstract Algebra*, 3rd edition, Addison-Wesley, Reading, MA, 1982.
2. I. Niven and H. S. Zuckerman, *An Introduction to the Theory of Numbers*, 4th edition, John Wiley and Sons, New York, NY, 1980.

Perfect Cuboids and Perfect Square Triangles

FLORIAN LUCA
Fakultät Mathematik
Universität Bielefeld
Postfach 10 01 31
33 501 Bielefeld
Germany

Among the many well-known unsolved diophantine problems is the following:

THE PERFECT CUBOID PROBLEM (PCP): *Is there a rectangular box with all edges, face diagonals, and main diagonals integers?*

An extensive list of references on this problem appears in **[1]**. In this note, we show that the existence of a solution for the PCP is equivalent to the existence of a solution for an apparently different problem:

THE PERFECT SQUARE TRIANGLE PROBLEM (PSTP): *Is there a triangle whose sides are perfect squares and whose angle bisectors are integers?*

Let us first observe that the word "integers" can be replaced by the word "rationals" in the statement of the PSTP. In order to show that the existence of a solution to the PCP is indeed equivalent to the existence of a solution to the PSTP, assume first that the PCP has a solution. Let x, y, and z be the edges of a perfect cuboid and set

$$a = y^2 + z^2, \quad b = x^2 + z^2, \quad c = x^2 + y^2. \tag{1}$$

Clearly, a, b, and c are the sides of a triangle and are perfect squares. Let $p = \frac{a+b+c}{2}$ be the semiperimeter of this triangle. Since

$$p = x^2 + y^2 + z^2, \quad p - a = x^2, \quad p - b = y^2, \quad p - c = z^2, \tag{2}$$

we conclude that all four numbers p, $p-a$, $p-b$, and $p-c$ are perfect squares. Let l_a, l_b, and l_c be the lengths of the angle bisectors drawn from the angles opposite to the sides a, b, and c, respectively. It is well known that the lengths of these angle bisectors are given in terms of a, b, and c by

$$l_a = 2 \cdot \frac{\sqrt{bcp(p-a)}}{b+c}, \quad l_b = 2 \cdot \frac{\sqrt{acp(p-b)}}{a+c}, \quad l_c = 2 \cdot \frac{\sqrt{abp(p-c)}}{a+b}. \tag{3}$$

Since all the numbers listed in (1) and (2) are perfect squares, it follows, by formula (3), that the triangle with sides a, b, and c is a solution of the PSTP (once "integers" has been replaced by "rationals" in the statement of the problem).

Conversely, assume now that the PSTP has a solution. Let a, b, and c be the sides of a triangle that solves this problem. We may assume that $\gcd(a, b, c) = 1$. Indeed, otherwise, let $d = \gcd(a, b, c)$. Since a, b, and c are all perfect squares, so is d. Then the triangle with sides a/d, b/d, and c/d still solves the PSTP, and $\gcd(a/d, b/d, c/d) = 1$.

By formula (3) and the fact that a, b, and c are perfect squares, it follows that all three integers

$$4p(p-a)=(b+c)^2-a^2,\quad 4p(p-b)=(a+c)^2-b^2,$$
$$4p(p-c)=(a+b)^2-c^2 \tag{4}$$

are perfect squares. Since $\gcd(a,b,c)=1$, it follows that not all of a, b, and c can be even. Reducing modulo 4 the integers listed in (4), one concludes that exactly one of the three numbers a, b, and c is even, and the other two are odd. It now follows that p is an integer, and formula (4) implies that all three integers

$$p(p-a),\quad p(p-b),\quad p(p-c) \tag{5}$$

are perfect squares. We now show that $\gcd(p-a,p-b,p-c)=1$. Indeed, let $e=\gcd(p-a,p-b,p-c)$. Clearly, $e \| (p-b)+(p-c)=a$. By a similar argument, one concludes that $e \mid b$ and $e \mid c$. Since $\gcd(a,b,c)=1$, it follows that $e=1$. Since all three numbers listed in (5) are perfect squares, so is their greatest common divisor. Hence,

$$\gcd(p(p-a),p(p-b),p(p-c))=pe=p$$

is a perfect square. It now follows (again from the fact that the three numbers in (5) are perfect squares) that all four numbers p, $p-a$, $p-b$, $p-c$ are perfect squares. If we now set

$$x=\sqrt{p-a}\,,\quad y=\sqrt{p-b}\,,\quad z=\sqrt{p-c}\,,$$

then one concludes easily that x, y, and z are the edges of a perfect cuboid.

REFERENCE

1. R. K. Guy, *Unsolved Problems in Number Theory*, Springer-Verlag, New York, NY, 1994, 173–181.

PROBLEMS

GEORGE T. GILBERT, *Editor*
Texas Christian University

ZE-LI DOU, KEN RICHARDSON, and SUSAN G. STAPLES, *Assistant Editors*
Texas Christian University

Proposals

To be considered for publication, solutions should be received by May 1, 2001.

1608. *Proposed by William D. Weakley, Indiana-Purdue University at Fort Wayne, Fort Wayne, Indiana.*

Let b be a positive integer, $b > 1$. We call a positive integer "onederful" in the base b if it divides some integer whose base b representation is all ones. Which positive integers are onederful in the base b?

1609. *Proposed by Yanir A. Rubinstein, student, Technion—Israel Institute of Technology, Haifa, Israel.*

Evaluate

$$\inf_{\substack{a,b\in\mathbb{C}\\ \operatorname{Im}(a\bar{b})\neq 0}} \frac{(|a|+|b|)(|a|+|b|+|a+b|)}{|\operatorname{Im}(a\bar{b})|}.$$

1610. *Proposed by Hassan A. Shah Ali, Tehran, Iran.*

Place n black pieces and n white pieces on distinct points on the circumference of a circle.

(a) Prove that for each natural number $k \leq n$, there exists a chain of $2k$ consecutive pieces on the circle of which exactly k are black.

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 containing the solver's name and full address.*

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 johnston@math.iastate.edu. *Readers who use e-mail should also provide an e-mail address.*

(b) Prove that there are at least two such chains that are disjoint if

$$k \leq \sqrt{2n+2} - 2.$$

1611. *Proposed by Ho-joo Lee, student, Kwangwoon University, Seoul, South Korea.*

Let P be in the interior of $\triangle ABC$, and let lines AP, BP, CP intersect the sides BC, CA, AB in L, M, N, respectively. Prove that P is the centroid of $\triangle ABC$ if

$$[APN] = [BPL] = [CPM],$$

where $[\cdot]$ denotes area.

1612. *Proposed by Ho-joo Lee, student, Kwangwoon University, Seoul, South Korea.*

Let P be in the interior of $\triangle\ ABC$, and let lines AP, BP, CP intersect the sides BC, CA, AB in L, M, N, respectively. Prove that P is the centroid of $\triangle ABC$ if

$$[APN] + [BPL] + [CPM] = [APM] + [BPN] + [CPL],$$

where $[\cdot]$ denotes area.

Quickies

Answers to the Quickies are on pages 410.

Q905. *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada.*

Determine the maximum volume of a tetrahedron given the lengths of three of its medians.

Q906. *Proposed by Răzvan Tudoran, University of Timişoara, Timişoara, Romania.*

Let n and k be positive integers with $k < n$. Prove the inequality

$$\binom{n}{k} \leq \left(1 + \frac{k}{n-k}(1 + \ln(n-k))\right)^{n-k}.$$

Solutions

A Triangular Number of Triangles — December 1999

1584. *Proposed by Ira Rosenholtz, Eastern Illinois University, Charleston, Illinois.*

Let n be a positive integer, and let $\triangle_n$ be the set of ordered triples of positive integers which are the side lengths of a nondegenerate triangle of perimeter n. Show that the cardinality of $\triangle_n$ is a triangular number.

I. *Solution by Michael Vowe, Therwil, Switzerland.*

Represent an element of Δ_n by $(a, b, n-a-b) \in \mathbb{N}^3$. The triangle inequality implies the conditions $a < n/2$, $b < n/2$, and $a + b > n/2$. Thus, the cardinality of Δ_n is the number of lattice points in the interior of the isosceles right triangle bounded by the lines $a = n/2$, $b = n/2$, and $a + b = n/2$. This number is clearly a triangular number. In particular, if n is odd, then

$$|\Delta_n| = 1 + 2 + \cdots + \frac{n-1}{2} = \frac{1}{2}\frac{n-1}{2}\frac{n+1}{2}.$$

If n is even, then

$$|\Delta_n| = 1 + 2 + \cdots + \frac{n-4}{2} = \frac{1}{2}\frac{n-4}{2}\frac{n-2}{2}.$$

II. *Solution by José H. Nieto, Universidad del Zulia, Maracaibo, Venezuela.*

We will show that

$$|\Delta_n| = \begin{cases} \frac{1}{2}\left(\frac{n}{2} - 2\right)\left(\frac{n}{2} - 1\right) & \text{if } n \text{ is even,} \\ \frac{1}{2}\frac{n-1}{2}\frac{n+1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

First we will establish the recurrence relation

$$|\Delta_n| - |\Delta_{n-3}| = \begin{cases} 0 & \text{if } n \text{ is even,} \\ \frac{3}{2}(n-3) & \text{if } n \text{ is odd.} \end{cases}$$

To prove this, note that if $(a, b, c) \in \Delta_n$ then $(a-1, b-1, c-1)$ either belongs to Δ_{n-3} or represents the side lengths of a *degenerate* triangle of perimeter $n - 3$. Since the perimeter of a degenerate triangle must be even, if n is even we conclude that $|\Delta_n| = |\Delta_{n-3}|$. If n is odd, let $m = (n-3)/2$. If m is even, the possible degenerate triangles of perimeter $2m$ have side lengths (without regarding the order) $(m, m, 0), (m, m-1, 1), \dots, (m, m/2+1, m/2-1), (m, m/2, m/2)$. Taking the order into account we observe that each of these triples may be permuted in 6 ways, except the first and the last, which may be permuted in only 3 ways. Therefore we obtain $6(m/2 - 1) + 3 \cdot 2 = 3m = 3(n-3)/2$ ordered degenerate triples. If m is odd, the possible degenerate triangles of perimeter $2m$ have side lengths (without regarding the order) $(m, m, 0), (m, m-1, 1), \dots, (m, (m+1)/2, (m-1)/2)$. Taking the order into account we observe that each of these triples may be permuted in 6 ways, except the first one, which may be permuted in only 3 ways. Summing up, we obtain again $6(m-1)/2 + 3 = 3m = 3(n-3)/2$ ordered degenerate triples.

Now it is straightforward to prove our claim by induction. It is true for $n = 1$, $n = 2$, and $n = 3$. Let $n > 3$ and assume the result valid for triangles with perimeter less than n.

If n is even, then $n - 3$ is odd and by the recurrence relation and the induction hypothesis we have

$$|\Delta_n| = |\Delta_{n-3}| = \frac{1}{2}\frac{n-4}{2}\frac{n-2}{2} = \frac{1}{2}\left(\frac{n}{2} - 2\right)\left(\frac{n}{2} - 1\right).$$

On the other hand, if n is odd we have

$$|\Delta_n| = |\Delta_{n-3}| + \frac{3}{2}(n-3) = \frac{1}{2}\left(\frac{n-3}{2}-2\right)\left(\frac{n-3}{2}-1\right) + \frac{3}{2}(n-3)$$
$$= \frac{1}{2}\frac{n-1}{2}\frac{n+1}{2}.$$

III. *Solution by Jim Delany, California Polytechnic State University, San Luis Obispo, California.*

We claim that $|\Delta_{2m}| = (m-2)(m-1)/2$ and $|\Delta_{2m+1}| = m(m+1)/2$.

Let D_n be the set of positive integers (a, b, c) for which $a+b+c=n$ and $a \equiv b \equiv c \pmod 2$. In light of the triangle inequality the mapping from Δ_n to D_n given by

$$(\alpha, \beta, \gamma) \to (n-2\alpha, n-2\beta, n-2\gamma)$$

is a bijection, so $|\Delta_n| = |D_n|$. We compute the latter, even perimeters first.

If $(a, b, c) \in D_{2m}$, then a, b, c are all even. Thus, $(a, b, c) = (2i, 2j, 2k)$, where i, j, k are positive integers such that $i+j+k=m$. The number of such triples is the coefficient of x^m in the generating function

$$(x + x^2 + x^3 + \cdots)^3 = \frac{x^3}{(1-x)^3} = \sum_{m=3}^{\infty} \frac{(m-1)(m-2)}{2} x^m.$$

Therefore, $|D_{2m}| = (m-2)(m-1)/2$. For odd perimeters, we use the bijection from D_{2m+1} to D_{2m+4} given by

$$(a, b, c) \to (a+1, b+1, c+1)$$

to conclude $|D_{2m+1}| = |D_{2m+4}| = m(m+1)/2$.

Also solved by Michael H. Andreoli, Michel Bataille (France), J. C. Binz (Switzerland), Jean Bogaert (Belgium), Keith Chavey, Michael P. Cohen, Leo Comerford, Con Amore Problem Group (Denmark), Daniele Donini (Italy), Kurt Dresner (student), Arthur H. Foss, Marty Getz and Dixon Jones, Georgi D. Gospodinov (student), Robert Heller, Kathleen E. Lewis, Kevin McDougal, Peter Schumer, Heinz-Jürgen Seiffert (Germany), Skidmore College Problem Group, Southwest Missouri State University Problem Solving Group, Philip D. Straffin, Monte J. Zerger, Li Zhou, and the proposer. There was one incorrect solution.

A Timely Sum, but not a Square **December 1999**

1585. *Proposed by Shahin Amrahov, Ankara, Turkey.*

Prove that the number

$$\sum_{n=1}^{1998}\left(\sum_{k=1}^{n} k^{1998}\right)1997^{n-1}$$

is not a perfect square.

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

More generally, we prove that

$$N := \sum_{n=1}^{p-1}\left(\sum_{k=1}^{n} k^{p-1}\right)(p-2)^{n-1}$$

is not a perfect square for any prime p congruent to 5 or 7 modulo 12.

For p prime and $1 \le k < p$, we have $k^{p-1} \equiv 1 \pmod{p}$. Hence

$$N \equiv \sum_{n=1}^{p-1} n(-2)^{n-1} \pmod{p}.$$

Differentiating the identity $\sum_{n=0}^{p-1} x^n = (1-x^p)/(1-x)$ yields

$$\sum_{n=1}^{p-1} nx^{n-1} = \frac{1 - px^{p-1} + (p-1)x^p}{(1-x)^2}.$$

Thus, letting $x = -2$ and clearing the denominator,

$$9N \equiv 1 - p(-2)^{p-1} + (p-1)(-2)^p \equiv 3 \pmod{p},$$

or $3N \equiv 1 \pmod{p}$.

We conclude that 3 and N are either both quadratic residues or both quadratic nonresidues modulo p. In terms of the Legendre symbol, the law of quadratic reciprocity states that, for odd primes p and q, $\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$ unless both primes are congruent to 3 modulo 4, in which case $\left(\frac{q}{p}\right) = -\left(\frac{p}{q}\right)$. Hence, if $p \equiv 5 \pmod{12}$, we have

$$\left(\frac{3}{p}\right) = \left(\frac{p}{3}\right) = \left(\frac{2}{3}\right) = -1.$$

If $p \equiv 7 \pmod{12}$, we have

$$\left(\frac{3}{p}\right) = -\left(\frac{p}{3}\right) = -\left(\frac{1}{3}\right) = -1.$$

In both cases 3 is a quadratic nonresidue modulo p, completing the proof.

II. *Solution by J. C. Binz, University of Bern, Bern, Switzerland.*

Set

$$N := \sum_{n=1}^{1998} \left(\sum_{k=1}^{n} k^{1998} \right) 1997^{n-1} = \sum_{k=1}^{1998} \left(\sum_{n=k}^{1998} 1997^{n-1} \right) k^{1998}.$$

We observe that $1997 \equiv 2 \pmod{7}$, $k^{1998} \equiv 1 \pmod{7}$ if $(k,7) = 1$, and $k^{1998} \equiv 0 \pmod{7}$ if $7 \mid k$. It follows that

$$\begin{aligned}
N &\equiv \sum_{\substack{k=1\\(k,7)=1}}^{1998} \left(\sum_{n=k}^{1998} 2^{n-1} \right) \equiv \sum_{\substack{k=1\\(k,7)=1}}^{1998} (1 - 2^{k-1}) \\
&\equiv \sum_{k=1}^{1998} (1 - 2^{k-1}) - \sum_{k=1}^{\lfloor 1998/7 \rfloor} (1 - 2^{7k-1}) \\
&\equiv \sum_{k=1}^{1998} (1 - 2^{k-1}) - \sum_{k=1}^{285} (1 - 2^{k-1}) \equiv \sum_{k=286}^{1998} (1 - 2^{k-1}) \\
&\equiv (1998 - 285) - (2^{1998} - 2^{285}) \equiv 5 - 0 \equiv 5 \pmod{7}.
\end{aligned}$$

Because 5 is not a square modulo 7, N cannot be a perfect square.

Comment. Georgi Gospodinov proved the assertion in a computation modulo 16 similar to the two above.

Also solved by Michel Bataille (France), Brian D. Beasley, David Clarke, Con Amore Problem Group (Denmark), Daniele Donini (Italy), Georgi D. Gospodinov (student), Heinz-Jürgen Seiffert (Germany), and the proposer.

A Comparison of Two Integrals **December 1999**

1586. *Proposed by Gerald A. Edgar, Ohio State University, Columbus, Ohio.*

Let w be a nonnegative, continuous, and nonincreasing function on $[0, \infty)$. Let g be a nonnegative, continuous function on $[0, \infty)$. For a given $\alpha \in (0, 1)$, assume that

$$\alpha x g(x) \leq \int_0^x \min\{w(t), g(x)\}\, dt \quad \text{for all } x > 0.$$

(a) Show that there is a positive constant c_α, independent of w and g, such that

$$\int_0^\infty g(x)\, dx \leq c_\alpha \int_0^\infty w(x)\, dx.$$

(b)* Find the smallest possible value of c_α.

(* Neither the proposer nor the editors have provided a solution to (b). Solvers of only (a) will be acknowledged.)

Composite of solutions due to the Proposer and the Editors.

(a) For $\kappa \in [0, 1)$, we have

$$\begin{aligned}\alpha x g(x) &\leq \int_0^x \min\{w(t), g(x)\}\, dt \leq \int_0^{\kappa\alpha x} g(x)\, dt + \int_{\kappa\alpha x}^x w(t)\, dt \\ &\leq \kappa\alpha x g(x) + (1 - \kappa\alpha) x w(\kappa\alpha x),\end{aligned}$$

so that

$$g(x) \leq \frac{1 - \kappa\alpha}{(1 - \kappa)\alpha} w(\kappa\alpha x)$$

for $x > 0$. Therefore,

$$\int_0^\infty g(x) dx \leq \frac{1 - \kappa\alpha}{(1 - \kappa)\alpha} \int_0^\infty w(\kappa\alpha x)\, dx = \frac{1 - \kappa\alpha}{(\kappa - \kappa^2)\alpha^2} \int_0^\infty w(x)\, dx.$$

We find that $\kappa = (1 - \sqrt{1 - \alpha})/\alpha < 1$ minimizes the fraction, hence

$$\int_0^\infty g(x)\, dx \leq \frac{1}{\left(1 - \sqrt{1 - \alpha}\right)^2} \int_0^\infty w(x)\, dx.$$

Although we cannot answer part (b), we obtain a lower bound for c_α. Choose any continuous, positive, nonincreasing, and integrable $w(x)$ and set $g(x) := w(\alpha x)$. Then

$$\begin{aligned}\int_0^x \min\{w(t), g(x)\}\, dt &= \int_0^{\alpha x} g(x)\, dt + \int_{\alpha x}^x w(t)\, dt \\ &= \alpha x g(x) + \int_{\alpha x}^x w(t)\, dt \geq \alpha x g(x).\end{aligned}$$

Furthermore,

$$\int_0^\infty g(x)\,dx = \int_0^\infty w(\alpha x)\,dx = \frac{1}{\alpha}\int_0^\infty w(x)\,dx,$$

so $c_\alpha \geq 1/\alpha$. For comparison,

$$\frac{1}{\left(1-\sqrt{1-\alpha}\right)^2} \sim \frac{4}{\alpha^2}$$

as $\alpha \to 0$.

Constructing Foci of Conics **December 1999**

1587. *Proposed by Kevin Ferland, Bloomsburg University, Bloomsburg, Pennsylvania, and Florian Luca, Czech Academy of Science, Prague, Czech Republic.*

Consider constructions using straightedge and compass. Prove or disprove the following:

(a) Given any ellipse, the foci can be constructed.
(b) Given any hyperbola, the foci and asymptotes can be constructed.
(c) Given any parabola, the focus and directrix can be constructed.

Solution by David M. Bloom, Brooklyn College of CUNY, Brooklyn, New York.

All of the constructions are possible.

(c) The construction of the focus of a given parabola was a problem on the 1955 Putnam examination. One solution follows. We may assume the parabola has equation $4cy = x^2$. If a line $y = mx + b$ cuts the parabola at $A_i = (x_i, y_i)$ $(i = 1,2)$, then $x_i = 2cm \pm 2\sqrt{c^2m^2 + cb}$ and hence the midpoint of A_1A_2 has x-coordinate $2cm$, which is independent of b. It follows that if we draw two parallel chords of the parabola, then the line L' joining their midpoints is parallel to the axis L of the parabola. We construct L as the perpendicular bisector of any chord of the parabola perpendicular to L'. Then the vertex V is the intersection of L and the parabola, and the x-axis is the perpendicular to L at V. If we now draw the line of slope $1/2$ through V, intersecting the parabola a second time in W, then the line through W perpendicular to L will meet L at the focus F.

To construct the directrix, locate point Q on L so that F and Q are equidistant from V and then draw the perpendicular to L through Q.

(a) The construction for the ellipse is based on a similar geometric fact: The line joining the midpoints of two parallel chords will pass through the center of the ellipse. To show this, assume the ellipse has equation $x^2/a^2 + y^2/b^2 = 1$. Then the midpoint of the chord satisfying $y = mx + r$ is

$$(x_0, y_0) = \left(\frac{-a^2mr}{b^2 + a^2m^2}, \frac{b^2r}{b^2 + a^2m^2}\right),$$

so that the ratio $y_0/x_0 = -b^2/(a^2m)$ depends only on m and not on r. Thus, the line through the center C having slope $-b^2/a^2m$ will pass through the midpoints of all chords of slope m, and our assertion follows. We construct the line between the midpoints of any two parallel chords of the ellipse; the center C is the midpoint of the chord, call it AB, determined by this line. The circle through A and B with center C intersects the ellipse in two more points, A' and B' (unless A and B are fortuitously vertices of the ellipse). The perpendicular bisectors of AA' and AB' are then the axes of the ellipse, which in turn give us the major vertices V_1 and V_2 and the minor vertices W_1 and W_2. Finally, the foci are the intersections of the major axis with the circle having center W_1 and radius CV_1.

(b) For the hyperbola, which we may assume has equation $x^2/a^2 - y^2/b^2 = 1$, the center C, major axis L, and vertices V_i are found the same way as for the ellipse. We then have $CV_i = a$ and the distance from C to the foci is $c = (a^2 + b^2)^{1/2}$, so we can construct the foci once we can construct b. To construct b, locate P on L so that $PC = 2a$; then construct the perpendicular to L at P, with Q one of its intersections with the hyperbola, so that $PQ = b\sqrt{3}$. Construct R on L such that $\angle PQR = 30°$. Then $PR = b$ and we may construct the foci. Because the slopes of the asymptotes are $\pm b/a$, the asymptotes are now constructible as well.

Also solved by Michel Bataille (France), Con Amore Problem Group (Denmark), Neela Lakshmanan, Peter Y. Woo, and the proposers. There was one incorrect solution.

An Iterative Sequence of Determinants **December 1999**

1588. *Proposed by Emeric Deutsch, Polytechnic University, Brooklyn, New York.*

Let $a = (a_0, a_1, a_2, \dots)$ be any sequence of complex numbers. Define the sequence transformation T by $T(a) = (b_0, b_1, b_2, \dots)$, where

$$b_n = \begin{vmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} & a_n \\ -1 & a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ 0 & -1 & a_0 & \cdots & a_{n-3} & a_{n-2} \\ & & & \cdots & & \\ 0 & 0 & 0 & \cdots & a_0 & a_1 \\ 0 & 0 & 0 & \cdots & -1 & a_0 \end{vmatrix}.$$

Find a determinant expression for the nth term of the sequence $T^{(q)}(a)$, where q is a positive integer. (Here $T^{(q)}$ denotes the q-fold composition of T.)

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

We will show that if $T^{(q)}(a) = (c_0, c_1, c_2, \dots)$, then

$$c_n = \begin{vmatrix} qa_0 & qa_1 & qa_2 & \cdots & qa_{n-1} & a_n \\ -1 & qa_0 & qa_1 & \cdots & qa_{n-2} & a_{n-1} \\ 0 & -1 & qa_0 & \cdots & qa_{n-3} & a_{n-2} \\ & & & \cdots & & \\ 0 & 0 & 0 & \cdots & qa_0 & a_1 \\ 0 & 0 & 0 & \cdots & -1 & a_0 \end{vmatrix}.$$

Clearly, $b_0 = a_0$. We claim that

$$b_n = a_n + \sum_{k=0}^{n-1} b_k a_{n-1-k}.$$

By expanding b_n in its last column, we obtain the claim by induction. Set

$$a(x) = \sum_{n=0}^{\infty} a_n x^n, \quad b(x) = \sum_{n=0}^{\infty} b_n x^n, \quad c(x) = \sum_{n=0}^{\infty} c_n x^n.$$

Then $b(x)=a(x)+xb(x)a(x)$, hence

$$b(x)=\frac{a(x)}{1-xa(x)}.$$

By induction on q, it follows that

$$c(x)=\frac{a(x)}{1-qxa(x)},$$

i.e., $c(x)=a(x)+qxc(x)a(x)$. Therefore,

$$c_n=a_n+\sum_{k=0}^{n-1}c_k(qa_{n-1-k}),$$

which can be written in the determinant form as claimed.

Also solved by J. C. Binz (Switzerland), David Callan, and the proposer.

Answers

Solutions to the Quickies on page 403.

A905. Let $ABCD$ be the tetrahedron, m_A, m_B, m_C, and m_D the median lengths, the last three given, and G the centroid. The medians are concurrent and are such that $AG=3m_A/4$, $BG=3m_B/4$, $CG=3m_C/4$, and $DG=3m_{AD}/4$. Thus the volume of $ABCD$ is four times the volume of $GBCD$. Furthermore, the latter volume will be a maximum when BG, CG, and DG are mutually orthogonal. Hence the maximal volume is $4(BG\cdot CG\cdot DG/6)=9m_Bm_Cm_D/32$.

This easily generalizes to determining the maximum volume of an n-dimensional simplex given the lengths of n of its medians.

A906. From the arithmetic mean-geometric mean inequality, we have

$$\begin{aligned}\binom{n}{k}&=\frac{(k+1)(k+2)\cdots(k+n-k)}{(n-k)!}\\&=(1+k)\left(1+\frac{k}{2}\right)\cdots\left(1+\frac{k}{n-k}\right)\\&\le\left(\frac{(1+k)+\left(1+\frac{k}{2}\right)+\cdots+\left(1+\frac{k}{n-k}\right)}{n-k}\right)^{n-k}\\&=\left(1+\frac{k}{n-k}\left(1+\frac{1}{2}+\cdots+\frac{1}{n-k}\right)\right)^{n-k}.\end{aligned}$$

The inequality now follows from

$$1+\frac{1}{2}+\cdots+\frac{1}{n-k}\le1+\int_1^{n-k}\frac{1}{x}\,dx=1+\ln(n-k).$$

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.

Devlin, Keith, *The Math Gene: How Mathematical Thinking Evolved and Why Numbers are Like Gossip*, Basic Books, 2000; xvii + 328 pp, $25. ISBN 0–465–01618–9.

Most of the U.S. believes that there is a "math gene"; if you don't have it, you can't do math and shouldn't have to try (the same goes for a "science gene"). Most of the rest of the world is convinced that people who try hard can do math. Despite the overwhelming explanatory success of this alternative (see any international comparison of math achievement), belief in a "math gene" persists. Unfortunately, the title of Devlin's book will tend to reinforce that belief, despite his disclaimer that there is no DNA sequence for mathematical ability and his main thesis that we all have "the math gene." He means an "innate facility for mathematical thought," and he details an argument that language and mathematics both proceeded from the same developments in the brain. "Thinking mathematically is just a specialized form of using our language facility." The book is written in a conversational style, rich with metaphors and examples, and it does a good job of explaining the nature of mathematics as the science of patterns. Devlin characterizes mathematics as just one kind of "off-line thinking" (i.e., abstraction) about a "world of internally generated symbols." But if we all have "the math gene," why do so many claim to find math impossible? Because it takes practice and hard work (concentration), and because of the "degree of rigor required in its reasoning processes." Lack of interest is the main difference between those who can do math and those who say they can't, says Devlin. Why should we encourage that interest? Not because you need math to function in a technological society, but because studying it develops scientific habits of mind and the ability to learn and adapt to changing circumstances. Devlin also cites the Riley Report ("Mathematics and Future Opportunities," `http://www.ed.gov/pubs/math/part3.html`) for concrete "benefits" from taking more mathematics: a greater chance of going to college and succeeding there, especially for children from low-income families. But such claims are like those that learning a musical instrument will make a child a better student—association does not imply causation.

Hass, Joel, General double bubble conjecture in $\mathbb{R}^3$ solved, *Focus* 20 (5) (May/June 2000) 4–5. Stewart, Ian, Bubble trouble: It's taken 170 years, but now the problem is solved, *New Scientist* (25 March 2000) 6. Cipra, Barry, Dana Mackenzie, and Charles Seife, Rounding out solutions to three conjectures, *Science* (17 March 2000) 1910–1912.

Michael Hutchins (Stanford University), Frank Williams (Williams College), Manuel Ritoré and Antonio Ros (University of Granada) have announced a proof of the general double bubble conjecture in $\mathbb{R}^3$. The conjecture says that the surface enclosing two given volumes that has smallest area consists of two bubbles connected by a common surface. The proof proceeds by showing that other candidates for optimality could be deformed into other feasible solutions with smaller surface area, i.e., cannot be even locally optimal. Undergraduate students of Morgan's had already proved the double bubble conjecture in two- and in four-dimensional space.

Peterson, I., The power of partitions: Writing a whole number as the sum of smaller numbers springs a mathematical surprise, *Science News* 157 (17 June 2000) 396. **http://www.sciencenews.org/20000617/bob2.asp** . Ono, Ken, Distribution of the partition function modulo m, *Annals of Mathematics* 151 (January 2000) 293–307 (MR 2000k: 1115). The genius factor, *The Penn Stater* (July/August 2000) 28.

The number of partitions $p(n)$ of a positive integer n is the number of ways to express it as a sum of (unordered) positive integers. Srinivasa Ramanujan proved that starting at a certain integer, every fifth integer thereafter has its number of partitions divisible by 5, viz., $p(5n+4) \equiv 0 \bmod 5$. The same holds for every seventh (divisible by 7), every eleventh (by 11), and similarly for multiples or powers of 5, 7, and 11. In the 80 years since, only one or two other isolated patterns of the kind $p(kn+l) \equiv 0 \bmod m$ (called *congruences*) have been found. Are they just flukes? No; we now know that there is a congruence for every prime (except possibly 2 and 3), as proved by Ken Ono (Penn State University and University of Wisconsin), and for every composite made from those primes, as proved by Scott Ahlgren (Colgate University). Ono's proof was not constructive; but an undergraduate at Penn State, Rhiannon L. Weaver, found an algorithm for generating examples.

Mackenzie, Dana, May the best man lose, *Discover* 21 (11) (November 2000). **http://www.discover.com/nov_00/featbestman.html** . **http://www.math.nwu.edu/~d_saari/** . Saari, Donald G., and Maria M. Tataru, The likelihood of dubious election outcomes, *Economic Theory* 13 (2) (1999) 345–363. Saari, Donald G., Mathematical structure of voting paradoxes, I: Pairwise votes, *Economic Theory* 15 (1) (2000) 1–53; II: Positional voting, 55–102. Saari, Donald G., and Fabrice Valognes, Geometry, voting, and paradoxes, this MAGAZINE 78 (October 1998) 243–259.

Is something wrong with the U.S. election system? Did the stunning closeness of the 2000 presidential election challenge your complacency, increase your dissatisfaction, or confirm your contentment with winning by plurality (no runoff) and winner-take-all (for electoral votes in almost all states)? Two voting theorists, mathematician Donald Saari (Northwestern University) and political scientist Steve Brams (New York University),find the presidential primary and election system fundamentally flawed. What should we do instead? Brams favors approval voting (used by the MAA in some elections) but Saari favors the Borda count (each voter ranks the candidates and the ranks are added over all voters).

Frucht, William (ed.), *Imaginary Numbers: An Anthology of Marvelous Mathematical Stories, Diversions, Poems, and Musings*, Wiley, 2000; xvi + 327 pp, $16.95 (P). ISBN 0–471–39341–X.

The last anthology of prose and poetry about mathematics was Rudy Rucker's *Mathenauts: Tales of Mathematical Wonder* (1987); but the spiritual predecessors of this book are Clifton Fadiman's *Fantasia Mathematica* (1958) and *The Mathematical Magpie* (1962) (both in print). Frucht has collected works (with three exceptions) that did not exist back then; among the 31 pieces by Rucker, Raymond Smullyan, Douglas Hofstadter, Martin Gardner, Jorge Luis Borges ("The library of Babel"), and others are six poems and two dialogues.

Fink, Thomas, and Yong Mao, *The 85 Ways to Tie a Tie: The Science and Aesthetics of Tie Knots*, Bantam Doubleday Dell, 2000; 144 pp, $14.95. ISBN 0–76790643–8. Tie knots, random walks, and topology, *Physica A* 276 (2000) 109–121; **http://www.tcm.phy.cam.ac.uk/~ym101/tie/aps97tie.html** .

Just in time for last-minute holiday shopping!: complete instructions on *every* way to tie a necktie, plus advice on good taste (which knot goes best with a particular tie and collar?).

NEWS AND LETTERS

Acknowledgments

Along with our associate editors, the following referees have assisted the MAGAZINE *during the past year. We thank them for their time and care.*

Adams, Colin, *Williams College, Williamstown, MA*

Adams, William W., *University of Maryland, College Park, MD*

Alavi, Yousef, *Western Michigan University, Kalamazoo, MI*

Andersen, Kurt Munk, *Technical University of Denmark, Copenhagen, Denmark*

Apostol, Tom, *California Institute of Technology, Pasadena, CA*

Appleby, Glenn D., *Santa Clara University, Santa Clara, CA*

Archibald, W. Thomas, *Acadia University, Wolfville, Canada*

Askey, Richard A., *University of Wisconsin, Madison, WI*

Axler, Sheldon, *San Francisco State University, San Francisco, CA*

Barbeau, Edward, *University of Toronto, Toronto, Canada*

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

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

Beals, Richard W., *Yale University, New Haven, CT*

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

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

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

Bergum, Gerald E., *South Dakota State University, Brookings, SD*

Berndt, Bruce, *University of Illinois, Urbana, IL*

Brawner, James N., *Armstrong Atlantic State University, Savannah, GA*

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

Bridger, Mark, *Northeastern University, Boston, MA*

Burger, Edward, *Williams College, Williamstown, MA*

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

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

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

Chinn, Phyllis Z., *Humboldt State University, Arcata, CA*

Crannell, Annalisa, *Franklin & Marshall College, Lancaster, PA*

Culp–Ressler, Wendell, *Franklin & Marshall College, Lancaster, PA*

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

Drobot, Vladimir, *San Jose State University, San Jose, CA*

Eisenberg, Bennett, *Lehigh University, Bethlehem, PA*

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

Emert, John, *Ball State University, Muncie, IN*

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

Firey, William, *Corvallis, OR*

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

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

Flanders, Harley, *Jacksonville Beach, FL*

Fraser, Craig G., *University of Toronto, Toronto, Canada*

Fredricks, Gregory A., *Lewis & Clark College, Portland, OR*

Gallian, Joseph, *University of Minnesota,* *Duluth, MN*

Gloor, Philip, *St. Olaf College, Northfield, MN*

Goldstein, Jerome A., *University of Memphis, Memphis, TN*

Goodson, Geoffrey R., *Towson University, Towson, MD*

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

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

Grant, John, *Towson University, Towson, MD*

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

Groeneveld, Richard, *Iowa State University, Ames, IA*

Grosshans, Frank D., *West Chester University of Pennsylvania, West Chester, PA*

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

Gulick, Denny, *University of Maryland, College Park, MD*

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

Haunsperger, Deanna, *Carleton College, Northfield, MN*

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

Henle, Michael, *Oberlin College, Oberlin, OH*

Hoehn, Larry P., *Austin Peay State University, Clarksville, TN*

Honsberger, Ross A., *University of Waterloo, Waterloo, Canada*

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

Isaac, Richard, *Herbert H. Lehman College (CUNY), Bronx, NY*

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

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

Johnson, Warren P., *Beloit College, Beloit, WI*

Johnsonbaugh, Richard, *DePaul University, Chicago, IL*

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

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

Kaplan, Sam, *University of North Carolina, Asheville, NC*

Kerckhove, Michael G., *University of Richmond, Richmond, VA*

Kimberling, Clark, *University of Evansville, Evansville, IN*

Kung, Sidney H.L., *University of North Florida, Jacksonville, FL*

Kuzmanovich, James, *Wake Forest University, Winston-Salem, NC*

Lagarias, Jeffrey, *AT&T Laboratories Research, Florham Park, NJ*

Larson, Loren, *St. Olaf College, Northfield, MN*

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

Lewand, Robert E., *Goucher College, Baltimore, MD*

LoBello, Anthony, *Allegheny College, Meadville, PA*

Mackiw, George B., *Loyola College, Baltimore, MD*

Manvel, Bennet, *Colorado State University, Ft. Collins, CO*

McCleary, John H., *Vassar College, Poughkeepsie, NY*

McDermot, Richard F., *Allegheny College, Meadville, PA*

Megginson, Robert E., *University of Michigan, Ann Arbor, MI*

Merrill, Kathy, *Colorado College, Colorado Springs, CO*

Morgan, Frank, *Williams College, Williamstown, MA*

Needham, Tristan, *University of San Francisco, San Francisco, CA*

Neidinger, Richard D., *Davidson College, Davidson, NC*

Nelsen, Roger B., *Lewis & Clark College, Portland, OR*

Nunemacher, Jeffrey, *Ohio Wesleyan University, Delaware, OH*

Pearce, Janice L., *Berea College, Berea, KY*

Pedersen, Jean, *Santa Clara University, Santa Clara, CA*

Pfiefer, Richard E., *San Jose State University, San Jose, CA*

Pomerance, Carl, *Lucent Technologies Bell Laboratories, Murray Hill, NJ*

Richman, Fred, *Florida Atlantic University, Boca Raton, FL*

Riddle, Lawrence H., *Agnes Scott College, Decatur, GA*

Rosenholtz, Ira, *Eastern Illinois University, Charleston, IL*

Rosenstein, Jr., George M., *Franklin & Marshall College, Lancaster, PA*

Ross, Kenneth, *University of Oregon, Eugene, OR*

Ross, Peter, *Santa Clara University, Santa Clara, CA*

Roth, Richard L., *University of Colorado, Boulder, CO*

Roy, Ranjan, *Beloit College, Beloit, WI*

Sadek, Jawad, *Northwest Missouri State University, Maryville, MO*

Samuels, Stephen M., *Purdue University, West Lafayette, IN*

Saunders, Sam C., *Scientific Consulting Services, Inc., Kirkland, WA*

Schaefer, Edward F., *Santa Clara University, Santa Clara, CA*

Schattschneider, Doris J., *Moravian College, Bethlehem, PA*

Schumer, Peter D., *Middlebury College, Middlebury, VT*

Scott, David R., *University of Puget Sound, Tacoma, WA*

Segal, Sanford L., *University of Rochester, Rochester, NY*

Sherman, Gary, *Rose-Hulman Institute of Technology, Terre Haute, IN*

Sloughter, Daniel C., *Furman University, Greenville, SC*

Smith, Bryan A., *University of Puget Sound, Tacoma, WA*

Steinbach, Peter M., *Albuquerque Technical-Vocational Institute, Albuquerque, NM*

Stone, Arthur, *Northeastern University, Boston, MA*

Straffin, Jr., Philip, *Beloit College, Beloit, WI*

Stroyan, Keith, *University of Iowa, Iowa City, IA*

Tikoo, Mohan Lal, *Southeast Missouri State University, Cape Girardeau, MO*

Tucker, Thomas W., *Colgate University, Hamilton, NY*

Van Brummelen, Glen, *Bennington College, Bennington, VT*

Wagon, Stan, *Macalester College, St. Paul, MN*

Walden, Byron, *Santa Clara University, Santa Clara, CA*

Watkins, John, *Colorado College, Colorado Springs, CO*

Whitehead, Tamsen, *Santa Clara University, Santa Clara, CA*

Zulli, Louis, *Lafayette College, Easton, PA*

Index to Volume 73

AUTHORS

TITLES

PROBLEMS

The letters P, Q, *and* S *refer to Proposals, Quickies, and Solutions, respectively; page numbers appear in parentheses. For example, P1602(240) refers to Proposal 1602, which appears on page 240.*

Letters to (and from) the Editor

Dear Editor:

In the note "Edge-length of Tetrahedra" (June 2000), Hans Samelson proves that $3E(\Delta') < 4E(\Delta)$, where Δ' is a tetrahedron contained in the tetrahedron Δ, and $E(\Delta)$ denotes the total edge length of Δ, etc. At the end the author notes that his proof generalizes to simplices and gives the corresponding inequalities depending on whether or not the dimension is even or odd.

The following much more general result is known (see C. Linderholm, An inequality for simplices, *Geom. Dedicata* 21 (1986), 67–73):

Let $m \in \{1, 2, \ldots, n\}$ and let M_m, M'_m be the total m-dimensional content of all the m-dimensional faces of the n-dimensional simplices S, S', respectively. Then if $S' \subset S$ and $n + 1 = q(m + 1) + r$, where q and r are integers with $0 \leq r \leq m$,

$$M'_m < M_m \frac{q^{m+1-r}(q+1)^r}{n+1-m},$$

and the fraction on the right hand side is the best possible. There can be equality if we allow degenerate simplices.

Murray S. Klamkin
University of Alberta, Edmonton, Canada T6G 2G1

Dear Readers:

With this issue I complete my 5-year term as Editor of *Mathematics Magazine*; the new Editor is Frank Farris, of Santa Clara University. Serving as Editor has been a pleasure and a privilege. Bringing the *Magazine* to print requires an enormous amount of work—the large majority of it done by others than the Editor. It's my pleasure to thank some of them.

Authors deserve our greatest thanks; over 300 of them have published pieces in the *Magazine* in the last 5 years. (Many more authors submitted articles we were unable to publish, but often for reasons having less to do with merit than with space.) Over 200 *referees* also served during the same period; their careful, generous, and too-little-requited advice and effort materially improved almost every Article and Note published in the *Magazine*. *Associate editors* (listed on the first inside page of each issue) selflessly and reliably contributed hard work and sage counsel. *Harry Waldman*, Journals Editorial Manager at the MAA, skillfully and unflappably handled production and other tasks large and small. *Mary Kay Peterson*, my editorial assistant at St. Olaf, brought (much-needed) organizational, technical, and grammatical expertise and efficiency to our office.

Finally, I thank *readers*. Your general interest, meticulous reading, letters to the editor (pro *and* con), phone calls, and e-mails have all improved the product, and helped complete the circle of communication for which the *Magazine* exists.

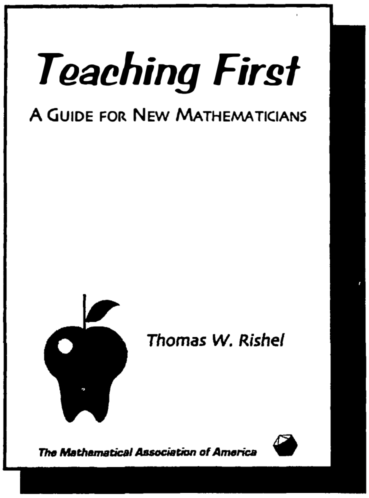
Paul Zorn
Saint Olaf College, Northfield, Minnesota 55057

The Mathematical Association of America

Teaching First: A Guide for New Mathematicians

Thomas W. Rishel

Series: MAA Notes



In this volume Thomas Rishel draws on his nearly forty years of teaching experience to address the "nuts and bolts" issues of teaching college mathematics. This book is written for the mathematics TA or young faculty member who may be wondering just where and how to start. Rishel opens the readers' eyes to pitfalls they may never have considered, and offers advice for balancing an obligation "to the student" with an obligation "to mathematics." Throughout, he provides answers to seemingly daunting questions shared by most new TAs, such as how to keep a classroom active and lively; how to prepare writing assignments, tests, and quizzes; how exactly to write a letter of recommendation; and how to pace, minute by minute, the "mathematical talks" one will be called upon to give.

This book is Rishel's answer to those who may suggest that good teaching is innate and cannot be taught. This he emphatically denies, and he insists that solid teaching starts with often overlooked "seeming trivialites" that one needs to master before exploring theories of learning. Along the way he also covers the general issues that teachers of all subjects eventually experience: fairness in grading, professionalism among students and colleagues, identifying and understanding student "types", technology in the classroom. All of the subjects in this book are considered within the context of Rishel's experience as a mathematics teacher. All are illustrated with anecdotes and suggestions specific to the teaching of mathematics.

Teaching First is a comprehensive guide for a mathematics TA, from the first semester preparations through the unforseen challenges of accepting a faculty position. Its aim is to prepare the new TA with clear suggestions for rapidly improving their teaching abilities.

Catalog Code: NTE-54/JR 150 pp., Paperbound, 2000 ISBN 088385-165-2 List: $19.00 MAA Member: $15.00

Name______________________ Credit Card No.______________________

Address______________________ Signature______________________Exp. Date____/____

City______________________ Qty________ Price $____________ Amount $____________

State____________ Zip ____________ Shipping and Handling $____________

Phone ______________________ **Catalog Code: NTE-54/JR** Total $____________

Shipping and Handling: Postage and handling are charged as follows: USA orders (shipped via UPS): $2.95 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 10 days 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: $3.50 per item ordered for books sent surface mail. Airmail service is available at a rate of $7.00 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

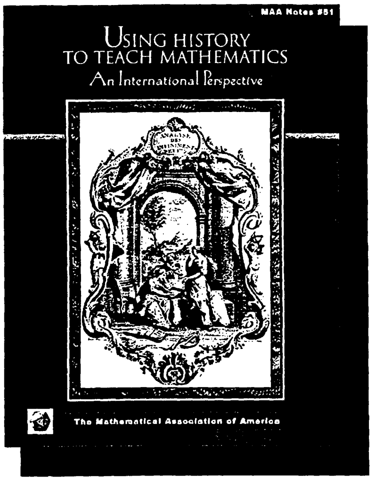
Web: www.maa.org

The Mathematical Association of America

Using History to Teach Mathematics

Victor Katz, editor

Series: MAA Notes



This book is a collection of articles by international specialists in the history of mathematics and its use in teaching, based on presentations given at an international conference in 1996. Although the articles vary in technical or educational level and in the level of generality, they show how and why an understanding of the history of mathematics is necessary for informed teaching of various subjects in the mathematics curriculum, both at secondary and at university levels. Many of the articles can serve teachers directly as the basis of classroom lessons, while others will give teachers plenty to think about in designing courses or entire curricula. For example, there are articles dealing with the teaching of geometry and quadratic equations to high school students, of the notion of pi at various levels, and of linear algebra, combinatorics, and geometry to university students. But there is also an article showing how to use historical problems in various courses and one dealing with mathematical anomalies and their classroom use.

Although the primary aim of the book is the teaching of mathematics through its history, some of the articles deal more directly with topics in the history of mathematics not usually found in textbooks. These articles will give teachers valuable background. They include one on the background of Mesopotamian mathematics by one of the world's experts in this field, one on the development of mathematics in Latin America by a mathematician who has done much primary research in this little known field, and another on the development of mathematics in Portugal, a country whose mathematical accomplishments are little known. Finally, an article on the reasons for studying mathematics in Medieval Islam will give all teachers food for thought when they discuss similar questions, while a short play dealing with the work of Augustus DeMorgan will help teachers provide an outlet for their class thespians.

Catalog Code: NTE-51/JR 300 pp., Paperbound, 2000 ISBN 088385-163-6 List: $32.95 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-51/JR** Total $_____

Shipping and Handling: Postage and handling are charged as follows: USA orders (shipped via UPS): $2.95 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 10 days 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: $3.50 per item ordered for books sent surface mail. Airmail service is available at a rate of $7.00 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:

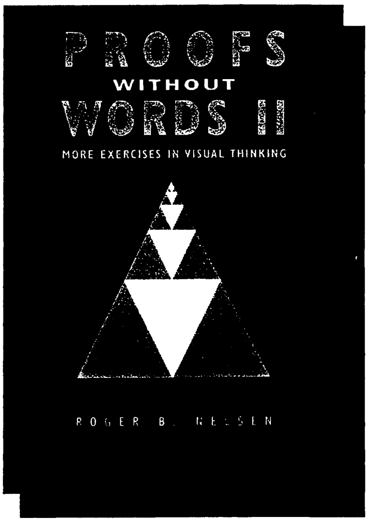
Phone: 1 (800) 331.1622
Fax: (301) 206.9789
Mail: Mathematical Association of America
PO Box 91112
Washington, DC 20090-1112
Web: www.maa.org

The Mathematical Association of America

Proofs Without Words II: More Exercises in Visual Thinking

Roger B. Nelsen

Series: Classroom Resource Materials



What are "proofs without words?" Many would argue that they are not really "proofs" (nor, for that matter, are many "without words," on account of equations which often accompany them). Like its predecessor *Proofs Without Words*, published by the MAA in 1993, this book is a collection of pictures or diagrams that help the reader see why a particular mathematical statement may be true, and also to see how one might begin to go about proving it true. The emphasis is on providing visual clues to the observer to stimulate mathematical thought.

Proofs without words have been around for a long time. In this volume you find modern renditions of proofs without words from ancient China, tenth century Arabia, and Renaissance Italy. While the majority of the proofs without words in this book originally appeared in journals published by the MAA, others first appeared in journals published by other organizations in the US and abroad, and on the World Wide Web.

The proofs in this collection are arranged by topic into five chapters: geometry and algebra; trigonometry, calculus and & analytic geometry; inequaltieis; integer sums, and infinite series & linear algebra. Although the proofs without words are presented primarily for the enjoyment of the reader, teachers will want to use them with students at many levels - in precalculus courses in high school, in college courses in calculus, number theory and combinatorics, and in pre-service and in-service classes for teachers.

Catalog Code: PW2/JR 142 pp., Paperbound, 2000 ISBN 088385-721-9 List: $24.00 MAA Member: $18.95

Name__________ Credit Card No.__________

Address__________ Signature__________ Exp. Date____/____

City__________ Qty_____ Price $_____ Amount $_____

State_____ Zip _____ Shipping and Handling $_____

Phone __________ Catalog Code: PW2/JR Total $_____

Shipping and Handling: Postage and handling are charged as follows: USA orders (shipped via UPS): $2.95 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 10 days 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:** $3.50 per item ordered for books sent surface mail. Airmail service is available at a rate of $7.00 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

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW
Washington, D.C. 20036