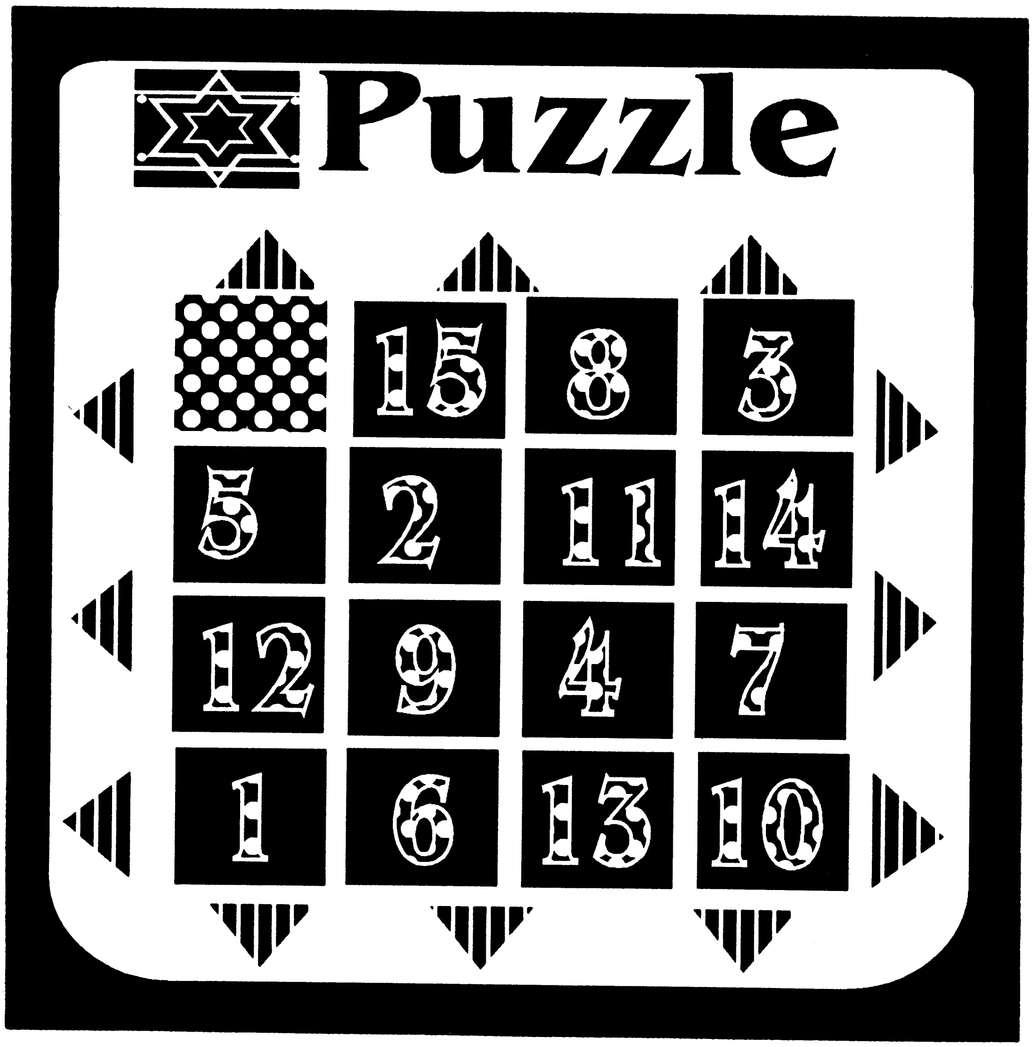
Vol. 66 No. 3 June 1993

MATHEMATICS MAGAZINE



- The Symbolic Dynamics of the Sequence of Pedal Triangles
- The Knight's Tour on the 15-Puzzle
- Fiber Optics and Fibonacci

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

EDITORIAL POLICY

The aim of *Mathematics Magazine* is to provide lively and appealing mathematical exposition. This is not a research journal and, in general, the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for an article 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. Articles on pedagogy alone, unaccompanied by interesting mathematics, are not suitable. Neither are articles consisting mainly of computer programs unless these are essential to the presentation of some good mathematics. Manuscripts on history are especially welcome, as are those showing relationships between various branches of mathematics and between mathematics and other disciplines.

The full statement of editorial policy appears in this *Magazine*, Vol. 64, pp. 71–72, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, nor published by another journal or publisher.

Send new manuscripts to: Martha Siegel, Editor, Mathematics Magazine, Towson State University, Towson, MD 21204. Manuscripts should be typewritten and double spaced and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit the original and two 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. Illustrations should be carefully prepared on separate sheets in black ink, the original without lettering and two copies with lettering added. Do not use staples.

Cover design by Carolyn Westbrook.

AUTHORS

James Alexander received his Ph.D. in algebraic topology from Johns Hopkins University in 1968. He has spent his career at the University of Maryland, although his mathematical interests have changed to applied mathematics. This year, he is also a program director in applied dynamics at the National Science Foundation. The present article arose when the author's interest was piqued while scanning journals in the faculty lounge. (The views in the article do not represent those of the NSF.)

Spencer P. Hurd was a physics undergraduate at Virginia Tech and got his Ph.D. in mathematics studying lattice-ordered groups under Steve McCleary at the University of Georgia. A former state chess champion, he likes research problems with a number-theoretic or combinatorial flavor that have associated diagrams, and, since he was 9 years old, has wondered if those "hard" positions on the 15-puzzle were really impossible.

David A. Trautman received his B.A. from Drake University in 1976 and his M.A. (1978) and Ph.D. (1983) from the University of Missouri-Columbia. He has published papers concerning Banach Space Theory and Electrical Engineering.

William P. Dotson earned his Ph.D. from the Air Force Institute of Technology in 1976. Trained as an electrical engineer, he regards himself as a mathematician in work clothes. His current research areas are computationally fast methods for solving the k-terminal reliability problem, and approximation methods for designing large communications networks. This paper came as a result of working on a problem involving the design of a highly reliable fiber optic ring network for NASA's Mission Control Center in Houston.

Frederick Norwood received his Ph.D. from the University of Southwestern Louisiana, a hotbed of algebraic topology down on the bayou. He has written numerous mathematical articles, also fiction and verse, and he is the editor of a book about the Prince Valiant comic strip. He is currently Assistant Professor at East Tennessee State University, where he does research in knot theory.

Charles Taylor received his B.S. degree from East Tennessee State College and his M.A. in mathematics from the University of Tennessee. He has taught mathematics for almost 32 years at East Tennessee State University. His primary interests are in algebra and number theory.

Vol. 66 No. 3 June 1993

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.

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

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

PRINTED IN THE UNITED STATES OF AMERICA

The Symbolic Dynamics of the Sequence of Pedal Triangles

J. C. ALEXANDER[†]
University of Maryland
College Park, MD 20742

Introduction

It is standard in plane geometry to construct the three medians, angle bisectors, or altitudes of a given triangle T. The three lines of any set intersect the opposite sides (or their extensions) of the triangle in three points (the feet) that can be taken as the vertices of a new triangle T'. The process can be iterated. Of course the successive triangles become smaller, but we concentrate only on their angles. For example, the triangle formed from the feet of the medians is similar to the original triangle; the angles are unchanged. The successive triangles formed from the angle bisectors limit to an equi-angular triangle. On the other hand, the angles of the successive triangles formed from the feet of the altitudes behave in a much more complicated manner. In some sense, every type of behavior is possible. For any r, there are triangles whose angles repeat after r iterations (for example, a triangle with angles $36°$, $72°$, $72°$ is similar to its second iterate; a triangle with angles $12°$, $36°$, $132°$ is similar to its fourth iterate). For any r and k, there are triangles whose angles repeat every r iterations after an initial delay of k iterations. For each k, there are triangles that after k iterations are equiangular (for example, $60°$, $105°$, $15°$ with $k = 2$). There are lots of triangles whose angles never repeat (uncountably many). Indeed, there are triangles whose angles come arbitrarily close to any given triangle. There are triangles that become almost equiangular and stay that way for, say 1,000,000 iterations, but that then do something completely different.

Hobson [3] called the triangle T' formed from the feet of the altitudes of a triangle T the *pedal triangle*. Recently Kingston and Synge [4] revisited and corrected Hobson's work and, for example, determined a criterion for some pedal iterate of T to have the same angles as T. The purpose of the present article is to revisit the issue again from a different mathematical point of view, which makes it routine to understand the behavior of the angles of successive pedal triangles.

Let us consider the question of triangles whose angles repeat after r pedal iterations. Determining criteria directly on the angles leads to rather technical conditions [4] (and in fact is where Hobson made mistakes). A recurring theme in mathematics is to relabel objects so that desirable properties are more evident, and that is our present theme. For each triangle T, we assign a label $E(T)$ that makes the behavior under the iterated pedal map obvious. Equally important, there is a straightforward way of determining the angles of the triangle from its label.

More precisely, we consider all infinite sequences $a_1a_2a_3\dots$ of four symbols, say each a_i equals 0, 1, 2, or 3. Every triangle T is labelled with one such sequence $E(T)$, and two triangles are labelled with the same sequence if and only if they are

[†]Supported by the National Science Foundation.

similar. Moreover, if $E(T)=a_1a_2a_3\ldots$, then $E(T')=a_2a_3\ldots$ obtained by erasing the first symbol in $E(T)$. Then if for example, 013013013... (infinitely repeating) is the label of a triangle T, it is clear that the third pedal iterate of T is similar to T. We show that almost all labels correspond to a triangle, and we develop an algorithm to find the triangle from its label.

This type of labelling is the subject of *symbolic dynamics*, which is quite powerful when it works. Unfortunately, there is no good introductory exposition of symbolic dynamics known to the author. Most expositions are rather terse discussions in advanced monographs, found under the subject headings of "measurable dynamics," "ergodic theory," or "Bernoulli shifts"; see for example [2], Ch. 7, 17, 25; [6], Ch. V; [1], Ch. 2. This situation may change in the near future; there are several elementary books on dynamical systems in the process of being written and soon the subject may be more available. For the present, the author hopes that, beyond whatever interest the pedal sequence has per se, this article serves as an elementary introduction to the power and beauty of symbolic dynamics. Accordingly, instead of proceeding directly to the question of pedal triangles, we develop some of the language and ideas of symbolic dynamics.

Let us thus begin with a triangle T. It is not difficult to determine a formula for the angles of the pedal triangle T' in terms of the angles A, B, and C of the original triangle T. The formula can be determined by classical synthetic methods [3] or by coordinate geometry: Position the triangle with one edge along the horizontal coordinate axis with the foot of the altitude at the origin and compute slopes—the computation is left to the reader. The result is:

$$A'=180^\circ-2A,\qquad B'=180^\circ-2B,\qquad C'=180^\circ-2C,\quad \text{if } A,B,C \text{ are all } \le 90^\circ \tag{$*$}$$

$$A'=2A-180^\circ,\qquad B'=2B,\qquad C'=2C,\qquad \text{if } A>90^\circ, \tag{$**$}$$

with similar formulae for obtuse B and C. Let us recast our object of study. We consider three angles (A,B,C) (in order) that add up to 180°, and (with a suggestive abuse of notation—but perhaps not etymology) call them a *triangle* $T=T_1$. The formulae (*) and (**) yield a new triangle $T'=T_2=F(T)=(A',B',C')$, called the pedal triangle of T. By iteration, we let T_n be the pedal triangle $F(T_{n-1})$. Note that we consider *ordered* triangles. That is, the three angles $(\angle 1, \angle 2, \angle 3)$ are listed in order—think of the vertices of the triangle as labelled 1, 2, and 3—and two sets of angles are equal if and only if they are equal as ordered triples. The case of unordered triangles is an easy consequence of the case of ordered triangles.

There is one technical difficulty that we must consider. The pedal triangle T' of a right triangle T has one or more angles of 0°. We call triangles with one or two angles of 0° *degenerate*. In fact, (*) and (**) make sense for degenerate triangles (and if T is degenerate, so is T') and we can define the pedal sequence for such triangles. A triangle $T=T_1$ such that some T_n is degenerate is called *eventually degenerate* (Kingston and Synge [4] call them *pedal deficient*). The set of eventually degenerate triangles can be completely characterized by our methods.

A mapping such as $T\mapsto T'=F(T)$, taking a set to itself (in this case the set of triangles), is called a *dynamical system*. The study of dynamical systems has its own terminology. For example, if $F(T)=T$, then T is called a *fixed point*. If $T_r=F^r(T)=T$, then T is called a *periodic point* of period r.

The essence of symbolic dynamics is to encode, or label, the objects in a way such that the properties of the mapping F are transparent. Indeed, a standard technique in mathematics is to find representations of objects that are better suited to the task at hand. Perhaps the following analogy will explain the philosophy. For most activities, our names work perfectly well as identifiers of us, as human beings. Sometimes, however, we are identified as numbers, for example, student ID numbers. This is a code, and for some purposes, e.g., registering for a course, the code is useful whereas the name may not be. Of course, one wants to be able to recover the name from the ID number, i.e., decode. In the same way, although triangles are usually "named" by their angles, for some purposes, such as studying the pedal map, the encoding developed below is more useful. For example, it is absolutely trivial to write down codes for all periodic or eventually equiangular triangles. Moreover, once the decoding is explained, it is straightforward to implement—either by hand or by computer. Thus, for example, it is straightforward to determine the triangles, named by their angles, which are periodic or eventually equiangular.

In addition, symbolic dynamics is modern applied mathematics. It is used by theoretical computer scientists to study the behavior of bits in a computer. Below, the concept of shift is defined, and multiplication by 2 in a binary computer is just a shift. Symbolic dynamics is also used in studying chaotic systems. One of the hallmarks of chaotic systems is that they (or part of them) can be encoded by shifts. In some sense, we can say that the pedal sequence is chaotic.

In the remainder of this section, we quickly explain the encoding, without details. In section 2, by way of introduction to symbolic dynamics, we consider an elementary example in detail. In section 3, we consider the details of the symbolic dynamics of the pedal sequence. In section 4, we use these symbolic dynamics to study the pedal sequence. It is this section that the reader should hope to understand and should follow with pencil and paper. In section 5, we consider some related questions.

Let us assign an *obtuseness label* 0, 1, 2, or 3 to any triangle T_1. If the triangle is obtuse with $\angle i$ larger than $90°$, we give T_1 the label $a_1 = i$. If all angles are no more than $90°$, we give it the label $a_1 = 0$. Similarly, let the obtuseness label of the pedal triangle $T_2 = T_1'$ be a_2, and so on. In this way, we generate a sequence of labels $a_1, a_2, \ldots$; where a_i is the obtuseness label of the ith pedal triangle. It is convenient to string them together to form an infinite *obtuseness word*

$$a_1 a_2 a_3 \ldots .$$

Clearly the pedal triangle T_2 corresponds, by definition, to the *shifted* sequence

$$a_2 a_3 a_4 \ldots$$

obtained by erasing the first symbol, and so on. A set of such infinite sequences with the transformation

$$G: a_1 a_2 a_3 \ldots \mapsto a_2 a_3 a_4 \ldots$$

is called a *shift*, in this case, on four symbols (clearly the choice of symbols is irrelevant—one could use ♠, ♥, ◇, ♣, but integers are both conventional and convenient).

For example, suppose the angles of T are $9°$, $89°$, $82°$. Then the sequence of pedal triangles is

$$\begin{pmatrix} 9° \\ 89° \\ 82° \end{pmatrix} \mapsto \begin{pmatrix} \triangleright 162° \\ 2° \\ 16° \end{pmatrix} \mapsto \begin{pmatrix} \triangleright 144° \\ 4° \\ 32° \end{pmatrix} \mapsto \begin{pmatrix} \triangleright 108° \\ 8° \\ 64° \end{pmatrix} \mapsto \begin{pmatrix} 36° \\ 16° \\ \triangleright 128° \end{pmatrix} \mapsto \ldots .$$

Here obtuse angles have been marked with a pointer. Thus the associated obtuseness word is 01113.... The pedal triangle T' is encoded as the shifted sequence 1113....

The following is the formal statement of our result:

THEOREM. *The correspondence between a triangle and its obtuseness sequence is an isomorphism between the dynamical system of the pedal transformation F and the shift on four symbols.*

We have not defined the notion of isomorphism, and in fact there are some technical points involving measure theory, but in the present case, it means that every triangle is uniquely specified by its obtuseness sequence $a_0a_1a_2\ldots$, and conversely, almost every sequence gives rise to a triangle. The most interesting part is that the isomorphism is explicit. The obtuseness sequence encodes a triangle as a shift. We develop an algorithm for determining the triangle from its code. The shift is easy to understand, and the algorithm makes it easy to translate to triangles. For example, it is then routine to write down a triangle that repeats after r iterates. Such is the power of symbolic dynamics.

Symbolic Dynamics

In this section, we make some general definitions and consider the most simple, and in fact seminal, example of coding a dynamical system by symbolic dynamics.

Consider the n symbols $0, 1, 2, \ldots, n-1$. Out of these we form infinite *words* $a = a_1a_2a_3\ldots$, where each a_i is one of the symbols $0, \ldots, n-1$. The set of all such words is the *sequence space* S_n on the n symbols. The use of integers for symbols is for convenience and custom; the particular choice of symbols is not important. There is a transformation on S_n, namely the *shift on n symbols*:

$$a_1a_2a_3\ldots \mapsto a_2a_3a_4\ldots.$$

Many features of the shift can be read off easily. Note, for example, that it is trivial to write down words that are periodic under the shift. These are the words that are periodic in their symbols. We use the standard notation

$$\overline{a_1a_2a_3\ldots a_r}$$

to mean the infinitely repeating word

$$a_1a_2a_3\ldots a_ra_1a_2a_3\ldots a_ra_1a_2a_3\ldots a_r\ldots.$$

Clearly under the shift, the word

$$\overline{a_1a_2a_3\ldots a_r}$$

is r-periodic. More generally

$$a_1a_2\ldots a_k\overline{a_{k+1}a_{k+2}\ldots a_{k+r}}$$

denotes the word

$$a_1a_2\ldots a_ka_{k+1}a_{k+2}\ldots a_{k+r}a_{k+1}a_{k+2}\ldots a_{k+r}a_{k+1}a_{k+2}\ldots a_{k+r}\ldots,$$

that is eventually periodic under the shift.

Suppose we have a set S with a mapping $F\colon S \to S$. To represent F by the shift G on n symbols means to assign to each $s \in S$ a word $E(s) \in S_n$ so that $GE = EF$. The assignment E is called the *encoding*. Thus the encoding of $F(s)$ is the shift of the encoding of s. The opposite process $D\colon S_n \to S$ of assigning an element of S to each word is called *decoding*.

To clarify all this abstraction, we consider an example. Consider the set S of reals x, $0 \le x < 1$ with the transformation $F(x) = 2x \pmod 1$. We find an encoding E. Write any x in binary notation:

$$x = .a_1 a_2 a_3 \dots, \qquad \text{each } a_i = 0 \text{ or } 1.$$

Then $F(x)$ has the binary representation

$$F(x) = .a_2 a_3 a_4 \dots$$

(multiply by 2 (= 10 in binary) and drop the integral part). Thus the encoding E that takes x to its string of binary digits represents F as the shift on two symbols 0 and 1. It is simple to decode, that is, go from a word $a_1 a_2 a_3 \dots$ to the represented number:

$$D(a_1 a_2 a_3 \dots) = x = \sum_{i=1}^{\infty} a_i 2^{-i}.$$

The encoding E is one-to-one, but not quite onto. There is an unavoidable ambiguity in binary notation for binary rationals—numbers that are integral multiples of some negative power of 2. That is, the two sequences

$$.a_1 a_2 \dots 0_k \bar{1}$$

and

$$.a_1 a_2 \dots 1_k \bar{0}$$

are binary representations of the same number. Here we use the overbar to denote infinite repetition, and 0_k, for example, means the entry in the kth place is a 0. We must make a choice. For example, we choose to encode each binary rational by the word terminating in an infinite string of 0s. Thus there are certain *prohibited words* that are not in the image of the encoding E. Here the prohibited words are those terminating in an infinite string of 1s.

Observe that, for example, it is easy to answer questions about numbers that are periodic under the mapping F. For example, $.\overline{01}$ should be periodic of period 2. Indeed it is: $D(.\overline{01}) = 1/3$ and the iterates of $1/3$ under F are $1/3, 2/3, 1/3, \dots$.

Consider the two subsets M_0 and M_1 of S where

$$M_0 = \left\{x \in S: 0 \le x < \frac{1}{2}\right\}, \qquad M_1 = \left\{x \in S: \frac{1}{2} \le x < 1\right\}.$$

The coding E can be defined in terms of F and the M_i. Namely, the kth digit in $E(x)$ is 0 or 1 according as $F^{k-1}(x)$ is in M_0 or M_1. Such sets M_0 and M_1 form what is called a *Markov partition* of S for F. Note that if we chose to include the point $\frac{1}{2}$ in M_1 instead of M_0, the resulting encoding would permit words terminating in an infinite string of 1s (except $.\bar{1}$) and prohibit those terminating in an infinite string of 0s (except $.\bar{0}$).

The two sets M_0 and M_1 can be further divided into

$$M_{00} = \left\{x \in S: 0 \le x < \frac{1}{4}\right\}, \qquad M_{01} = \left\{x \in S: \frac{1}{4} \le x < \frac{1}{2}\right\},$$

$$M_{10} = \left\{x \in S: \frac{1}{2} \le x < \frac{3}{4}\right\}, \qquad M_{11} = \left\{x \in S: \frac{3}{4} \le x < 1\right\}.$$

The binary expansions of the points in the $M_{a_1a_2}$ are $.a_1a_2\ldots$. Similarly, the $M_{a_1a_2}$ can be divided into $M_{a_1a_2a_3}$, etc. That is, the Markov partition leads to further partitions such that the subscripts on the partition sets correspond to the initial string of the encoding. This of course is nothing more than the standard process of localizing on the number line the point corresponding to a given binary expansion. It is a geometric way of decoding.

Encoding a transformation is thus a quite natural representation of the transformation. Markov partitions are an easy way to see the encoding and also the decoding. The idea that a binary (or ternary or decimal or whatever) expansion of a number is an encoding was the beginning of symbolic dynamics.

The Pedal Sequence

Following [4], we coordinatize the set of all triangles. Given any triangle with vertices labelled 1, 2, and 3, consider the triple

$$(x,y,z) = (\angle 1/180^\circ, \angle 2/180^\circ, \angle 3/180^\circ).$$

The set of such triples form the triangle

$$M = \{(x,y,z): x+y+z=1,\ x\geq 0,\ y\geq 0,\ z\geq 0\},$$

called the *moduli space* (FIGURE 1). The boundary of M consists of degenerate triangles, with one or more angles of zero degrees. The four sets M_i, $i=0,1,2,3$, consisting respectively of triangles with all angles acute, $\angle 1 > 90^\circ$, $\angle 2 > 90^\circ$, $\angle 3 > 90^\circ$, are four smaller triangles in FIGURE 1. The boundaries between the M_i consist of right triangles. By fiat, these boundaries between the M_i belong to M_0.

The following simple process goes from a point in the moduli space to a triangle: Drop segments from the point perpendicularly to the three boundary edges of M; the angles of the desired triangle are proportional to the lengths of the three segments.

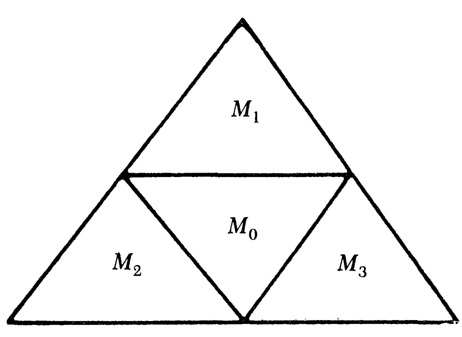


FIGURE 1

The moduli space of triangles. There are four sets M_i in a Markov partition; consisting of acute triangles and triangles obtuse in angles 1, 2, 3. The boundaries between the M_{α_1} consist of right triangles. The pedal map has a simple geometric interpretation in terms of the moduli space, namely each M_{α_1} is dilated by a linear factor of 2 and laid back on the moduli space (M_0 is also flipped). The reader is invited to sketch in the sets $M_{\alpha_1\alpha_2}$ of the text.

The four sets M_i are a Markov partition, and lead to the obtuseness encoding of the theorem. That is, if the nth pedal iterate of a triangle (counting the triangle as the first) is in M_i, the nth symbol in the corresponding word is i. The pedal maps (*) and (**) have a very simple description in terms of the moduli space M. Each of the three regions M_i, $i = 1, 2, 3$ is dilated by a factor of 2 from their exterior corners. The region M_0 is dilated by a factor of -2 from its center. That is, if the M_i were made of rubber, they would be stretched by a factor 2 and laid back over M. After stretching, M_0 would be flipped before being laid back over M. With this description, it is easy (tautological) to see that this is an encoding of the pedal transformation; that is, the shift corresponds to the pedal map.

To understand this better, we study M a little more. Divide each M_{a_1} into four equilateral triangles $M_{a_1 a_2}$ just as M is divided by the M_{a_1}. Just as a triangle corresponding to a point in M_{a_1} is encoded by a word beginning in i, a triangle corresponding to a point in $M_{a_1 a_2}$ is encoded by a word beginning in $a_1 a_2$. Similarly the $M_{a_1 a_2}$ could be divided into triangles $M_{a_1 a_2 a_3}$. The subscripts correspond to the initial segment of the encoding. One can imagine an infinite sequence of such subdivisions, locating the triangle coming from any given word.

There are certain prohibited words in this encoding. All have to do with the fact that the boundaries between M_0 and the other M_i are (arbitrarily) part of M_0. There are two types of prohibited words: (i) those that terminate in $i0\bar{i}$ for $i = 1, 2, 3$, and (ii) those that terminate in ia where $i = 1, 2, 3$, and a is any string formed from the other two nonzero numbers (for example $211\overline{31}$ is prohibited). It is a bit technical to see this, and is not particularly important for the computations below. For example, the process of the previous paragraph locates the triangle corresponding to the word $0\bar{1}$ at the middle bottom of the moduli space of FIGURE 1. Thus the word $10\bar{1}$ would "like to" be located at the middle bottom of the region M_1. However, the bottom of M_1 actually is part of M_0 and the point must be coded $00\bar{1}$. Similarly, ia leads to a triangle that is encoded $0\hat{a}$, where if a is a string of two numbers j and k, $\hat{a}$ is the string obtained by interchanging all instances of j and k by each other.

The Markov partition exhibits the encoding geometrically. We next express it algebraically. Starting with a triangle $T = T_1$, we write the three angles as fractions of $180°$, the fraction written in binary. That is

$$\begin{pmatrix} \angle 1 \\ \angle 2 \\ \angle 3 \end{pmatrix} = 180° \begin{pmatrix} .\alpha_1\alpha_2\alpha_3\ldots \\ .\beta_1\beta_2\beta_3\ldots \\ .\gamma_1\gamma_2\gamma_3\ldots \end{pmatrix}$$

where the $.\alpha_1\alpha_2\alpha_3\ldots$, etc., are strings of 0s and 1s. Call the matrix on the right the *angle matrix*. For example, if T has angles $120°, 30°, 30°$, the angle matrix is

$$\begin{pmatrix} .1\overline{01} \\ .0\overline{01} \\ .0\overline{01} \end{pmatrix}.$$

The first digit in each of the three strings of the angle matrix determines in which M_i the triangle lies and hence what the first digit is in the encoding of T. If there is a 1 in the first place of the ith row, the first digit of the encoding is i. If all three first places are 0, the first digit of the encoding is 0.

It is not possible that all three of the first digits are 1, since the sum of the three angles would be greater than 180°. Suppose two of them, say the first and second, are 1. The only possibility is

$$\begin{pmatrix} .1\bar{0} \\ .1\bar{0} \\ .\bar{0} \end{pmatrix},$$

a degenerate 90°, 90°, 0° triangle. If this is rewritten either

$$\begin{pmatrix} .1\bar{0} \\ .0\bar{1} \\ .\bar{0} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} .0\bar{1} \\ .1\bar{0} \\ .\bar{0} \end{pmatrix},$$

we see that the triangle is represented in the moduli space in the middle of the left edge. By our conventions, the correct encoding is $0\bar{3}$.

Having determined the first digit in the encoding, we find the angle matrix of the pedal triangle and repeat the process. If one of the first digits $\alpha_1, \beta_1, \gamma_1$ is 1, by (**), the angle matrix of the pedal triangle is

$$\begin{pmatrix} .\alpha_2\alpha_3\alpha_4 \dots \\ .\beta_2\beta_3\beta_4 \dots \\ .\gamma_2\gamma_3\gamma_4 \dots \end{pmatrix}.$$

If none of the first digits is 1 the angle matrix of the pedal triangle is

$$\begin{pmatrix} .\alpha'_2\alpha'_3\alpha'_4 \dots \\ .\beta'_2\beta'_3\beta'_4 \dots \\ .\gamma'_2\gamma'_3\gamma'_4 \dots \end{pmatrix}$$

where $\alpha'_i = 1 - \alpha_i$, the *complement* of α_i, etc. This of course effects the subtraction of (*) in binary.

From this angle matrix, the second digit in the encoding is determined, and so on. For example, our $120° - 30° - 30°$, triangle is encoded $1\bar{0}$. Triangles that are eventually deficient can lead to two 1s in a lead column; this case is handled by fiat, as above.

To algorithmically decode a word, we go backwards through the above process. It is straightforward to see that this leads to the following *decoding algorithm*, starting with a nonprohibited word $a = a_1a_2a_3\dots$ (with minor technical modifications that are left to the interested reader, it in fact also works for prohibited words).

A. Given a nonprohibited word $a = a_1a_2a_3\dots$, for each digit a_i, associate a column matrix

$$\begin{pmatrix} \tilde{\alpha}_i^{(1)} \\ \tilde{\beta}_i^{(1)} \\ \tilde{\gamma}_i^{(1)} \end{pmatrix}$$

as follows:

$$0 \mapsto \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad 1 \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad 2 \mapsto \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad 3 \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Concatenate these together to form an infinite 3-row matrix,

$$P_1 = \begin{pmatrix} .\alpha_1^{(1)}\alpha_2^{(1)}\alpha_3^{(1)}\dots \\ .\beta_1^{(1)}\beta_2^{(1)}\beta_3^{(1)}\dots \\ .\gamma_1^{(1)}\gamma_2^{(1)}\gamma_3^{(1)}\dots \end{pmatrix}.$$

B. Inductively obtain matrices

$$P_i = \begin{pmatrix} .\alpha_1^{(i)}\alpha_2^{(i)}\alpha_3^{(i)}\dots \\ .\beta_1^{(i)}\beta_2^{(i)}\beta_3^{(i)}\dots \\ .\gamma_1^{(i)}\gamma_2^{(i)}\gamma_3^{(i)}\dots \end{pmatrix}.$$

To obtain P_{i+1} from P_i, we consider the ith column of P_i. If this column consists of three 0s or three 1s, then for $j > i$, all elements are replaced by their complements. That is, for $j > i$,

$$\alpha_j^{(i+1)} = 1 - \alpha_j^{(i)}, \qquad \beta_j^{(i+1)} = 1 - \beta_j^{(i)}, \qquad \gamma_j^{(i+1)} = 1 - \gamma_j^{(i)}.$$

If some one element in the ith column is 1 or 0, $P_{i+1} = P_i$.

C. Note that the elements in the ith column stabilize after i iterations. That is $\alpha_j^{(i+1)} = \alpha_j^{(i)}$, etc. for $i \geq j$. Thus there is a limit that we denote

$$P = \begin{pmatrix} .\alpha_1\alpha_2\alpha_3\dots \\ .\beta_1\beta_2\beta_3\dots \\ .\gamma_1\gamma_2\gamma_3\dots \end{pmatrix},$$

that is in fact the angle matrix of the triangle.

D. The angles of the triangle are determined by evaluating the binary expansions of the angle matrix.

Computations and Consequences

The philosophy of encoding a dynamical system is that shifts are easy to work with. Thus the utility of the theorem toward investigating the sequence of pedal triangles is the decoding algorithm, which gives us an explicit triangle from any nonprohibited word. In this section, we give several illustrations. First we just work through an explicit computation. Consider the word $20\overline{1123}$. To decode it, according to the algorithm we form the matrix

$$P_1 = \begin{pmatrix} 00\overline{1100} \\ 10\overline{0010} \\ 00\overline{0001} \end{pmatrix}.$$

After processing this matrix we end up with the angle matrix

$$P = \begin{pmatrix} 00\overline{0011} \\ 10\overline{1101} \\ 00\overline{1110} \end{pmatrix}.$$

Evaluating these binary expansions (by summing geometric series), and multiplying by 180°, we find this word corresponds to the triangle with angles 9°, 129°, 42°. Note that the pedal sequence of this triangle is

$$\begin{pmatrix} 9^\circ \\ \triangleright 129^\circ \\ 42^\circ \end{pmatrix} \mapsto \begin{pmatrix} 18^\circ \\ 78^\circ \\ 84^\circ \end{pmatrix} \mapsto \begin{pmatrix} \triangleright 144^\circ \\ 24^\circ \\ 12^\circ \end{pmatrix} \mapsto \begin{pmatrix} \triangleright 108^\circ \\ 48^\circ \\ 24^\circ \end{pmatrix} \mapsto$$

$$\mapsto \begin{pmatrix} 36^\circ \\ \triangleright 96^\circ \\ 48^\circ \end{pmatrix} \mapsto \begin{pmatrix} 72^\circ \\ 12^\circ \\ \triangleright 96^\circ \end{pmatrix} \mapsto \begin{pmatrix} \triangleright 144^\circ \\ 24^\circ \\ 12^\circ \end{pmatrix} \mapsto \cdots .$$

Again the obtuse angles have been identified with a pointer, and it is clear that the obtuseness sequence is $20\overline{1123}$, as required. In particular, note that the pedal sequence cycles among the last four triangles in the sequence, as dictated by the original word. Kingston and Synge [4] call this a *cycle with a delay of* 2.

Clearly, periodic sequences correspond to periodic words. Consider for example, the word $\overline{1020}$. The decoding leads to the angle matrix

$$\begin{pmatrix} \overline{1011} \\ \overline{0001} \\ \overline{0011} \end{pmatrix}$$

and periodic pedal sequence

$$\begin{pmatrix} \triangleright 132^\circ \\ 12^\circ \\ 36^\circ \end{pmatrix} \mapsto \begin{pmatrix} 84^\circ \\ 24^\circ \\ 72^\circ \end{pmatrix} \mapsto \begin{pmatrix} 12^\circ \\ \triangleright 132^\circ \\ 36^\circ \end{pmatrix} \mapsto \begin{pmatrix} 24^\circ \\ 84^\circ \\ 72^\circ \end{pmatrix} \mapsto \begin{pmatrix} \triangleright 132^\circ \\ 12^\circ \\ 36^\circ \end{pmatrix} \mapsto \cdots .$$

Note that there are pedal sequences of any period. The reader can readily determine that the triangles encoded $\overline{0i}$ have period 2. The persistent reader can easily determine all 11 cycles of nondegenerate triangles of minimal period 3 (especially if he or she has written computer code to implement the decoding algorithm), or the two cycles of nondegenerate obtuse triangles of minimal period 3. The reader can also verify that the three sequences ([4], see (5.1), (5.2), (5.3)) are encoded, respectively, $3321\overline{0}$, $\overline{00000120}$, $\overline{0003220}$, and decode to triangles of angles

$$(180^\circ/12, 7\cdot 180^\circ/48, 37\cdot 180^\circ/48),$$
$$(83\cdot 180^\circ/255, 87\cdot 180^\circ/255, 85\cdot 180^\circ/255),$$
$$(47\cdot 180^\circ/63, 9\cdot 180^\circ/63, 7\cdot 180^\circ/63),$$

respectively. The first of these becomes equiangular after four pedal iterations, and the other two are periodic of periods 8 and 7 respectively.

We consider several questions that can be investigated via symbolic dynamics. For which triangles are all the pedal triangles acute? There is only one; it is encoded as $\overline{0}$ and is of course the equilateral triangle. This was noted in [4]. For which triangles are all the pedal triangles obtuse? This was also investigated in [4]. These triangles are the ones whose encoded words do not contain the symbol 0. There are an uncountable number of them. Incidentally, it is interesting to picture the set of such triangles in the moduli space. One deletes M_0, then the inverse image of M_0, which is an additional three smaller triangles M_{10}, M_{20}, M_{30}, then the preimage of these three triangles, etc. The resulting Cantor set may be familiar as coming from the Sierpinski triangle or Sierpinski gasket. The pedal dynamics restricted to this set is a shift on the three symbols $1, 2, 3$.

It is clear what we mean by two triangles being close to each other—their angles are close. Note that two triangles that are not eventually degenerate are close if their associated words agree for a long initial segment. It is possible to more-or-less write down the word of a triangle whose sequence of pedal triangles come arbitrarily close to all triangles. We form a word such that iterates contain any given finite string of symbols infinitely often. The orbit of such a word is dense in the moduli space. We simply take all strings of length 1 (4 of them), then all strings of length 2 (4^2 of them), etc. and concatenate them together. Thus we form the word (the wedges indicate the concatenation points)

$$0_\wedge 1_\wedge 2_\wedge 3_\wedge 00_\wedge 01_\wedge 02_\wedge 03_\wedge 10_\wedge 11_\wedge 12_\wedge 13_\wedge 20_\wedge \cdots {}_\wedge 33_\wedge 000_\wedge 001_\wedge \cdots {}_\wedge 333_\wedge \cdots .$$

This is the word for a triangle with dense orbit.

In general, the symbolic dynamics permits efficient investigation of any property of the pedal sequence that depends only on the angles. On the other hand it gives no information about geometric questions, such as whether some triangle in a periodic pedal sequence is parallel to the original triangle (a question considered in [4]).

Discussion

In this section, we discuss a variety of topics and questions suggested by the previous sections.

We leave it to the interested reader to show that the pedal sequence of isosceles triangles is the shift on two symbols.

Suppose the analogous pedal construction is made on tetrahedra in 3-space. That is, given a tetrahedron, the *pedal tetrahedron* is one with vertices at the feet of the altitudes of the given tetrahedron. Is there an analogous encoding of the dynamics?

The "tent map" (named for its graph) is defined on $0 \le x \le 1$:

$$x \mapsto \begin{cases} 2x & \text{if } 0 \le x \le \frac{1}{2}, \\ 2(1-x) & \text{if } \frac{1}{2} < x \le 1. \end{cases}$$

Using the same Markov partition as in section 2, the tent map also can be encoded as the shift on two symbols.

We have worked with ordered triangles—there are a first, second, and third angle. However, it also makes sense to study unordered triangles—the order of the angles is immaterial. It is clear that two ordered triangles correspond to the same unordered triangle if their encodings differ by a permutation of 1, 2, and 3. For example, the triangle encoded $\overline{123}$ is similar to its pedal triangle if we ignore the ordering of the angles.

It is worth noting that not every map can be represented as a shift. For example, the sequence of triangles formed from the feet of meridians are all similar to each other. The angles do not change; every element is a fixed point. Shifts have lots of elements that are not fixed points and so the sequence of meridinal triangles cannot be represented by a shift. Only maps that have particular complicated dynamics (for example, lots of periodic orbits, lots of aperiodic orbits, etc.) are isomorphic to shifts.

Symbolic dynamics actually involves something we have avoided, namely a probability measure. Maps and isomorphisms are measure-preserving, but need be defined only up to a set of measure zero. This avoids the technical difficulties we had with the binary rationals and the eventually degenerate triangles—both sets are of measure zero. On the other hand, it would be more trouble than it is worth to develop the necessary measure theory for our particular application. An important concept of dynamics on a probability space is *ergodicity*. A probability space with a map F is ergodic if the only measurable functions f such that $fF = F$ almost everywhere (an *invariant* function) are constant almost everywhere. Shifts are ergodic and hence any map isomorphic to a shift is ergodic. Lax [5] shows directly on the moduli space that the pedal sequence is ergodic. In a recent note (which appeared after this paper was written) Ungar [7] uses the shift to show that the pedal map is mixing, a stronger property than ergodicity.

REFERENCES

1. Mike Boyle, Brian Marcus, and Paul Trow, Resolving maps and the dimension group for shifts of finite type, *Mem. Amer. Math. Soc.* 377 (1987).
2. Manfred Denker, Christian Grillenberger, and Karl Sigmund, *Ergodic Theory on Compact Spaces*, Lect. Notes in Math. 527, Springer-Verlag, New York-Heidelberg-Berlin, 1976.
3. E. W. Hobson, *A Treatise on Plane Geometry*, Cambridge University Press, New York, 1897.
4. John G. Kingston and John L. Synge, The sequence of pedal triangles, *Amer. Math. Monthly* 95 (1988), 609–622.
5. Peter D. Lax, The ergodic character of sequences of pedal triangles, *Amer. Math. Monthly* 97 (1990), 377–381.
6. William Parry and Selim Tuncel, *Classification Problems in Ergodic Theory*, London Mathematical Society Lecture Notes 67, Cambridge University Press, New York, 1982.
7. Peter Ungar, Mixing property of the pedal mapping, *Amer. Math. Monthly* 97 (1990), 898–900.

The Knight's Tour on the 15-Puzzle

SPENCER P. HURD
DAVID A. TRAUTMAN
The Citadel
Charleston, SC 29409

Introduction

The 15-puzzle (or "Sliding Fifteen Puzzle," or 14-15-Puzzle) is a game that consists of a 4-by-4 array of cells, 15 of which contain blocks that are numbered from 1 to 15. The 16th cell is empty and is called the blank. A "move" consists of sliding a block horizontally or vertically into the empty space. This move of course creates a new empty space, namely the hole left by the piece just moved. The object of the most common modern variation of the game is to start with an arbitrary configuration and, by sliding the blocks around, to end with the configuration of DIAGRAM S.

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

DIAGRAM S

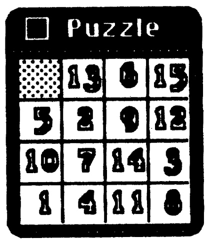
In the original form of the game, the blocks were to be arranged, by removing the blocks from the box, exactly as in DIAGRAM S except that the 14 and 15 blocks were to be transposed. The object was then to slide the blocks within the box so as to return to DIAGRAM S. In 1879, two articles appeared in the *American Journal of Mathematics* [**4**, **6**] that showed that this objective is in fact unobtainable. It is fascinating to learn that the introduction of the 15-puzzle in the 1870's caused an international sensation akin to the commotion caused by Rubik's cube almost exactly 100 years later. Hordern [**3**] gives a good account of the early history of the game.

We shall call DIAGRAM S the standard solution or the basic position in the 15-puzzle. Then, a diagram or position, say A, is defined to be *legal* if, and only, if A can be obtained on the 15-puzzle from S by sliding the numbered squares in legal fashion (removing squares from the box is not allowed). We will say, in this case, that A is reachable from S. Clearly, by reversing one's moves, S is also reachable from A. It is further clear that, if A and B are any two legal positions, then A is reachable from B (starting at B, move to the standard position S; then move to A), and vice-versa. If A is legal and if C is illegal, then A is not reachable from C, and vice-versa.

Let us digress briefly to introduce the problem of the knight's tour. Given a chessboard of m rows and n columns (for example, the standard 8-by-8 chessboard), can one place the knight in one of the squares and then move the knight by legal chess moves so that it lands in each square exactly once? Such a sequence of moves is known as a knight's tour. [Before going further, let us recall that the knight is the chess piece which, from a given square on a chessboard, can move two spaces horizontally and one vertically, or one space horizontally and two vertically. That is, a knight on square (a, b) can move to any of the eight squares $(a \pm 2, b \pm 1)$ or $(a \pm 1, b \pm 2)$ provided these squares are on the board.]

It has long been known that there is no knight's tour on a 4-by-4 chessboard. See [1], [2], or [7]. However, 15 of the 16 squares can be covered, as shown in Tour A below. This inspired us to investigate the relationship between the 15-puzzle and the knight's tour.

Our goal in this paper is to determine all the knight's tours on the 15-puzzle and to make a sublist of all the tours that yield legal positions in the 15-puzzle. In particular, we prove that Tour A is legal.



Tour A

Two Counting Principles

In determining which positions are legal, the idea of a *transposition* is fundamental. The interchanging of two *numbered* squares in a diagram will be called a transposition of that diagram. If we interchange the 14 and 15 numbers in the standard position S, we obtain a transposition, say T, of the original position. As we have mentioned, the original form of the puzzle was to transpose the 14 and 15 squares from the standard position (by removing the squares from the box) and then challenge someone to return the position to the standard one by sliding the pieces. The question is, are S and T reachable from each other? (Or, is T legal?) The original puzzle was shown by Johnson and Story [4, 6] to be impossible to solve. An interesting explanation can be found in the article by Spitznagel [5]. The main result was that starting from the standard solution, and finishing with the blank returned to the bottom right position, the only legal positions obtainable are those that could also be obtained by an even number of transpositions of numbered squares. Since there is nothing special about putting the blank in the bottom right square, Spitznagel's proof actually gives:

The Even Number Principle.

Two positions with the blank in the same square are reachable from each other if and only if they differ by an even number of transpositions of numbered squares.

We have already seen that any two legal positions are reachable from each other. We now show that the Even Number Principle implies that any two illegal positions are reachable from each other. First, let M and N be any two illegal positions that leave the same square blank, and let P be any legal position leaving that same square blank. By the Even Number Principle, there are odd positive integers m and n so that M differs from P by m transpositions and so that N differs from P by n transpositions. So M and N differ from each other by $m+n$ transpositions. Since $m+n$ is even, M and N are reachable from each other by the Even Number Principle. Now suppose that M and N are any two illegal positions. From N, slide the blank to the square that M leaves blank. This new position, say K, was reached from N by legal moves; therefore, the new position K is illegal also. But now K and M are reachable from each other by the argument just above since they are both illegal and have the blank in the same square. By moving N to K and then to M, we see that M is reachable from N. This shows any two illegal positions are reachable from each other.

Clearly reachability is an equivalence relation on the set of possible positions in the 15-puzzle, and the argument above shows that there are only two equivalence classes, the legal and the illegal positions.

As useful as it is, however, the Even Number Principle is insufficient for our purposes. It doesn't tell us about the blank square, only numbered squares. For example, if we interchange the blank square and the number 8 in Tour A just above, we get another knight's tour. Is this new tour legal?

In dealing with the blank square, a version of the taxicab metric will be helpful. The taxicab metric in geometry is a distance function that gives the distance between two points as the sum of the north-south distance and the east-west distance (as if a taxi were traveling along a grid from one point to the other). On the 4-by-4 board, let us identify each square with a pair of integer (x, y)-coordinates, placing the board in quadrant I and the lower left square at $(1, 1)$ and the upper right square at $(4, 4)$. The (x, y)-coordinates satisfy $1 \leqslant x,\ y \leqslant 4$. The taxicab distance from square (x_1, y_1) to (x_2, y_2) is $|x_1 - x_2| + |y_1 - y_2|$. The modified principle is:

The Odd Number Principle.

Suppose in diagram M *that* s *denotes a numbered square and* b *denotes the blank square. Then the position with* s *and* b *transposed is reachable from* M *if and only if the taxicab distance from* s *to* b *is an odd number.*

Proof. The proof of the Odd Number Principle is by induction on n, the taxicab distance from the blank b to the numbered square s. In any position, if $n = 1$, s is adjacent to the blank b, and so one can slide s into b. This position is certainly reachable from the first, and this transposes b and s. Suppose now that the principle is true for some number $k \geqslant 1$ and that, in the diagram M, the square s is a distance $k + 1$ from b. Without loss of generality, suppose M is legal. Let C denote the position with s and b transposed. We need to show that the Odd Number Principle applies to position C. Now, one transposition of numbered squares in M can bring s to a new square t, which is a distance k from b. Call this position B. Note that, since M is legal, B is illegal (by the Even Number Principle). Since in diagram B, b and s are a distance k from each other, the induction hypothesis can be applied. If k is odd (and $k + 1$ even), then C is reachable from B (specifically, transpose s and b in B, which is okay as k is odd, and then slide t into b). Therefore, C must be illegal as B is, and hence C is not reachable from M. If k is even (and $k + 1$ odd), then C is not reachable from B; hence C is legal and is reachable from M. Thus, if $k + 1$ is odd, C is reachable from M; if $k + 1$ is even, then C is not. It follows that the Odd Number Principle is true for $k + 1$. Therefore, the principle is true for any n, by mathematical induction. This is an unusual induction since $1 \leqslant n \leqslant 6$.

Counting the Knight's Tours

There is an argument showing that any 15-move tour must leave a corner blank and must begin (or end) in a corner on the long diagonal opposite the blank [see Appendix]. It follows, therefore, that corresponding to every tour, say M, is a corresponding one, which we shall call $t(M)$, obtained by transposing b and the opposite diagonal corner. We shall refer to $t(M)$ as the "twin" of M.

PROPOSITION 1. *The tour M is legal if and only if t(M) is illegal.*

Proof. The proposition follows immediately from the Odd Number Principle and the fact that the taxicab distance between opposite corners is the even number 6.

We can now prove that TOUR A given above is legal. By Proposition 1, TOUR A is legal if and only if t(A), pictured below, is illegal.

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

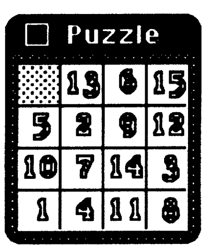
TOUR t(A)

But t(A) is obtained from S by the following sequence of transpositions:

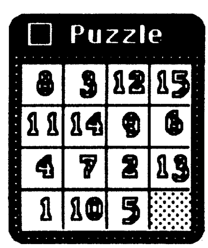
$$\langle 1,8\rangle \quad \langle 2,13\rangle \quad \langle 3,6\rangle \quad \langle 4,15\rangle \quad \langle 2,3\rangle \quad \langle 7,9\rangle$$
$$\langle 1,12\rangle \quad \langle 10,7\rangle \quad \langle 11,14\rangle \quad \langle 1,3\rangle \quad \langle 11,4\rangle.$$

Since this is a total of 11 transpositions, t(A) is illegal, and TOUR A is legal.

We now turn our attention to reflections of tours. TOUR B below is the reflection of TOUR A about the (1, 1)-(4, 4) diagonal.



Tour A



Tour B

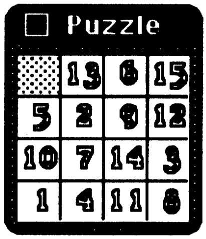
Note that the diagonal numbers are unchanged. We can think of the reflection as follows. Start with A and apply the following transpositions:

$$\langle 10,4\rangle \quad \langle 5,11\rangle \quad \langle 14,2\rangle \quad \langle 3,13\rangle \quad \langle 12,6\rangle.$$

This reaches a position we will call C. By the Even Number Principle, C is illegal. On the other hand, B = t(C); therefore, B is legal. This argument is perfectly general, and it proves the following:

PROPOSITION 2. *Suppose T is any tour, and let r(T) denote the reflection of T about its long diagonal (i.e., the diagonal not containing the blank). Then, T is legal if and only if r(T) is legal.*

Given any tour T, we let σ(T) denote the 90° clockwise rotation of T. We shall use σ^2(T) to denote $\sigma(\sigma(\text{T}))$, the 180° clockwise rotation of T. Here is an example.



A

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

σ(A)

PROPOSITION 3. *A tour T is legal if and only if $\sigma(T)$ is illegal.*

Proof. We prove more generally that, if P is any legal position, then σ(P) is illegal. For convenience, let the corner squares $(1,1),(1,4),(4,4),(4,1)$ be denoted w, x, y, and z, respectively. Then the rotation σ, restricted to these squares, cyclically shifts w to x, x to y, y to z, and z to w. This can be accomplished by the following sequence of transpositions:

$$\langle w,x\rangle \quad \langle x,z\rangle \quad \langle x,y\rangle.$$

Suppose the blank is one of w, x, y, or z. First suppose the blank is x. Then each of the transpositions above moves the blank successively an odd distance (3 units), and so by the Odd Number Principle the set of three preserves the legality of P. Now suppose w, y, or z is the blank. In these cases the blank is moved only once (but again for the odd distance 3), and the other two transpositions move numbered squares. Thus, applying both principles, the legality of P is preserved again by the set of three transpositions. Now, there are three other groups of four squares, and the cyclic action of σ on each of them is accomplished by three sets of three transpositions, nine in all. Since 9 is odd, the Even Number Principle now asserts that σ(P) is illegal.

If the blank is in the group $(1,2),(2,4),(4,3),(3,1)$ or in the group $(1,3),(3,4),(4,2),(2,1)$, then the argument just above can be applied without change. If the blank is one of the 4 center squares, then the argument is the same except that the blank will only move a distance of one unit. But 1 is odd, so the argument goes through in this last possibility. The converse is proved by observing that we have proved that P and σ(P) have opposite legality. This ends the proof.

From Proposition 3 and the earlier results we can prove the following:

PROPOSITION 4. *A tour* T *is legal if and only if* σ^2(T) *is legal. Furthermore,* $\sigma(t(\mathrm{T})) = t(\sigma(\mathrm{T}))$. *And* T *is legal if and only if* $\sigma(t(\mathrm{T}))$ *is legal.*

There is one more class of tours we need to consider. Given any tour, the knight could travel it backwards. Therefore, for any tour T, let us use b(T) to denote the "backward" version of T. Where T has $1,2,\ldots,15$, b(T) will have $15,14,\ldots,2,1$. That is, one obtains b(T) from T by the following transpositions:

$$\langle 1,15\rangle \quad \langle 2,14\rangle \quad \langle 3,13\rangle \ldots \langle 7,9\rangle.$$

Notice that the 8 and b squares are the same for T and b(T).

PROPOSITION 5. *The tour* T *is legal if and only if* b(T) *is illegal. Furthermore,* $b(t(\mathrm{T})) = t(b(\mathrm{T}))$, *and* T *is legal if and only if* $t(b(\mathrm{T}))$ *is legal.*

Proof. The proof is immediate upon noticing that there are an odd number of transpositions needed to transform T into b(T).

We now call attention to FIGURE 1. Twelve tours are shown with TOUR A at the top left. Actually, all the tours are legal. A proof, which we omit, would require only eleven arguments similar to that for the legality of TOUR A.

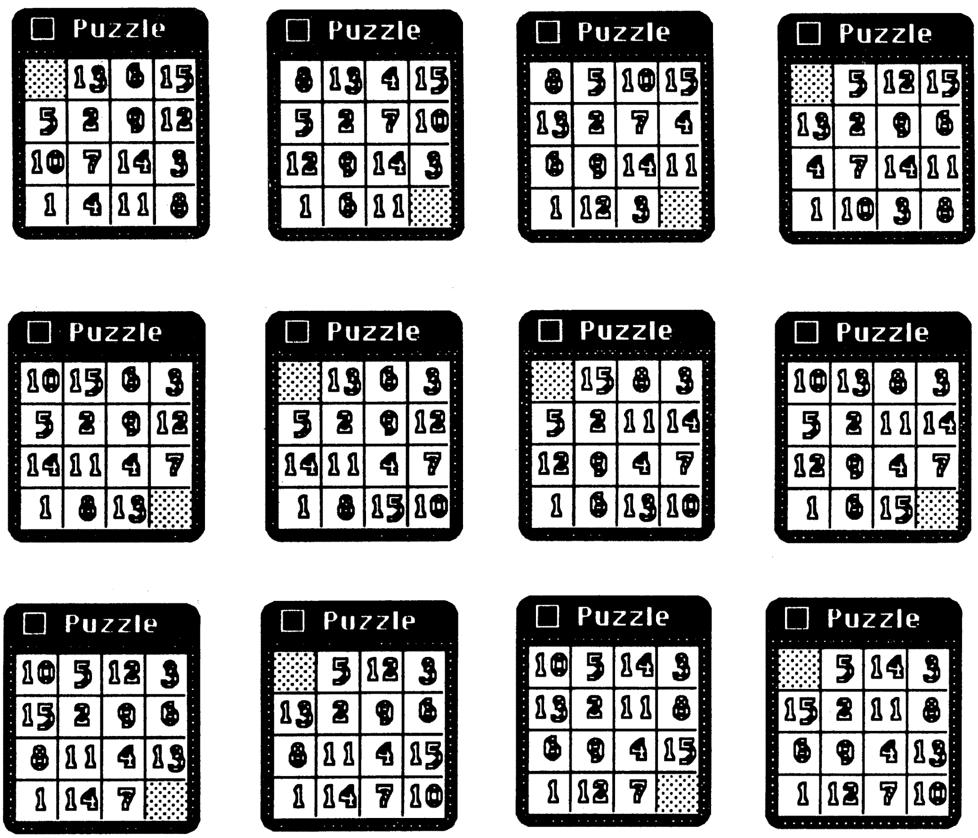


FIGURE 1
All 15-puzzle knight's tours are one of these or are constructed from one of these.

Let us look more closely at TOUR A. A little experimentation with TOUR A shows that, except for $t(\mathrm{A})$, there is no way to finish any other knight's tour that has the same first four moves (i.e., beginning with the squares numbered 1, 2, 3, 4 in TOUR A). Specifically, the only possible alternate 5th move, a move to the (3, 3) square (labeled 9 in TOUR A), does not lead to a completed tour. Furthermore, later alternate knight moves in TOUR A also do not lead to successful tours. However, an alternative 4th move, going to the (3, 4) square, does lead to a successful tour, and this is shown in FIGURE 1 in the diagram next to TOUR A. By comparing adjacent tours, ignoring twins throughout, one can follow the succession of successful alternatives. We leave the confirmation of these details to interested readers. However, we are now in a position to count the total number of possible tours.

PROPOSITION 6. *There are* 320 15-*move knight's tours on the* 4-*by*-4 *chessboard, and* 160 *of them are legal tours.*

Proof. First, the 12 legal tours pictured in FIGURE 1 all start in (1, 1), that is, in the bottom left square, and all have the first move (1, 1)-(2, 3). An inspection of alternate possibilities, alluded to above, shows that there are no other legal tours with this same first move. By Proposition 2, each of these 12 tours has a legal reflection, each of which will have first move (1, 1)-(3, 2). These exhaust the possibilities for tours starting at (1, 1). Eight of these tours from FIGURE 1 do not end in a corner. For each of these eight, apply Proposition 5. That is, the twin of the "*backward*" tour is also legal, and each of these eight, after the transformations b and t are applied, end in (1, 1) (and do not begin in a corner). Each of these eight has a legal reflection likewise ending in (1, 1) and not beginning in a corner. Thus, there are 40 legal tours beginning at (1, 1) or ending at (1, 1), and, in the latter case, not beginning in a corner.

By Proposition 4, for any T in this group of 40, $t(\sigma(\mathrm{T}))$ is legal also. This means that there are 40 legal tours based at (1, 4), and so on for squares (4, 4) and (4, 1). Thus there are 160 legal tours. There are 160 illegal tours since each of the legal ones has an illegal twin. None are duplicated in the counting since each tour is "based" uniquely at one of the four corners, either beginning there, or ending there without having begun at another corner. This ends the proof.

The transformation of one diagram on the 15-puzzle to another can be regarded as a permutation of a 16-element set. Therefore, in closing, let us refer briefly to the theory of finite permutation groups. It is well known that every permutation of a finite set can be accomplished by a sequence of transpositions, generally by several different sequences in fact. However, if one sequence of transpositions representing a particular permutation has even length, then no sequence of transpositions of odd length can also represent that transformation, and vice-versa. A particular consequence of this important theorem is that there is no ambiguity about odd or even in the Even Number Principle or in any of our proofs.

We leave for the reader the task of composing a list of all permutations on a 16-element set which, on the 15-puzzle, preserve knight's tours, or the task of deciding which of our principles and propositions are valid on the "24-puzzle" or larger sliding-piece puzzles.

APPENDIX. We shall show that, on a 4-by-4 grid, any 15-move knight's tour must begin (or end) on a corner square, and that the blank must be on an adjacent corner square. Let us suppose that we have completed a tour and that the knight did not begin at (1, 1) or (4, 4). Without loss of generality, we further suppose that the knight reached (1, 1) before arriving at (4, 4). Then the tour must have the move sequence $\ldots,(2,3),(1,1),(3,2),\ldots,$ or the sequence $\ldots,(3,2),(1,1),(2,3),\ldots.$ In either case the knight must now move into (4, 4), or it can never get there later. And as there is no way out of (4, 4) on an unused square, the tour must end at (4, 4). Therefore, the knight not having started at (1, 1) or (4, 4) must end at one of them. However, the same argument shows that, not having started at (4, 1) or (1, 4), the supposed tour must also end at one of them. Since no tour can end at two squares, we conclude that any tour that includes all four corners must begin in one corner and end in a corner on the other main diagonal. Our supposed tour thus must look like DIAGRAM C, possibly with 15 and 13 transposed, 12 and 14 transposed, 1 and 3 transposed, or 2 and 4 transposed. However, as the reader will see, the following argument is not affected by any of these transpositions.

15	X	Y	3
Y	4	14	X
X	12	2	Y
1	Y	X	13

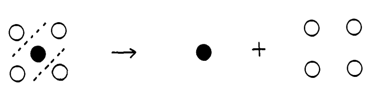
DIAGRAM C

Now, from the square containing 4, the knight may move to an X square or to a Y square. Whichever is chosen, a look at the diagram shows that moves 5 through 8 fill in the X squares or they fill in the Y squares. But now there is no place to move—there is no direct connection between the X and the Y squares. We are well short of a tour wherever the blank may be. This contradicts our assumption about a tour with four numbered squares in the four corners. So, after all, there is no tour in which the blank is not a corner, and even with the blank in a corner, the previous argument shows any tour must begin or end in a corner square on the long diagonal opposite the blank.

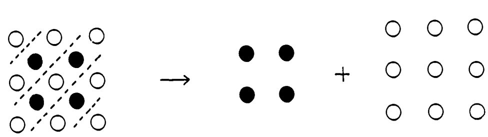
REFERENCES

1. W. W. Rouse Ball, *Mathematical Recreations and Essays*, 11th edition, Macmillan, New York, 1939.
2. Henry Ernest Dudeney, *Amusements in Mathematics*, Dover Publications, New York, 1958.
3. Edward Hordern, *Sliding Piece Puzzles*, Oxford University Press, Oxford, 1986.
4. Wm. Woolsey Johnson, Notes on the '15' puzzle, I, *Amer. J. Math.* 2 (1879), 397–399.
5. Edward L. Spitznagel, Jr., A new look at the fifteen puzzle, this MAGAZINE 40 (1967), 171–174.
6. W. E. Story, Notes on the "15" puzzle, II, *Amer. J. Math.* 2 (1879), 399–404.
7. David A. Trautman and Spencer P. Hurd, The knight's tour, preprint.

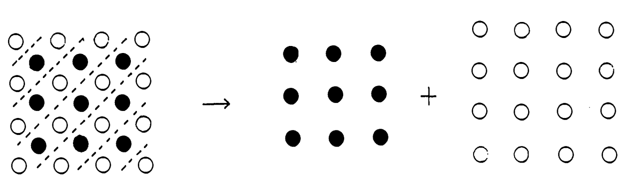
Proof without words:



$$1 + 3 + 1 = 1^2 + 2^2$$



$$1 + 3 + 5 + 3 + 1 = 2^2 + 3^2$$



$$1 + 3 + 5 + 7 + 5 + 3 + 1 = 3^2 + 4^2$$

$$\vdots$$

$$1 + 3 + 5 + \cdots + (2n-1) + (2n+1) + (2n-1) + \cdots + 5 + 3 + 1 = n^2 + (n+1)^2$$

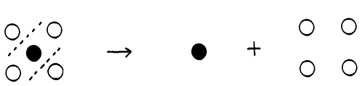
—HEE SIK KIM
CHUNGBUK NATIONAL UNIVERSITY
CHEONGJU, 360-763, KOREA

Now, from the square containing 4, the knight may move to an X square or to a Y square. Whichever is chosen, a look at the diagram shows that moves 5 through 8 fill in the X squares or they fill in the Y squares. But now there is no place to move—there is no direct connection between the X and the Y squares. We are well short of a tour wherever the blank may be. This contradicts our assumption about a tour with four numbered squares in the four corners. So, after all, there is no tour in which the blank is not a corner, and even with the blank in a corner, the previous argument shows any tour must begin or end in a corner square on the long diagonal opposite the blank.

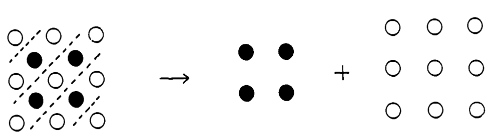
REFERENCES

1. W. W. Rouse Ball, *Mathematical Recreations and Essays*, 11th edition, Macmillan, New York, 1939.
2. Henry Ernest Dudeney, *Amusements in Mathematics*, Dover Publications, New York, 1958.
3. Edward Hordern, *Sliding Piece Puzzles*, Oxford University Press, Oxford, 1986.
4. Wm. Woolsey Johnson, Notes on the '15' puzzle, I, *Amer. J. Math.* 2 (1879), 397–399.
5. Edward L. Spitznagel, Jr., A new look at the fifteen puzzle, this MAGAZINE 40 (1967), 171–174.
6. W. E. Story, Notes on the "15" puzzle, II, *Amer. J. Math.* 2 (1879), 399–404.
7. David A. Trautman and Spencer P. Hurd, The knight's tour, preprint.

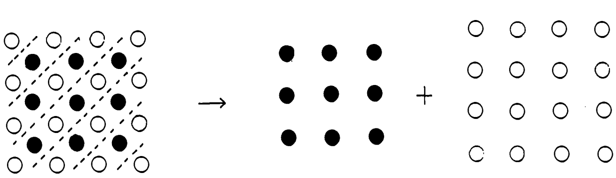
Proof without words:



$$1+3+1=1^2+2^2$$



$$1+3+5+3+1=2^2+3^2$$



$$1+3+5+7+5+3+1=3^2+4^2$$

$$\vdots$$

$$1+3+5+\cdots+(2n-1)+(2n+1)+(2n-1)+\cdots+5+3+1=n^2+(n+1)^2$$

—HEE SIK KIM
CHUNGBUK NATIONAL UNIVERSITY
CHEONGJU, 360-763, KOREA

Fiber Optics and Fibonacci

WILLIAM DOTSON
FREDERICK NORWOOD
CHARLES TAYLOR
East Tennessee State University
Johnson City, TN 37601

Introduction

A recent problem in fiber optic networks led to a surprising answer. NASA Mission Control Center wanted a design for a fiber optic network that had extremely high reliability. A fiber optic network consists of a number of stations connected by fiber optic links. Stations send and receive messages, and also pass on messages intended for other stations.

One common configuration for fiber optic networks is called a token ring. In a *single ring configuration*, the stations are arranged in a circle, with each station connected to the next by a single, one way fiber optic link (FIGURE 1). A more robust *dual ring network* has two links between each pair of adjacent stations, so that communications can flow in either direction (FIGURE 2). In this configuration, each station is connected to exactly two neighboring stations, and can transmit or pass on information to either of them. The advantage of a dual ring network is that a single failed station does not break the ring. Any number of consecutive stations in the ring can fail, and every working station can still communicate with every other working station.

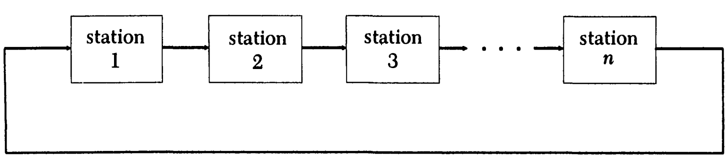


FIGURE 1
A Single Ring Network.

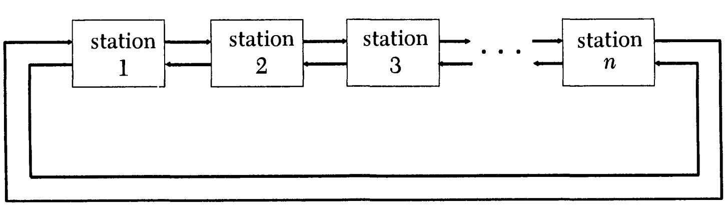


FIGURE 2
A Dual Ring Network.

A still more efficient configuration is to add bypass switches and bypass circuits to a dual ring network (FIGURE 3). If one or more of the stations fail, a bypass circuit can carry the signal past exactly one failed station, provided the next station in line is working. If, however, the next station is also down, the next bypass circuit will absorb enough additional signal power to cause an unacceptably high bit error rate in the network. Bypass circuits are passive devices, they do not supply power, they absorb it.

The stations provide the power for transmission. Stations also require a minimum amount of power at the receiving end in order to detect a signal and decode it properly. Failure to contact a working station further down the line automatically reverses the direction of the signal, and sends it back around the ring in the other direction. If a working station has a pair of consecutive failed stations on either side, then it cannot transmit across either pair and it is effectively cut off from any other working stations in the ring.

NASA chose the dual fiber optic ring with bypass circuits and bypass switches for its installation, and this is the system, as shown in FIGURE 3, that we consider in the remainder of this paper. For more technical discussions of these systems, including hardware considerations, see [3], [4], and [5].

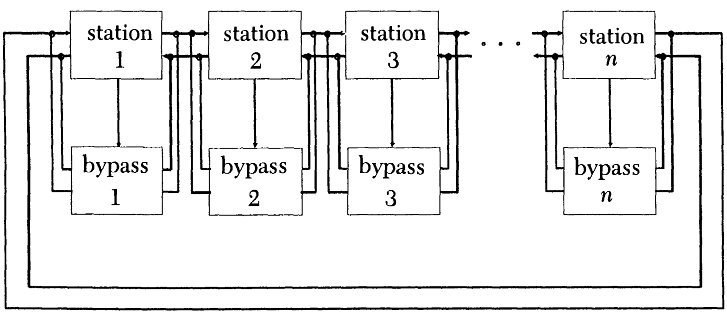


FIGURE 3
A Dual Ring Network with Bypass Switches.
The failure of a station automatically activates the bypass.

To evaluate the reliability of a dual fiber optic ring with bypass circuits, we need to count the number of different patterns in which the following can occur: Some or none, but not all, of the stations fail, but every working station that remains can still communicate with every other working station. For example, if we number the stations $1, 2, 3, 4, 5$, then one pattern would be WFWFW, meaning that stations 2 and 4 have failed, while stations 1, 3, and 5 are working. Stations can communicate around the ring in either direction, but, in this original formulation of the problem, the stations only have enough power to bypass a single failed station. Later in this paper, we will consider the case where stations have enough power to bypass more than one failed station.

The Original Problem

A ring network is called a *connected ring* when at least one station is still working and every working station can still communicate with every other working station. Thus another way to formulate the question we want to answer is this: In how many different patterns can some or none of the stations fail, while still leaving the ring connected, in a dual ring with bypass circuits?

A Linear Simplification

Since the stations are arranged in a dual ring, any number of consecutive stations can fail, and the ring will remain connected, as long as there is at least one working station. Suppose there are n stations in the ring and k consecutive stations have failed, $2 \leq k < n$. The remaining $n - k$ stations form a linked linear array (FIGURE 4). If the ring is still connected, then this linear array has the following two properties:

1) the first and the last stations have not failed, and
2) no two consecutive stations have failed.

Now, to take the problem one step further, let the number of stations in the linked linear array be $t = n - k$. In how many ways can we have s failures in a linear array of t stations that are subject to conditions 1 and 2? If s stations fail, then $t - s$ stations are still working. The s failed stations can then occupy any of the $t - s - 1$ spaces between the $t - s$ working stations, with at most one failed station in any space between any two consecutive working stations. Thus the number of different ways in which s stations can fail is the number of combinations of $t - s - 1$ spaces taken s at a time, or $C(t - s - 1, s)$. Therefore the total number of ways we can have some or no failures in a linear array of t stations subject to conditions 1 and 2 is the sum for $s = 0$ to $[(t-1)/2]$ of $C(t - s - 1, s)$ (where $[x]$ means the greatest integer in x). This total number is the sum of one of the diagonals of Pascal's triangle (FIGURE 5). Calculating these numbers for $t = 0, 1, 2, 3, 4, \ldots$ we discover that the answer is $1, 1, 2, 3, 5, \ldots$, the Fibonacci sequence.

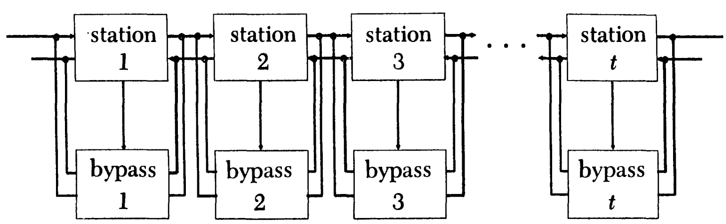


FIGURE 4
A Linear Network with Bypass Switches.

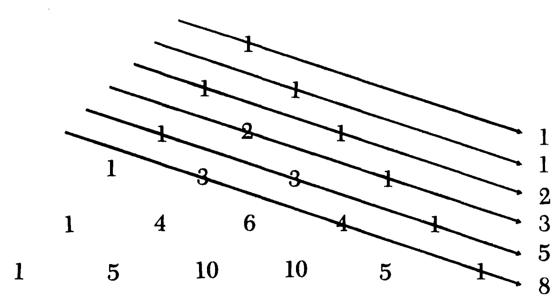


FIGURE 5
Diagonals of Pascal's Triangle sum to the Fibonacci sequence.

This comes as a surprise. Why is the answer to this question the Fibonacci sequence? In the original computation, there is nothing to suggest the defining properties of the Fibonacci sequence: that the first two terms are one, while each term thereafter is the sum of its two predecessors. It would be more satisfying to perform this computation so that the Fibonacci sequence arises in a natural manner.

We can accomplish this if we count the number of patterns in a different way. Once again, we ask: In how many ways can we have some or no failures in a linear array of t stations that are subject to conditions 1 and 2? Represent this number by $F(t)$. If t is three or more, these arrays can be separated into two disjoint classes: (i) arrays in which the next to last station fails; and, (ii) arrays in which the next to last station does not fail. If the next to last station does not fail, the number of ways that some or none of the remaining stations can fail subject to conditions 1 and 2 is $F(t-1)$. If the next to last station does fail, the number of ways some or none of the remaining stations can fail is $F(t-2)$. Therefore $F(t) = F(t-1) + F(t-2)$. Finally, subject to conditions 1 and 2, in a one-station network some or none of the stations can fail in only one way: Because of condition 1, the one station does not fail. Therefore $F(1) = 1$. In a two-station network, also, some or none of the stations can fail in only one way: Because of condition 1, neither station fails. Therefore $F(2) = 1$. It follows that $F(t)$ yields the Fibonacci sequence.

Stations in a Ring

Next we answer the original question: In how many different patterns can some or none of the stations fail, while still leaving the ring connected, in a dual ring with n stations and bypass switches (shown in FIGURE 3). Call this number $G(n)$.

First, we count the n rotations of k consecutive failures and multiply this number by $F(n-k)$. We then sum the numbers, $nF(n-k)$, as k goes from 2 to $n-1$. To this sum we add the number of ways we can have only isolated failures. The number of ways we can have only isolated failures can be calculated as follows. Since station number one and station number n are consecutive in a circular array, they cannot both fail in the case of only isolated failures. Therefore, either neither fails or exactly one of the two fails. If neither fails, then we can treat the circular array as a linear array by breaking the circle between station number n and station number one. In this case, conditions 1 and 2 are met, and there are $F(n)$ ways that the ring can sustain some or no failures and remain connected. If, on the other hand, the first station fails, then the second station and the nth station must not fail, and there are $F(n-1)$ ways that the ring can sustain some failures and remain connected. Similarly, if the nth station fails there are $F(n-1)$ ways the ring can sustain some failures and remain connected.

Therefore the total number of patterns in which some or none of the stations can fail, while still leaving the ring connected, in a dual ring with bypass switches, where the number of stations is n and where $F(n)$ is the nth Fibonacci number, is given by

$$G(n) = F(n) + 2F(n-1) + n\sum_{k=2}^{n-1} F(n-k).$$

We would like to thank the referee for pointing out that this formula appears in [2].

For $n = 1$ to 10, $G(n) = 1, 3, 7, 15, 31, 60, 113, 207, 373, 663$. Note that $G(n)$ is $2^n - 1$ for $n = 1$ to 5, because in this range every possible pattern of failures is allowed, with the single exception of all stations failing. This does not, of course, hold for values of n large enough to allow two separate strings of two or more consecutive failures. The first value of n for which this can happen is $n = 6$.

This formula, together with the probability that a station will fail, is used in calculating the reliability of a dual ring with bypass circuits.

Better Bypass Circuits

The next question it is natural to ask is: What if the bypass circuits can bypass more than one failed station? Let f represent the maximum number of failed stations that the bypass circuits can skip over. In this case the number of different patterns in which some or none of the stations fail, while still leaving the ring connected, becomes a computation using multinomial coefficients in place of binomial coefficients. Multinomial coefficients can be computed using an extended Pascal triangle. Some of the properties of an extended Pascal triangle are discussed in [1].

A Linear Simplification Again

As before, we begin by considering linear arrays. Suppose we have a linear array of t stations in which s stations have failed and in which the bypass circuits can carry a signal across f or fewer consecutive stations, all of which have failed. Let $N(t, s, f)$ be the number of different ways that this can happen while still maintaining these two conditions:

1) the first and last stations do not fail, and

2′) no more than f consecutive stations fail.

The number of working stations is $t-s$. The s failed stations can be distributed among the $t-s-1$ gaps between the working stations in such a way that no gap contains more than f failed stations. We need to know how many ways we can partition s failed stations, with every summand in the partition less than or equal to f, so that each summand in the partition can be placed in one of the $t-s-1$ gaps between working stations. This last requirement forces the number of summands in a partition to be less than or equal to $t-s-1$. This leads us to consider all partitions of s that satisfy two conditions:

a) each summand in the partition must be less than or equal to f, and

b) the number of summands in the partition cannot exceed $t-s-1$.

Let $S(t, s, f)$ denote the set of all partitions of s that satisfy conditions a and b. For example, consider $N(10,5,3)$. The partitions of 5 are: $5, 4+1, 3+2, 3+1+1, 2+2+1, 2+1+1+1, 1+1+1+1+1$. Here we reject the first two partitions, 5 and $4+1$, because they do not satisfy condition a, and we reject the last partition, $1+1+1+1+1$, because it does not satisfy condition b. Therefore $S(10,5,3) = \{3+2, 3+1+1, 2+2+1, 2+1+1+1\}$. Consider the case $2+2+1$. In how many ways can two pairs of failures and one isolated failure be distributed among $10-5-1=4$ gaps? The isolated failure can be distributed in $C(4,1)$ ways and the two pairs of failures can then be distributed in $C(4-1,2)$ ways, so the total number of ways is $C(4,1) \cdot C(3,2)$. This product of binomial coefficients is the multinomial coefficient: $4!/(1! \cdot 2! \cdot (4-1-2)!)$, which counts the number of different ways we can, from four objects, first choose one object, and then, from the remaining three objects, choose two objects. In general, if the natural number n appears k times in a partition π, we can write $\pi(n) = k$. Then $N(t, s, f)$ is equal to the sum over the partitions $\pi \in S(t, s, f)$ of:

$$(t-s-1)!/(\pi(1)! \cdot \pi(2)! \cdots \pi(f)! \cdot (t-s-1-\pi(1)-\cdots-\pi(f))!).$$

In other words, from the $t-s-1$ gaps, we choose $\pi(1)$ gaps to put isolated failures in, then, from the remaining $t-s-1-\pi(1)$ gaps we choose $\pi(2)$ gaps to put pairs of failures in, and so on.

Let $M(t,f)$ be the sum of $N(t,s,f)$ as s goes from 0 to $t-2$, so

$$M(t,f)=\sum_{s=0}^{t-2} N(t,s,f).$$

Then $M(t,f)$ is the number of different ways some or no stations in a linear array can fail, subject to conditions 1 and 2′.

Another way of looking at $M(t,f)$ is this: It is the number of different integers base $f+1$ (digits 0 through f) with the property that the sum of the digits plus the number of digits is $t-1$. The reason for this is that the nth digit can represent the number of failures (0 through f) between the nth and the $(n+1)$st working station. Then the sum of the digits is s, the number of failures, and the number of digits is $t-s-1$, the number of gaps between working stations. Therefore the sum of the digits plus the number of digits is $s+(t-s-1)=t-1$. For example, suppose $t=10$ and $f=3$. Then the base 4 number 1013 would correspond to one failed station between the first two working stations, no failed stations between the second and third working stations, one failed station between the third and fourth working stations, and three failed stations between the fourth and fifth working stations: WFWWFWFFFW.

To calculate $M(t,f)$, we can use recursion, just as we did in the case of $F(t)$.

If $f \geq t-2$, then condition 2′ is no restriction at all, since condition 1 does not permit more than f failures in any case. Therefore, every pattern of failures is allowed, as long as the first and last stations do not fail. In this case, $M(1,f)=1$ while for $1<t\leq f+2$, we have $M(t,f)=2^{t-2}$. If, on the other hand, $f<t-2$, then we can partition the ways in which some or no stations can fail subject to conditions 1 and 2′ into $f+1$ disjoint subsets, as follows. The first subset consists of all cases where station number $t-2$ has not failed. The next subset consists of all cases where station number $t-2$ has failed but station $t-3$ has not. The next subset consists of all cases where station $t-2$ and station $t-3$ have failed but $t-4$ has not. In general, cases are grouped together into a subset according to the number of consecutive stations that have failed, counting backward from station $t-2$. The last subset consists of all cases where the maximum number, f, consecutive stations have failed, counting backward from station $t-2$. Thus, in the case where $f<t-2$, we have $M(t,f)=M(t-1,f)+M(t-2,f)+\cdots+M(t-f-1,f)$. This expresses $M(t,f)$ as a sum of $f+1$ terms. In most applications, f, the number of consecutive failed stations that can be bypassed, is fixed and is small relative to the size of t, so this is a convenient expression for calculating $M(t,f)$ recursively, using the observation at the beginning of this paragraph to get us started. The values of $M(t,f)$, $t=1,2,3,\ldots$, form a recurrent sequence of order $f+1$, where the first $f+1$ terms are $1,1,2,4,8,\ldots,2^f$ and where each term thereafter is the sum of the preceding $f+1$ terms.

For example, if $f=3$, then $M(1,3)=1$, $M(2,3)=1$, $M(3,3)=2$, $M(4,3)=4$, and $M(5,3)=8$. This takes care of all of the cases with $f\leq t-2$. Since this gives us $f+2$ terms, we always have enough terms to begin using the recursion formula for $M(t,f)$, $f>t-2$. For example, $M(6,3)=8+4+2+1=15$ and $M(7,3)=15+8+4+2=29$.

When $f=1$, the values of $M(t,1)$ form the Fibonacci sequence. A table of values for $M(t,f)$ for $1\leq t\leq 10$, $1\leq f\leq 8$, is given in Table 1.

TABLE 1. $M(t,f)$
The number of different patterns in which some or none of the stations fail, in which the first and last stations do not fail, and in which no string of more than f consecutive stations contains all failures, in a linear array of t stations.

$t=$	1	2	3	4	5	6	7	8	9	10
$f=1$	1	1	2	3	5	8	13	21	34	55
2	1	1	2	4	7	13	24	44	81	149
3	1	1	2	4	8	15	29	56	108	208
4	1	1	2	4	8	16	31	61	120	236
5	1	1	2	4	8	16	32	63	125	248
6	1	1	2	4	8	16	32	64	127	253
7	1	1	2	4	8	16	32	64	128	255
8	1	1	2	4	8	16	32	64	128	256

Stations in a Ring with Better Bypass Circuits

Finally, we consider a dual ring with n stations and bypass circuits, in which the bypass circuits can bypass f or fewer failed stations. Let $H(n,f)$ represent the number of different patterns in which some or none of the stations fail, while still leaving the ring connected.

First, we count patterns that contain one sequence of more than f but less than n consecutive failed stations. Because of the dual nature of the ring, one such sequence does not disconnect the ring, while two or more such sequences would disconnect the ring. Let k represent the number of stations in this longest sequence of consecutive failed stations, $f<k<n$. Such a sequence can start at any station in the ring, so there are n positions in which such a sequence can occur. For each one of these n positions, the number of patterns in which some or none of the remaining $n-k$ stations can fail, subject to conditions 1 and 2′ of the previous section of this paper, is $M(n-k,f)$. We therefore multiply n times $M(n-k,f)$. To count all of the patterns that contain one sequence of more than f but less than n consecutive failed stations, we sum this number, $nM(n-k,f)$, as k goes from $f+1$ to $n-1$.

To this sum we must add the number of ways we can have some or no failures, but never have more than f consecutive stations all fail. Number the stations in the ring 1 through n (so n and 1 are consecutive). Represent the number of failed stations in the initial string of failed stations by i, where $0 \leq i \leq f$. For example, if stations number 1, 2, and 3 fail, while station number 4 does not fail, then $i=3$. Represent the number of failed stations in the terminal string of failed stations by j, where $0 \leq j \leq f$. For example, if stations numbered n, $n-1$, $n-2$, and $n-3$ fail, while station numbered $n-4$ does not fail, then $j=4$. Since the first station and the last station in a ring are consecutive, it follows from the condition that no string of more than f stations all fail that $i+j \leq f$. Let $z=i+j$. There are $z+1$ possibilities that can yield the same value for $i+j$: $i=0$, $i=1,\ldots,i=z$. In each of these $z+1$ cases, there are $M(n-z,f)$ ways in which some or none of the other stations, which are not part of the initial or terminal strings of failures, can fail, subject to conditions 1 and 2′. Therefore we need to multiply $M(n-z,f)$ by $i+j+1=z+1$. Thus the number of ways in which we can have some or no failures, but in which no string of more than f consecutive stations all fail, is the sum as z goes from 0 to f of $(z+1)M(n-z,f)$.

Putting these two pieces together, we have

$$H(n,f) = \sum_{z=0}^{f} (z+1)M(n-z,f) + n \sum_{k=f+1}^{n-1} M(n-k,f).$$

A table of values for $H(n,f)$, for $1 \le n \le 10$, $1 \le f \le 6$, is given in Table 2.

TABLE 2. $H(n,f)$
The number of different patterns in which some or none of the stations fail, while leaving the ring connected, in a dual ring with n stations, in which the bypass circuits can bypass f or fewer failed stations.

$n =$	1	2	3	4	5
$f = 1$	1	3	7	15	31
2	1	3	7	15	31
3	1	3	7	15	31
4	1	3	7	15	31
5	1	3	7	15	31
6	1	3	7	15	31

$n =$	6	7	8	9	10
$f = 1$	60	113	207	373	663
2	63	127	241	493	963
3	63	127	255	511	1018
4	63	127	255	511	1023
5	63	127	255	511	1023
6	63	127	255	511	1023

It is interesting to note that, even though the terms in the series evaluating $H(n,f)$ are different, the sums of those terms are the same, for different values of f, provided $2f+2 \ge n$, since we need at least $2f+2$ stations to have two strings of consecutive failed stations with f stations in each string. In other words, once $2f+2 \ge n$, increasing f will not increase the number of ways in which some or no stations can fail and still leave the ring connected. Therefore, if $2f_1 + 2 \ge n$ and $2f_2 + 2 \ge n$, then $H(n,f_1) = H(n,f_2)$.

We began with a real world problem, designing a fiber optic network. We isolated one small part of that problem, counting failure patterns in a particular configuration of that network. In solving that part of the problem we discovered some interesting combinatorial mathematics. We should note that this is just one piece of the puzzle. The solution of the larger problem, the design and installation of a token ring network in NASA Mission Control Center, consisted of a great many such pieces.

REFERENCES

1. Richard Bollinger and Charles Burchard, Lucas's theorem and some related results for extended Pascal triangles, *Amer. Math. Monthly*, 97 (1990) 198–204.
2. C. A. Church, Jr., Lattice paths and Fibonacci and Lucas numbers, *Fibonacci Quarterly* 12(4) (1974) 336–338.
3. William Stallings, *Handbook of Computer Communications Standards*, Vol. 2, Howard W. Sams & Company, Indianapolis, Indiana, 1988.
4. Werner Bux, Token-ring local-area networks and their performance, *Proceedings of the IEEE*, Vol. 77, No. 2, Feb. 1989.
5. Marjory J. Johnson, Reliability mechanisms of the FDDI high bandwidth token ring protocol, *Computer Networks and ISDN Systems*, Vol. 11, February 1986.

NOTES

There Are Three Times as Many Obtuse-Angled Triangles as There Are Acute-Angled Ones

RICHARD K. GUY
The University of Calgary
Calgary, Alberta, Canada T2N 1N4

At least! In fact almost all triangles are obtuse-angled, as you will see if you draw one side, AB (FIGURE 1) and then decide where to put the third vertex, C. Unless you put it in the shaded region, which constitutes a negligible fraction of the whole plane, you'll get an obtuse triangle. How can we get a more realistic estimate?

If AB is to be the shortest side, then C must be outside both the circles in FIGURE 2, and we get a similar result.

If AB is to be the middle side, then C must be outside one of the circles in FIGURE 2 and inside the other. The fraction of the area giving an obtuse-angled triangle is

$$\frac{\pi}{2\pi/3+\sqrt{3}} \quad \text{or} \quad 82.1\%.$$

If AB is to be the longest side, then C must be inside both circles in FIGURE 2, and the fraction is

$$\frac{\pi/4}{2\pi/3-\sqrt{3}/2} \quad \text{or} \quad 63.9\%.$$

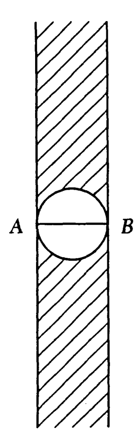


FIGURE 1
Strip in which C forms an acute-angled triangle.

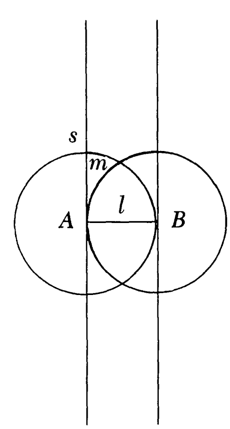


FIGURE 2
Where C makes AB the shortest (s), middle (m), or longest (l) side.

In practice, we have only a finite region to work in, and we want our triangle to be visible to the naked eye, so a natural region to which to restrict C might be an ellipse with foci A and B (FIGURE 3).

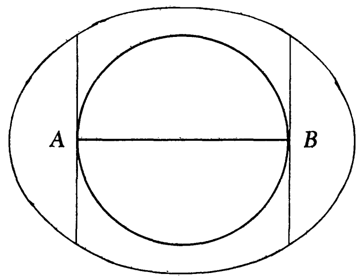


FIGURE 3
Triangles confined to a golden ellipse.

The proportion of obtuse-angled triangles with C in such a region is easily seen to be

$$1+\frac{e^2}{\sqrt{1-e^2}}-\frac{2}{\pi}\left(\arcsin e+e\sqrt{1-e^2}\right)$$

where e is the eccentricity of the ellipse. As a founding member of SECT[1] let me reassure you that the formula was found by orthogonally projecting a circle and not by using calculus. Of course, if we take the eccentricity to be very small, or very large, then we get an unduly large proportion of obtuse-angled triangles.

But with reasonable numbers, such as $e = 0.618\dots$, the golden ratio (depicted in FIGURE 3), or $e=\frac{1}{4}$, we get the reasonable proportions of 75.24% or 74.96% respectively. However you choose the eccentricity, there'll be more than twice as many obtuse triangles as acute ones.

But most of us work on rectangular sheets of paper. To make the picture "fairly fill the paper" as they used to say in the art exams, draw a $2\frac{1}{2}$-in. line in the middle of an 11-in. $\times$ $8\frac{1}{2}$-in. sheet of paper, and the triangle will be obtuse-angled

$$\frac{11\times\left(8\frac{1}{2}-2\frac{1}{2}\right)+\pi\left(\frac{5}{4}\right)^2}{11\times 8\frac{1}{2}}=75.83\%$$

of the time. Or, a 6-cm. line in the middle of an A4 sheet (and it's high time we caught up with the rest of the world) yields an obtuse triangle 75.98% of the time.

Another way would be to break a stick, say of unit length, into three pieces, lengths $x+y+z=1$, and try to make a triangle. Three-quarters of the time it doesn't work, because the pieces must satisfy the triangle inequalities $x+y>z$, etc. This is clear from FIGURE 4, which is an equilateral triangle of unit height.

A point at distances x, y, z from its sides represents a set of numbers with $x+y+z=1$. Since x, y, z are positive, we are inside the triangle, and unless we are in the quarter-sized triangle in the middle, one of the triangle inequalities will be violated.

But even when we are inside the middle triangle, the angle opposite z will be obtuse if $x^2+y^2<z^2=(1-x-y)^2$, i.e., if the point is inside the branch of the

[1]The Society for the Elimination of Calculus Teaching, *not*, as an undergraduate recently suggested, the Society for the Extermination of Calculus Teachers.

hyperbola $(1-x)(1-y) > \frac{1}{2}$. The chance that the point is inside one of the three hyperbolas is $9 - 12\ln 2$ or 68.22%.

"But all this is mere experimental nonsense," I hear you cry, "what about some real mathematics?" Very well, here are no fewer than five (count them) separate genuine proofs.

Proof 1. Every triangle has an **orthocentre**, the common point of the altitudes, that is outside the triangle just if the triangle is obtuse-angled. In any case the vertices and the orthocentre form a set of four **orthocentric points**, each of which is the orthocentre of the other three. Any conic passing through them is a rectangular hyperbola, the locus of whose centres is the nine-point circle..., but I digress! Observe that any three points uniquely determine a fourth and just three of the four triangles thus formed are obtuse-angled.

Proof 2. Every triangle has a **circumcircle**. So choose three points at random on a circle. It's convenient to think of the circumference of the circle as being bent from the stick of unit length that we earlier broke into three pieces. The new interpretation of FIGURE 4 is that x, y, z are *arcs of the circumcircle* instead of straight lines. A triangle is *always* determined by a point inside the big triangle: The sides are now the associated *chords*. But we know that 3/4 of the time the lengths of the *arcs* don't satisfy the triangle inequality: The three vertices all lie in one half of the circumcircle and the triangle is obtuse-angled.

Proof 3. Every triangle has an **incircle**. So, as in Proof 2, we choose three points randomly on a circle, but this time draw tangents there to form the triangle. But the points must not all lie in the same semicircle, else we would have an **excircle** instead of an incircle. So we are in a similar situation to that of FIGURE 4: We confine our attention to the middle quarter of FIGURE 5.

Our triangle will be obtuse-angled just if two of the points of tangency lie in the same quadrant of the circle, i.e. just if one of x, y, or $z < \frac{1}{4}$. These three inequalities are represented by three of the four small triangles in the middle of FIGURE 5.

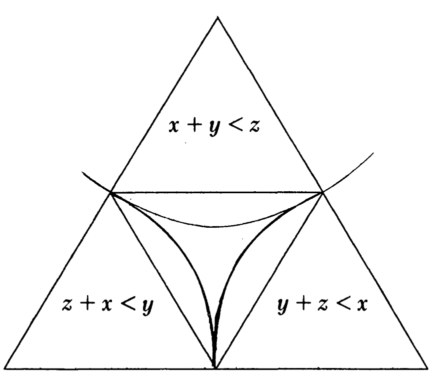


FIGURE 4
If we break a stick into three parts, x, y, z, then three-quarters of the time we can't make a triangle.

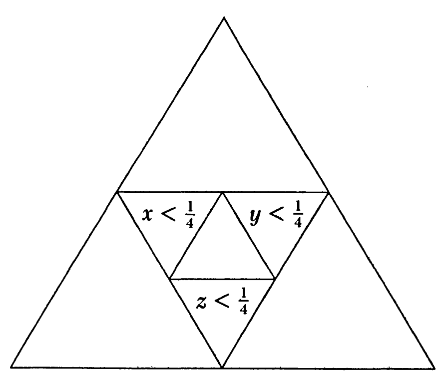


FIGURE 5
If any of the three arcs of the incircle is less than a quarter, the tangents form an obtuse angle.

Proof 4. Throw down three long straight poles so that they form a triangle. The second pole crosses the first at A as in FIGURE 6. Of the two supplementary angles at A, call the smaller one θ.

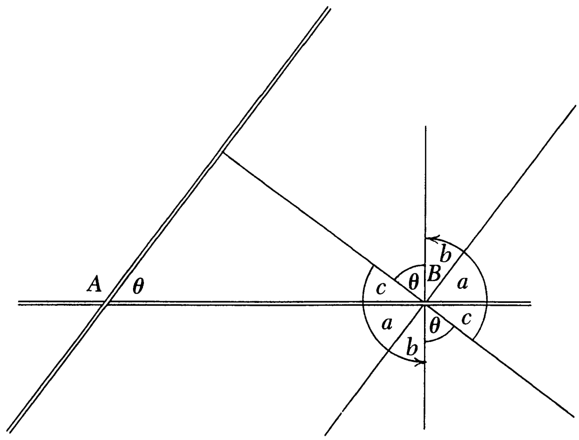


FIGURE 6
Three poles form a random triangle.

The third pole crosses the first at B (if B falls on the other side of A, rotate the figure through 180 degrees). If the third pole lies in the sectors indicated by the circular arrows, the triangle will be obtuse-angled (at A, B, or C according as the pole lies in the sectors a, b, or c). The fraction of obtuse-angled triangles is

$$\frac{\pi-\theta}{\pi}.$$

We must average this over the interval $0<\theta<\pi/2$, i.e., $1-\theta/\pi$ averaged over $0<\theta/\pi<\frac{1}{2}$; a linear function that runs from 1 to $\frac{1}{2}$; average $\frac{3}{4}$.

Proof 5. Every triangle has a largest angle. It lies in the range from 60 degrees to 180 degrees. For $\frac{3}{4}$ of this range the triangle is obtuse-angled. This proof may be presented in another way. Consider the largest of the three excircles of the triangle. The minor arc of this, between the points of tangency with the arms of the largest angle of the triangle, has length at most 120°. If its length is less than 90°, the largest angle is obtuse, and this occurs $\frac{3}{4}$ of the time.

In five well-known geometry books I found figures depicting 146 triangles that purported to be "general" triangles, but only 36 of them were obtuse (often in illustrations of Desargues's theorem, or inversion, or multiplication of complex numbers), almost exactly the reverse proportion of what we *now* know to be the correct one. You will no doubt immediately join me in supporting the GROAT[2] cause.

Historical Note A good deal of the content of the paper must be scattered about in the literature. Singmaster's *Sources*, under the head "Probability that three lengths form a triangle", traces the "breaking the stick into three pieces" problem at least as far back as Lemoine, and states that there were several later articles based on Lemoine, but neither he nor I have actually seen any of these. He also refers to Fourrey [2], who gives the answer $P=1/4$ and cites Lemoine [3].

Under the head "Probability that a triangle is acute" he quotes Sylvester and Lewis Carroll. Again he says he hasn't seen these, but gives Miles and Serra [4] as his source for Sylvester. I've located the Sylvester reference [5], a discursive article, with no results. The problem of finding "the chance of three points within a circle or

[2]Greater rights for obtuse-angled triangles

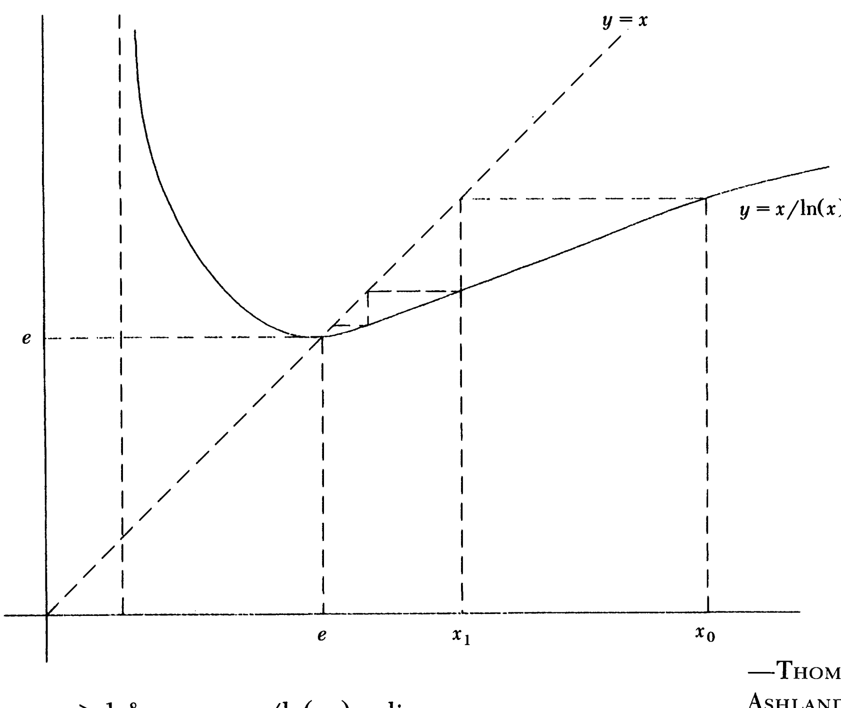
sphere being apices of an acute or obtuse-angled triangle" is attributed to Woolhouse, and his conclusion is that "the form in which ... problems ... originally proposed... without a specified boundary,... do not admit of a determinate solution." Lewis Carroll [1] states the problem in the form: "Three Points are taken at random on an infinite Plane. Find the chance of their being the vertices of an obtuse-angled Triangle." His solution is essentially that given above under the assumption that *AB* is the longest side, with the answer

$$\frac{3}{8 - \frac{6\sqrt{3}}{\pi}}.$$

REFERENCES

1. Lewis Carroll, *Pillow Problems*, 1893. 4th edition (1895) Problem 58, pp. 14, 25, 83–84; reprinted (with A Tangled Tale), Dover Publications, New York, 1958.
2. E. Fourrey, *Curiosités Géometriques*, (1st edition Vuibert and Nony, Paris 1907) 4th edition, Vuibert, Paris 1938. Part 3, chap. 1 §5: Application au calcul des probabilités, pp. 360–362. [Break a stick into three pieces.]
3. E. Lemoine, *Bull. Soc. Math. France*, 1872–1873.
4. R. E. Miles and J. Serra, En matière d'introduction, in *Geometrical Probability and Biological Structures: Buffon's 200th Anniversary*, Springer Lecture Notes in Biomathematics, 23(1978) 3–28 (esp. p. 18).
5. J. J. Sylvester, On a special class of questions in the theory of probabilities, *Birmingham British Association Report* (1865), p. 8. Vol. 2 of *The Collected Mathematical Papers of James Joseph Sylvester*, Cambridge, 1908; Chelsea reprint, 1973, **75**. pp. 480–481.

Proof Without Words



$x_0 > 1$ & $x_{n+1} = x_n/\ln(x_n) \rightarrow \lim x_n = e$

—THOMAS P. DENCE
ASHLAND UNIVERSITY
ASHLAND, OH 44805

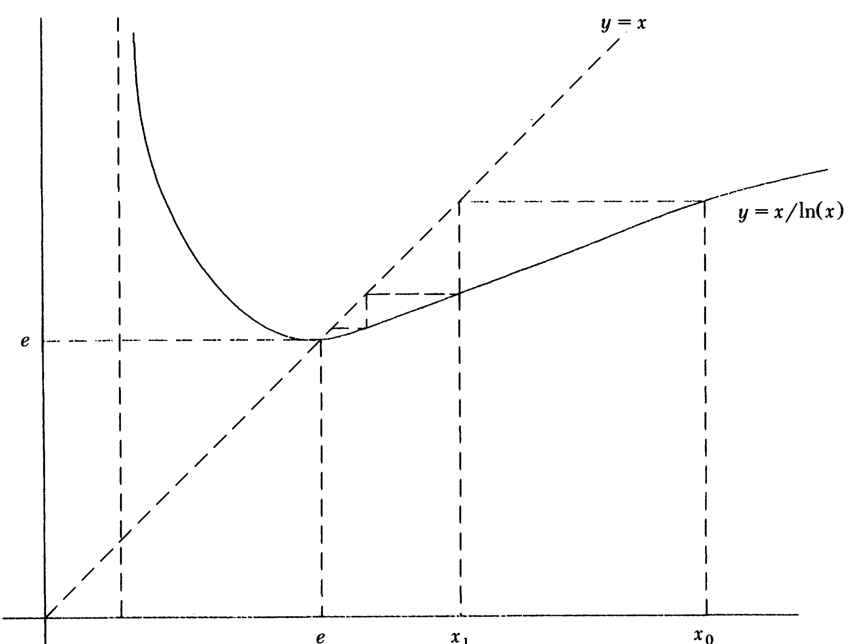
sphere being apices of an acute or obtuse-angled triangle" is attributed to Woolhouse, and his conclusion is that "the form in which . . . problems . . . originally proposed . . . without a specified boundary, . . . do not admit of a determinate solution." Lewis Carroll [1] states the problem in the form: "Three Points are taken at random on an infinite Plane. Find the chance of their being the vertices of an obtuse-angled Triangle." His solution is essentially that given above under the assumption that AB is the longest side, with the answer

$$\frac{3}{8 - \frac{6\sqrt{3}}{\pi}}.$$

REFERENCES

1. Lewis Carroll, *Pillow Problems*, 1893. 4th edition (1895) Problem 58, pp. 14, 25, 83–84; reprinted (with A Tangled Tale), Dover Publications, New York, 1958.
2. E. Fourrey, *Curiosités Géometriques*, (1st edition Vuibert and Nony, Paris 1907) 4th edition, Vuibert, Paris 1938. Part 3, chap. 1 §5: Application au calcul des probabilités, pp. 360–362. [Break a stick into three pieces.]
3. E. Lemoine, *Bull. Soc. Math. France*, 1872–1873.
4. R. E. Miles and J. Serra, En matière d'introduction, in *Geometrical Probability and Biological Structures: Buffon's 200th Anniversary*, Springer Lecture Notes in Biomathematics, **23**(1978) 3–28 (esp. p. 18).
5. J. J. Sylvester, On a special class of questions in the theory of probabilities, *Birmingham British Association Report* (1865), p. 8. Vol. 2 of *The Collected Mathematical Papers of James Joseph Sylvester*, Cambridge, 1908; Chelsea reprint, 1973, **75**. pp. 480–481.

Proof Without Words

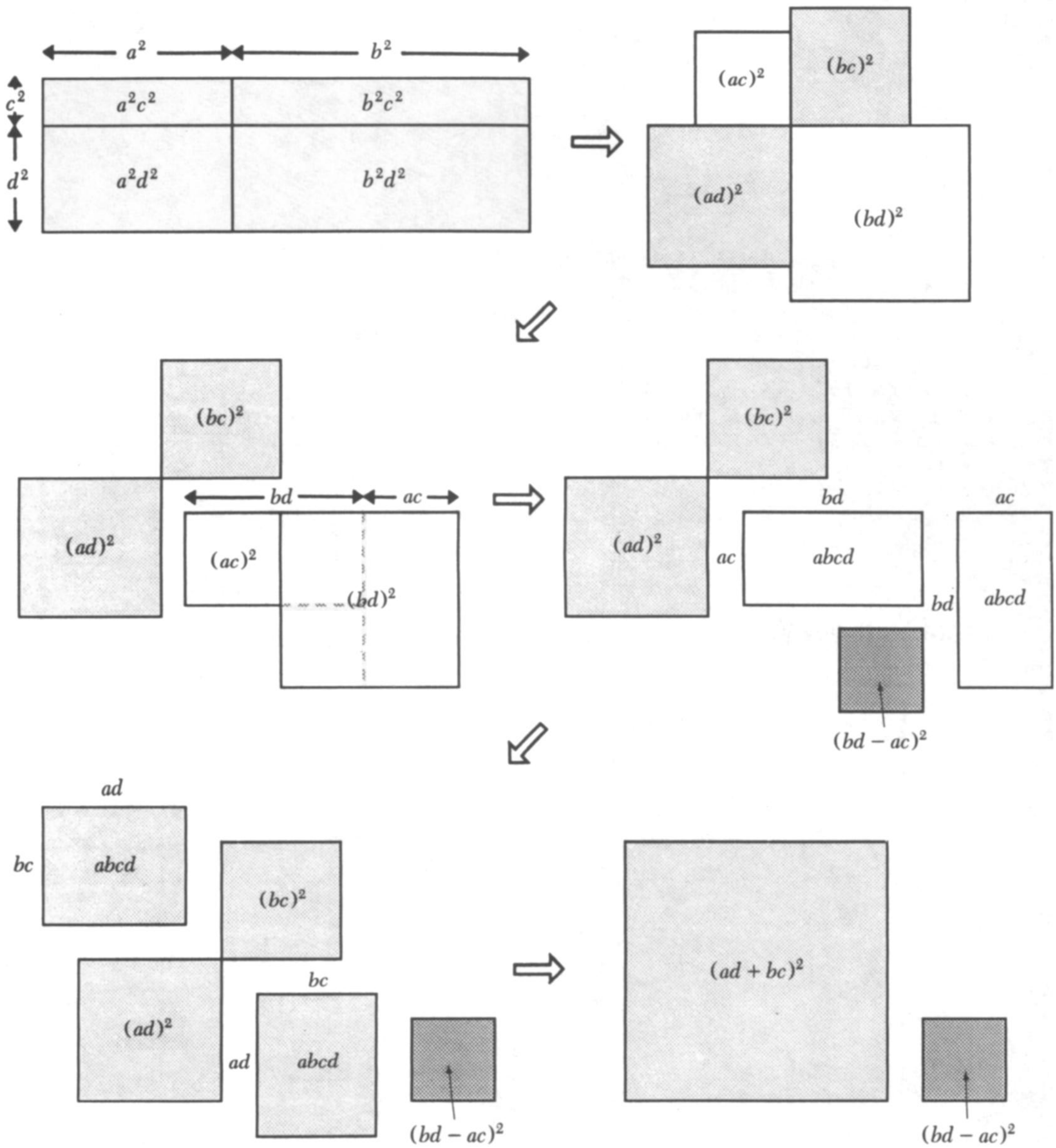


$x_0 > 1 \ \& \ x_{n+1} = x_n/\ln(x_n) \to \lim x_n = e$

—THOMAS P. DENCE
ASHLAND UNIVERSITY
ASHLAND, OH 44805

Proof without Words: Diophantus of Alexandria's "Sum of Squares" Identity

$$(a^2+b^2)(c^2+d^2)=(ad+bc)^2+(bd-ac)^2$$



—Roger B. Nelsen
Lewis and Clark College
Portland, OR 97219

Euler Numbers and Skew-Hooks

RICHARD M. GRASSL
Muhlenberg College
Allentown, PA 18104

The *Euler number* E_n is the number of permutations of the integers $1, 2, 3, \ldots, n$ that first rise, then alternately fall and rise. For $n = 4$ the permutations of $1, 2, 3, 4$ satisfying this property are 1324, 1423, 2314, 2413, and 3412. So we see that $E_4 = 5$. If one throws in the term $E_0 = 1$, the first eleven Euler numbers $1, 1, 1, 2, 5, 16, 61, 272, 1385, 7936, 50521$ are given as the beginning terms of sequence #587 in the N. J. A. Sloane *Handbook of Integer Sequences* [9].

Such permutations, sometimes called "up-down" permutations, were first studied in 1879 by D. André [2]; they make extensive appearances in the literature, showing up in a variety of contexts. One frequently used recursive technique for finding E_n involves the recurrence relation $E_{n+1} = \frac{1}{2}\sum_{i=0}^{n}\binom{n}{i}E_iE_{n-i}$ along with the initial conditions $E_0 = 1$, $E_1 = 1$. (The reader is invited to try to prove this recurrence.) One disadvantage in using this scheme, of course, is that in computing E_{n+1}, one needs to know all the values $E_0, E_1, E_2, \ldots, E_n$. In this paper, we first highlight several settings in which the Euler numbers play prominent roles and then give a method for finding a relatively simple, nonrecursive formula for E_n, employing a technique that is accessible to most undergraduates.

Appearances The even-subscripted terms of the sequence $\{E_i\}$ of Euler numbers appear in the Maclaurin series expansion for $\sec x$:

$$\sec x = 1 + \frac{x^2}{2!} + 5\frac{x^4}{4!} + 61\frac{x^6}{6!} + \cdots = E_0 + E_2\frac{x^2}{2!} + E_4\frac{x^4}{4!} + E_6\frac{x^6}{6!} + \cdots$$

The numbers E_{2n} are often called the Euler numbers, or the secant numbers. If one rewrites the usual Maclaurin series for $\tan x$, namely,

$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \cdots$$

as

$$\tan x = x + 2\frac{x^3}{3!} + 16\frac{x^5}{5!} + 272\frac{x^7}{7!} + \cdots = E_1 + E_3\frac{x^3}{3!} + E_5\frac{x^5}{5!} + E_7\frac{x^7}{7!} + \cdots$$

it becomes clear why the numbers E_{2n+1} are called the *tangent numbers*. E. Netto [8] expressed these relationships in the form:

$$\sec x + \tan x = \sum_{n=0}^{\infty} E_n\frac{x^n}{n!}.$$

André wrote the expansion in the form

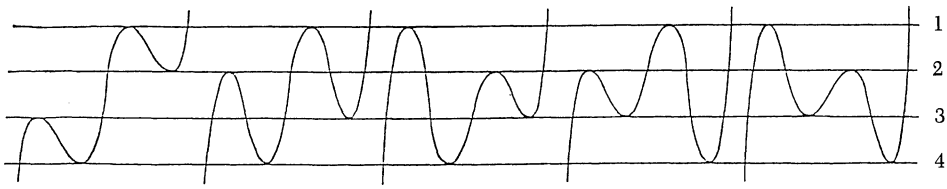
$$\tan\left(\frac{x}{2} + \frac{\pi}{4}\right) = \sum_{n=0}^{\infty} E_n\frac{x^n}{n!}.$$

In [3], L. Carlitz notes that the hyperbolic secant gets into the act:

$$\operatorname{sech} x = \frac{2}{e^x + e^{-x}} = \sum_{n=0}^{\infty} E_{2n}(-1)^n \frac{x^{2n}}{(2n)!}.$$

Every beginning calculus student investigates the shape of polynomial curves in determining relative maximum and minimum points. A. Kempner [7] showed in 1933 that the order in which maxima and minima may be distributed is restricted only by the fact that such points must alternate.

Suppose $p(x)$ is a polynomial of degree five such that its derivative has its four roots x_1, x_2, x_3, x_4 all real and distinct. Agreeing that $x_1 > x_2 > x_3 > x_4$ and that the leading coefficient of $p(x)$ is positive, in how many different ways may we prescribe the order in which the distinct extremes $p(x_1), p(x_2), p(x_3), p(x_4)$ can be arranged? The following listing of graphs reveals that there are $E_4 = 5$ ways. Each graph, scanning left to right, begins with a maximum. A *type*, indicating the order of magnitudes of the extrema, is assigned to each graph. The left-most graph below is of type (3412), since, reading from left to right, that graph has a maximum at 3, followed by a minimum at 4, then a maximum at 1, and finally a minimum at 2.



(3412) (2413) (1423) (2314) (1324)

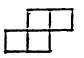
As the following table shows (verified by Kempner), the number of types is an Euler number.

n = degree of $p(x)$	2	3	4	5	6	7	8	$\cdots$
number of types	1	1	2	5	16	61	272	$\cdots$
E_n	1	2	5	16	61	272	$\cdots$	

The Problem The goal in this paper is to give a reasonably nice, nonrecursive formula for the Euler number E_n. Toward this goal, $E_4 = 5$ is looked at first.

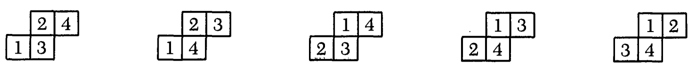
Recalling that the five "up-down" permutations of $1, 2, 3, 4$ are:

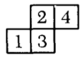
1324 1423 2314 2413 3412

we notice that each such permutation can be associated with a particular arrangement of the integers $1, 2, 3, 4$ in the shape  in which the entries are increasing in rows (reading left to right) and columns (top to bottom); each horizontal row corresponds to a "rise" and each vertical column corresponds to a "fall." In the remainder of this paper we shall refer to this increasing property by saying that the entries are *monotonic in rows and columns*. The following displays this association:

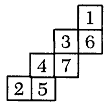
1234 1423 2314 2413 3412

$\updownarrow$ $\updownarrow$ $\updownarrow$ $\updownarrow$ $\updownarrow$



The arrangement 

, for example, corresponds with 1324 by reading across rows left to right, from bottom to top. Similarly for $n = 7$, the arrangement



corresponds to the "up-down" sequence 2547361.

With this in mind, our general problem can be rephrased: In how many ways can one insert the integers $1, 2, 3, \ldots, n$ into the shape

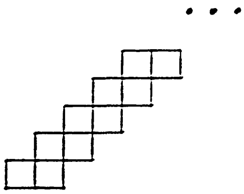


FIGURE 1

consisting of n boxes so that the entries are monotonic in rows and columns? We call such a shape a *skew-hook*. Foulkes [5] was one of the first to link "up-down" permutations with skew-hooks. His approach to the enumeration involved the concepts of Schur functions and irreducible representations of the symmetric group.

The Box Algebra Here we explore the basics of the enumeration technique that yields a formula for E_n in terms of multinomial coefficients. The cases $n = 3$, 4, 5 and 6 are examined closely. Our approach is to express the skew-hooks in FIGURE 1 as a linear combination of simpler shapes consisting of disjoint rows, and enumerate these latter shapes.

We call a shape consisting of disjoint rows such as

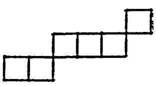
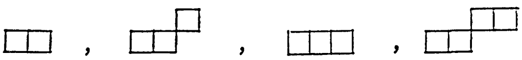
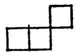
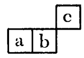
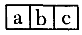


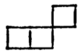
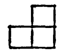
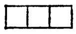
FIGURE 2

a *row product*. There is no overlapping of rows in a row product and no restriction on either the row lengths or the number of rows. Further examples of row products are

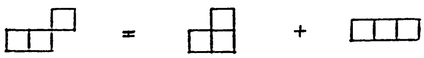


In how many ways can one insert the integers $1, 2, 3, 4, 5, 6$ into the row product in FIGURE 2 maintaining monotonicity in the rows (there are no columns)? The answer is $\binom{6}{2}\binom{4}{3}\binom{1}{1}$–choose 2 of the 6 integers in $\binom{6}{2}$ ways and place them in the bottom row (monotonicity is forced!), then choose 3 of the remaining 4 for the middle row in $\binom{4}{3}$ ways; the remaining integer must be placed into the top row, which consists of just one box.

A simple *two-step procedure* links row products as in FIGURE 2 with the skew-hooks in FIGURE 1. To illustrate, label, for convenience, the shape  as  where $\{a, b, c\} = \{1, 2, 3\}$. Now compare b and c. If c is smaller than b the box containing c could be slid over on top of b, maintaining monotonicity. If c is larger than b, drop it down and in the resultant configuration  we have monotonicity. The statement:

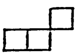
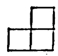
"The number of ways of inserting 1, 2, 3 into , maintaining monotonicity (in rows), is equal to the number of ways of inserting 1, 2, 3 into  plus the number of ways of inserting 1, 2, 3 into , maintaining monotonicity."

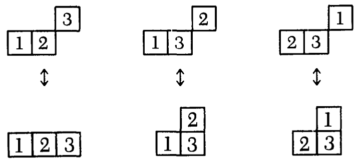
could be expressed symbolically as

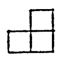


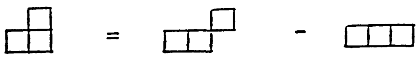
(slide left or drop down)

FIGURE 3

To illustrate this statement, take each of the $\binom{3}{2}\binom{1}{1} = 3$ ways of inserting 1, 2, 3 into  while maintaining monotonicity (in rows) and assign them to either  or  using the sliding-dropping process. The following matching results:

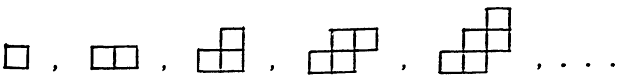


The Euler number E_3 is the number of ways of inserting 1, 2, 3 into  maintaining monotonicity in rows and columns. "Solve" for the skew-hook in FIGURE 3, arriving at

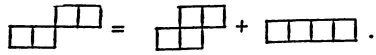


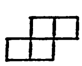
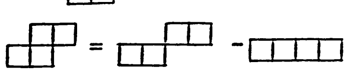
From this box algebra equation we obtain $E_3 = \binom{3}{2}\binom{1}{1} - \binom{3}{3} = 2$.

In the process of finding an equation involving one of the skew-hooks



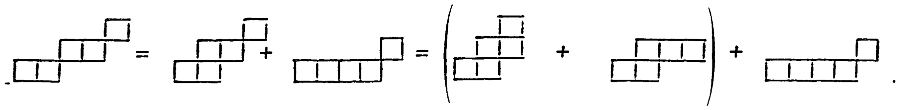
while it is possible to start with a variety of shapes, it is most efficient to start with shapes consisting of disjoint rows of length one or two.

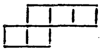
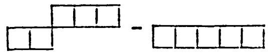
The case $n = 4$ is .

Solving for  gives ,

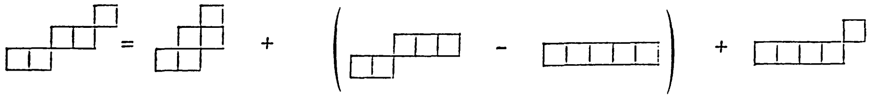
and we obtain $E_4 = \binom{4}{2}\binom{2}{2} - \binom{4}{4} = 5$.

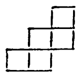
The case $n = 5$:

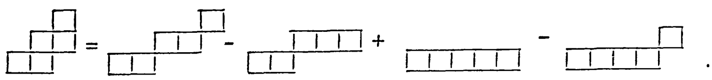


Here, due to the more complex shapes involved, a little trouble surfaces. What should be done with the term  ? Simply rewrite it as 

Then

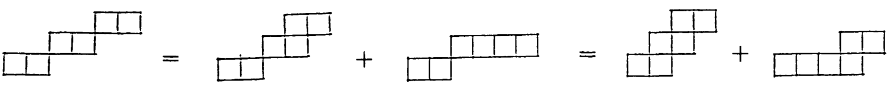


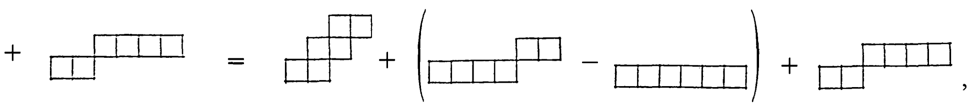
and we now solve for our unknown  , obtaining



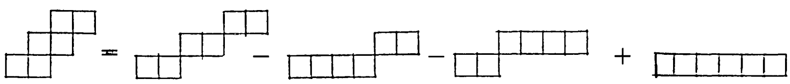
The corresponding count is $E_5 = \binom{5}{2}\binom{3}{2}\binom{1}{1} - \binom{5}{2}\binom{3}{3} + \binom{5}{5} - \binom{5}{4}\binom{1}{1} = 16.$

The case $n = 6$: From





we obtain



and $E_6 = \binom{6}{2}\binom{4}{2}\binom{2}{2} - \binom{6}{4}\binom{2}{2} - \binom{6}{2}\binom{4}{4} + \binom{6}{6} = 61.$

Observing the Pattern The box algebra has done its work. It has provided us with sufficient data from which a general pattern can be perceived. Rewriting products of binomial coefficients as the multinomial coefficient $\binom{n}{a,b,c,\dots,k} = n!/a!b!c!\cdots k!$, the first few E_n look like:

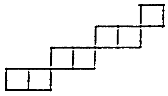
$$E_1 = \binom{1}{1} \qquad E_2 = \binom{2}{2}$$

$$E_3 = \binom{3}{1,2} - \binom{3}{3} \qquad E_4 = \binom{4}{2,2} - \binom{4}{4}$$

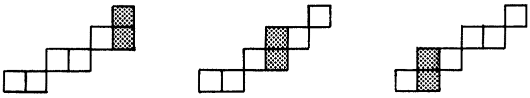
$$E_5 = \binom{5}{1,2,2} - \binom{5}{2,3} - \binom{5}{1,4} + \binom{5}{5} \qquad E_6 = \binom{6}{2,2,2} - 2\binom{6}{2,4} + \binom{6}{6}.$$

These expressions suggest that the principle of inclusion-exclusion could be used for determining a formula for E_n as a linear combination of multinomial coefficients. A detailed look at the evaluation of E_7 will tell us precisely how to proceed for any n.

For $n = 7$ our previous results suggest that we start with the shape



There are three places where monotonicity in columns could fail. These failure locations are indicated by the shading:



But each of these shapes could then be redrawn, respectively, as

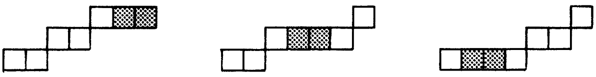
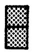
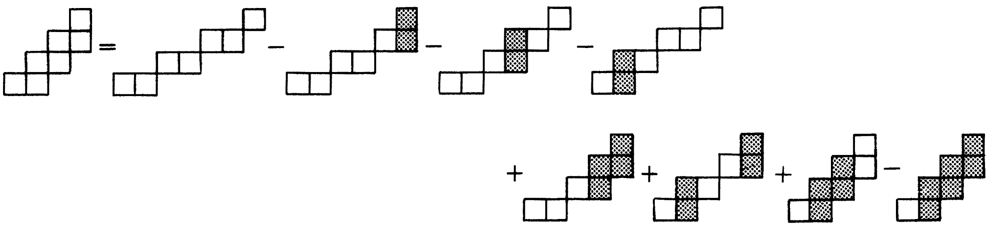


FIGURE 4

since in each shaded column  the entry on top is larger than that below.

The arrangements problem for these latter three, being row products, is easily resolved.

If one deletes, using inclusion-exclusion, all the cases where monotonicity fails, the following relationship is obtained:



Redrawing the shaded shapes to look like those in FIGURE 4, we obtain

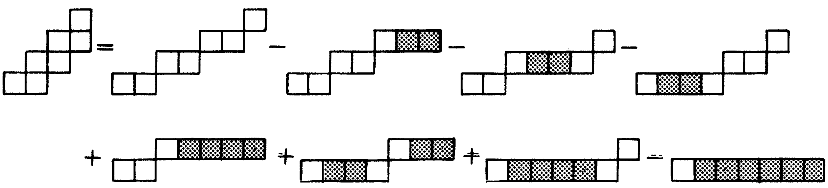


FIGURE 5

From FIGURE 5 we deduce the count

$$E_7 = \binom{7}{1,2,2,2} - \binom{7}{3,2,2} - 2\binom{7}{1,2,4} + \binom{7}{5,2} + \binom{7}{3,4} + \binom{7}{1,6} - \binom{7}{7}.$$

In a similar manner we derive

$$E_8 = \binom{8}{2,2,2,2} - 3\binom{8}{2,2,4} + 2\binom{8}{2,6} + \binom{8}{4,4} - \binom{8}{8}.$$

The General Case Let $[a_1, a_2, \dots, a_n]$ denote a skew-hook shape as in FIGURE 1 having row lengths $a_1, a_2, \dots, a_n$ reading top to bottom and let the "product" $[a_1][a_2]\cdots[a_n]$ denote the shape consisting of n disjoint rows having lengths $a_1, a_2, \dots a_n$. Then the box algebra identity in FIGURE 5 can be expressed as

$$\begin{aligned}[1,2,2,2] &= [1][2][2][2] - [1+2][2][2] - [1][2+2][2] - [1][2][2+2] + [1+2+2][2] \\ &\quad + [1+2][2+2] + [1][2+2+2] - [1+2+2+2] \\ &= [1][2][2][2] - [2][2][3] - 2[1][2][4] + [2][5] + [3][4] + [1][6] - [7]\end{aligned}$$

FIGURE 6

An algorithm for determining this decomposition is described in the following. Consider the row product [1][2][2][2]. Observe here that replacing one of the three pairs of opposing brackets][in [1][2][2][2] by a + sign (this can be done in $\binom{3}{1}$ ways) produces the three terms

$$[1+2][2][2], \qquad [1][2+2][2], \qquad [1][2][2+2],$$

each being a product of three rows. Replacing two such bracket pairs by two + signs in $\binom{3}{2}$ ways produces the three terms [1 + 2 + 2][2], [1 + 2][2 + 2], [1][2 + 2 + 2], each a product of two rows. With three + signs, just one product consisting of one row can be formed, namely [1 + 2 + 2 + 2].

As the principle of inclusion-exclusion demands, we alternate signs in forming the linear combination in FIGURE 6. Assign a plus to products of four rows, a minus to products of three rows, a plus to products of two rows, and finally a minus to the product consisting of one row.

This algorithm easily yields

$$[2,2,2,2] = [2][2][2][2] - 3[2][2][4] + 2[2][6] + [4][4] - [8]$$

from which follows

$$E_8 = \binom{8}{2,2,2,2} - 3\binom{8}{2,2,4} + 2\binom{8}{2,6} + \binom{8}{4,4} - \binom{8}{8}.$$

Similarly,

$$\begin{aligned}E_{10} &= \binom{10}{2,2,2,2,2} - 4\binom{10}{2,2,2,4} + 3\binom{10}{2,4,4} + 3\binom{10}{2,2,6} \\ &\quad - 2\binom{10}{2,8} - 2\binom{10}{4,6} + \binom{10}{10} = 50521.\end{aligned}$$

In general, from the product $[a_1][a_2]\cdots[a_n]$ of n rows, subtract all properly formed products of $n-1$ rows,

$$[a_1+a_2][a_3]\cdots[a_n],\ [a_1][a_2+a_3][a_4]\cdots[a_n],\ \ldots,\ [a_1][a_2]\cdots[a_{n-1}+a_n],$$

add all products of $n-2$ rows,

$$[a_1+a_2+a_3][a_4]\cdots[a_n],\ [a_1][a_2+a_3+a_4]\cdots[a_n],\ \ldots,$$
$$[a_1+a_2][a_3+a_4][a_5]\cdots[a_n],\ \ldots,$$

subtract all products of $n-3$ rows, etc., and finally assign the appropriate signs to the products of two rows,

$$[a_1][a_2+\cdots+a_n],[a_1+a_2][a_3+\cdots+a_n],\ldots,[a_1+a_2+\cdots+a_{n-1}][a_n],$$

and the term consisting of one row,

$$[a_1+a_2+\cdots+a_n].$$

The appropriate multinomial coefficients can now be readily assigned to each term and a formula for E_n as a linear combination of such multinomial coefficients is achieved.

Conclusion We restricted our results here to the cases $[2,2,\ldots,2]$ and $[1,2,2,\ldots,2]$ since our main goal was to give a nonrecursive formula for E_n. It is clear that the two-step procedure can be applied to the more general skew-hook $[a_1,a_2,\ldots,a_n]$. In fact, Carlitz [3] has developed generating functions for certain special cases of this more general situation. While Carlitz has stated a theoretical result that includes the present main result, it was an effort to make the result more accessible that prompted this present treatment; it has its own intrinsic appeal due to the pleasantness of the box algebra.

A very extensive bibliography containing 81 references can be found in the Abramson paper [1]. Entringer [4] obtains a recurrence relation involving the number of alternating permutations. See Gould [6] for some history on "up-down" permutations.

REFERENCES

1. Morton Abramson, Permutations related to secant, tangent and Eulerian numbers, *Canad. Math. Bull.* 22 (1979).
2. D. André, Sur les permutationes alternées, *J. Math. Pures Appl.* 7 (1881), 167–184.
3. L. Carlitz, Permutations with prescribed pattern, *Math. Nachr.* 58 (1973), 31–53.
4. R. C. Entringer, A combinatorial interpretation of the Euler and Bernoulli numbers, *Nieuw Arch. Wisk.* 14, (1966), 241–246.
5. H. O. Foulkes, Enumeration of permutations with prescribed up-down and inversion sequences, *Discrete Math.* 15 (1976), 235–252.
6. H. W. Gould, Comment on Problem 67-5, "up-down" permutations, *SIAM Rev.* 10 (1968), 225–226.
7. A. J. Kempner, On the shape of polynomial curves, *Tôhoku Math. J.* 37 (1933), 347–362.
8. E. Netto, *Lehrbuch der Combinatorik*, Leipzig und Berlin, 1927.
9. N. J. A. Sloane, *Handbook of Integer Sequences*, Academic Press, New York, 1973.

Complex Power Series—a Vector Field Visualization

ALAN D. GLUCHOFF
Villanova University
Villanova, PA 19085

How can we visualize the behavior of complex power series and convey to students a clear understanding of convergence and divergence? In this note we describe a visualization technique based on the use of Pólya vector fields, as introduced by Pólya and Latta [3] and further developed by Braden [1], [2].

In [3], Pólya introduced the notion of a vector field representation of a complex function of a complex variable: If $f(x+iy)=u(x,y)+iv(x,y)$, then the associated Pólya field is $F(x,y)=(u(x,y),-v(x,y))$. A plot of the vector field can then be used to visualize the behavior of f. The complex conjugate is used rather than (u,v) itself because the Cauchy-Riemann equations for analyticity of f translate into zero divergence and zero curl for F; thus a nice interpretation of analyticity in vector field terms is available. Of course, the conjugate field preserves all important features of f.

The vector field representation is ideally suited to depict the process of convergence of a sequence of complex functions. Each function f_n is represented by a vector field F_n on the common domain of definition, and convergence of f_n to a limiting function f can be envisioned as a "stabilizing" of the vector fields F_n to a limiting configuration, that of the vector field for f. Convergence at a point is thus represented by a sequence of arrows based at a common point tending toward a limiting arrow. Divergence appears as either an oscillation or a growth in magnitude without bound of a sequence of common-based arrows.

The production and exploration of such sequences of vector fields is easy in an interactive computer system such as *Mathematica*. We describe a program written in the *Mathematica* language and implemented on the Macintosh II that allows construction of such fields for a complex power series. The program displays the partial sums of a series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ in vector field form.

The first input to the program is a domain, which can be a disc or an annulus with center z_0 and any radii. The next input is a "tolerance pair" $[a,b]$; only vectors whose magnitudes are between a and b will be shown. Then the user gives the number of annuli into which the display will be divided, and the closed form expression for a_n.

The program computes and displays the fields in succession; vectors are plotted as arrows with bases on circles determined by the subdivision annuli. The lengths of the vectors are scaled logarithmically to avoid messy overlaps; this does not alter the behavior of the sequence. Arrows whose lengths exceed the upper tolerance level are deleted and replaced by dots. Termination of the program leaves us with a field that may be taken as a representation of $\sum_{n=0}^{\infty} a_n(z-z_0)^n$. The animation feature of *Mathematica* allows the fields to be presented in an animated sequence, showing the meaning of convergence and divergence dynamically.

Example. We explore the series $\sum_{n=1}^{\infty} n(z-(1+3i))^n$ with domain $D=\{z:|z-(1+3i)|<2\}$, five subdivision annuli, and tolerance pair $[.01,100]$. FIGURES 1–5 show the fields for the 4th, 6th, 7th, 14th, and 48th partial sums. In the animated sequence, the viewer sees the vectors in each of the outer three rings gyrate, grow in length, then vanish as their vector lengths exceed the prescribed tolerance. (Some of this behavior can be observed in the third ring from the center in FIGURES 1–3; note, for example, the arrow in the upper right whose base is marked by a heavy dot.) The inner two rings continue to gyrate, but eventually stabilize to the configuration shown in FIGURE 5. So we see convergence on a disc centered at $1 + 3i$ having radius somewhere between 0.8 and 1.2. Further runs with different domains and tolerances can more accurately pin down the disc of convergence and show details on how divergence occurs outside this disc and how stability is achieved inside. Animation then makes this come alive!

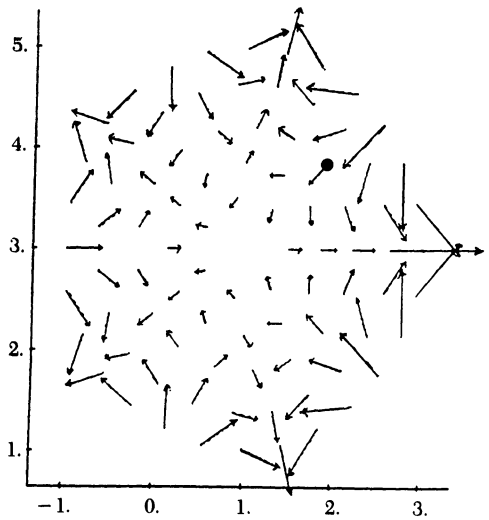


FIGURE 1
4th partial sum of $\Sigma_{n=1}^{\infty} n(z-(1+3i))^n$ on D.

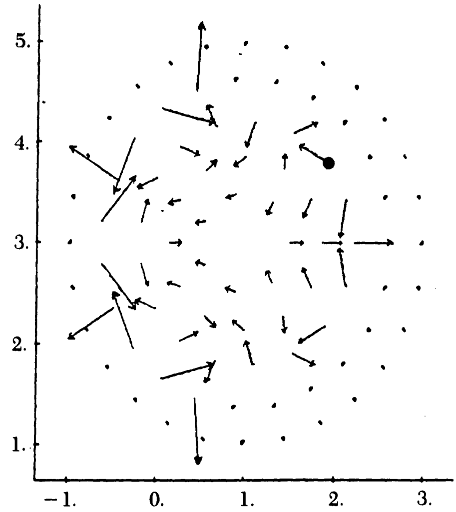


FIGURE 2
6th partial sum.

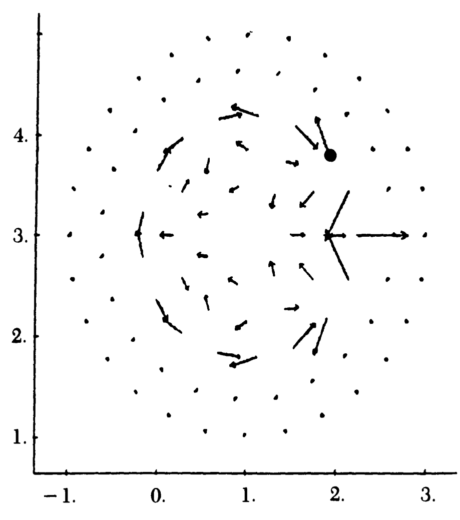


FIGURE 3
7th partial sum.

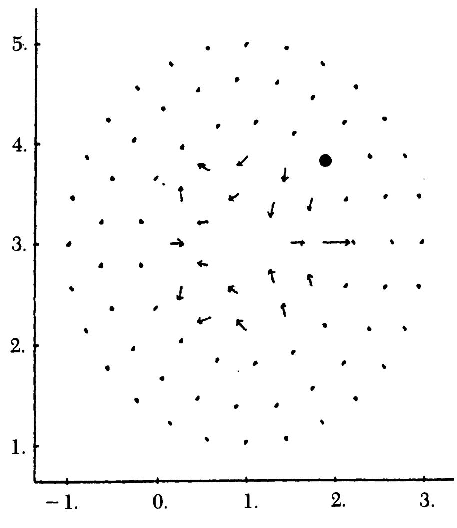


FIGURE 4
14th partial sum.

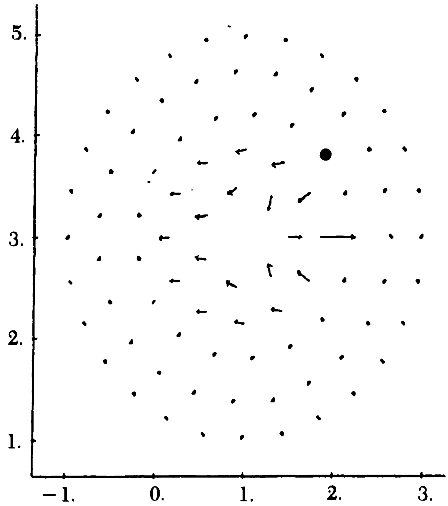


FIGURE 5
48th partial sum.

Even if you are very familiar with the usual treatment of this topic, you may wish to test yourself on the following questions about this example. Answers can be found by looking at the entire run and by making other runs. Try to answer without looking at the figures!

1. On the interior of the disc of convergence, which rings will stabilize first, the inner or outer rings? Why?

2. On any ring on which divergence occurs, will it occur more rapidly (in length) on the right or left portion of the ring? Why?

3. On the rings on which stability occurs, is there more angular gyration on the left or right portion of the ring? Why?

4. If we used a higher upper tolerance, could we see divergence without disappearance of the outer rings, or would they vanish regardless of tolerance used? Why?

5. How would runs for, say, $\Sigma_{n=1}^{\infty}(z-(1+3i))^n$, compare to our example? Would we see convergence occur more or less rapidly? What about divergence? Why?

A power series is only one of many classes of sequences that can be represented by vector fields; a Laurent series is another obvious candidate. I hope animation complements the standard formal treatment and provides a tangibility that the usual computations, e.g., finding radii of convergence, do not.

I would like to thank Bart Braden for numerous discussions and encouragement in this investigation. I would also like to thank Najib Nadi for technical assistance on the implementation and execution of the program.

REFERENCES

1. Bart Braden, Picturing functions of a complex variable, *College Math. J.* 16 (1985), 63–72.
2. Bart Braden, Pólya's geometric picture of complex contour integrals, this MAGAZINE 60 (1987), 321–327.
3. George Pólya and Gordon Latta, *Complex Variables*, John Wiley & Sons, Inc., New York, 1974.

PROBLEMS

LOREN C. LARSON, *editor*
St. Olaf College

GEORGE GILBERT, *associate editor*
Texas Christian University

Proposals

To be considered for publication, solutions should be received by November 1, 1993.

1423. *Proposed by Mirel Mocanu, University of Craiova, Craiova, Romania.*

Two equilateral triangles, of side-lengths a and b respectively, are enclosed in a unit equilateral triangle so that they have no common interior points. Prove that $a+b\leq 1$.

1424. *Proposed by J. C. Binz, University of Bern, Bern, Switzerland.*

Find all positive integers n such that

$$M_n=\left\{\binom{1}{2},\binom{2}{2},\binom{3}{2},\dots,\binom{n}{2}\right\}$$

is a complete set of residues modulo n.

1425. *Proposed by William P. Wardlaw, U.S. Naval Academy, Annapolis, Maryland.*

Let p be a prime number and let $\mathbf{A}$ be a $(p-1)\times(p-1)$ matrix over the field of rational numbers such that $\mathbf{A}^p=\mathbf{I}\neq\mathbf{A}$. Show that if $f(x)$ is any nonzero polynomial with rational coefficients and degree less than $p-1$, then $f(\mathbf{A})$ is nonsingular.

1426. *Proposed by Jiro Fukuta, Gifu-ken, Japan.*

Consider a circle with center at O, and a regular n-gon, $A_1A_2\dots A_n$, contained entirely within the given circle. Let C denote the center of the n-gon. Let P_iQ_i, $i=1,2,\dots,n$ be the chords of the given circle that are perpendicular to CA_i at A_i. Prove that $\sum_{i=1}^{n}(CP_i^2+CQ_i^2)$ is a constant.

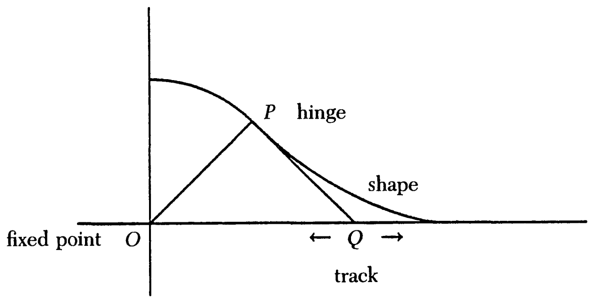
ASSISTANT EDITORS: Clifton Corzat, Bruce Hanson, Richard Kleber, Kay Smith, *and* Theodore Vessey, *St. Olaf College and* Mark Krusemeyer, *Carleton College. We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals should be accompanied by solutions, if at all possible, and by any other information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution. An asterisk (*) next to a problem number indicates that neither the proposer nor the editors supplied a solution.*

Solutions should be written in a style appropriate for Mathematics Magazine. Each solution should begin on a separate sheet containing the solver's name and full address.

Solutions and new proposals should be mailed in duplicate to Loren Larson, Department of Mathematics, St. Olaf College, 1520 St. Olaf Ave., Northfield, MN 55057-1098 or mailed electronically via fax: (507) 663-3549 or e-mail: larson@stolaf.edu.

1427. *Proposed by Gary L. Van Velsir, Anne Arundel Community College, Arnold, Maryland.*

A bi-fold closet door consists of two one-foot-wide panels, hinged at point P. One of the panels is fixed at the point O (see figure). Assume that as the endpoint Q moves to the right, the door rubs against a thick carpet. What shape will be swept out on the carpet?



Quickies

Answers to the Quickies are on page 201.

Q805. *Proposed by Murray S. Klamkin, and A. H. Rhemtulla, University of Alberta, Edmonton, Alberta, Canada.*

Let S be a semigroup.

(i) Given that $x^r y^r = y^r x^r$ for all $x, y \in S$ and for $r = 2, 3, \ldots,$ must S be commutative?

(ii) Given that $x^2 = x^3 = x^4 = \cdots$ for all $x \in S$, must S be commutative?

Q806. *Proposed by Gerry Ladas, The University of Rhode Island, Kingston, Rhode Island.*

Show that for $k = 1, 2, \ldots,$

$$\frac{1}{k} \leq 2\cos\left(\frac{k\pi}{2k+1}\right).$$

Q807. *Proposed by Norman Schaumberger, Hofstra University, Hempstead, New York.*

Evaluate

$$\sum_{k=1}^{\infty}\left(\frac{1}{3k-2} + \frac{1}{3k-1} - \frac{2}{3k}\right).$$

Solutions

A pair of periodic functions **June 1992**

1398. *Proposed by Florin S. Pîrvănescu, Slatina, Romania.*
Let $f, g : \mathbf{R} \to \mathbf{R} - \{0, 1\}$ be functions such that

$$f(x+1) = \frac{g(x)}{f(x)} \quad \text{and} \quad g(x+1) = \frac{g(x)-1}{f(x)-1}$$

for every $x \in \mathbf{R}$. Prove that f and g are periodic.

I. *Solution by Richard Holzsager, American University, Washington, D.C.*
Let $H(x) = x/(x-1)$. Then $H(H(x)) = x$, and by direct calculation, $g(x+2) = H(f(x))$ and $f(x+3) = H(g(x))$. Therefore $f(x+5) = H(g(x+2)) = H(H(f(x))) = f(x)$, and $g(x+5) = H(f(x+3)) = H(H(g(x))) = g(x)$.

II. *Solution by Gerry Ladas, University of Rhode Island, Kingston, Rhode Island.*
Eliminating $f(x)$ we get

$$g(x+2) = \frac{g(x+1) + g(x) - 1}{g(x) - 1}.$$

Now set $g(x) = 1 + h(x)$ to obtain

$$h(x+2) = \frac{h(x+1)+1}{h(x)}.$$

It is known that h is periodic of period 5 (see problem and solution **1343**, this Magazine, April 1991, p. 133). It follows that g has period 5. Since $f(x) = \dfrac{g(x)-1}{g(x+1)} + 1$, f also has period 5.

Additional results along these lines are included in a paper, "On Rational Recursive Sequences" by V. Lj. Kocic, G. Ladas, and I. W. Rodrigues that will appear in the *Journal of Mathematical Analysis and Applications*. Among other things, the authors investigate the periodic nature of solutions to the recursive sequences

$$x_{n+1} = \frac{a + \sum_{i=0}^{k-1} b_i x_{n-i}}{x_{n-k}}, \qquad n = 0, 1, 2, \ldots,$$

and

$$x_{n+1} = \frac{a + bx_n}{A + x_{n-k}}, \qquad n = 0, 1, 2, \ldots,$$

where $a, b, b_0, \ldots, b_{k-1}$ and A are nonnegative numbers and k is a positive integer.

III. *Solution by David Callan, University of Wisconsin, Whitewater, Wisconsin.*
We find from the defining relations that

$$f(x+2) = \frac{f(x)(g(x)-1)}{g(x)(f(x)-1)}, \qquad g(x+2) = \frac{f(x)}{f(x)-1},$$

and then successively,

$$f(x+3)=\frac{g(x)}{g(x)-1},\qquad g(x+3)=\frac{g(x)}{g(x)-f(x)},$$

$$f(x+4)=\frac{g(x)-1}{g(x)-f(x)},\qquad g(x+4)=\frac{f(x)(g(x)-1)}{g(x)-f(x)},$$

$$f(x+5)=f(x),\qquad g(x+5)=g(x).$$

Thus, f and g both have period 5.

It is interesting to note that these equations are sufficient to characterize f and g. To see this, let g be defined arbitrarily on $[0,2)$, subject of course to $g(x)\neq 0,1$ and also subject to $g(t)+g(1+t)\neq 1$ for $0\leq t<1$. Then g is determined on $[0,5)$ by the relations

$$g(2+t)=\frac{g(t)+g(1+t)-1}{g(t)-1},$$

$$g(3+t)=\frac{g(t)g(1+t)}{(g(t)-1)(g(1+t)-1)},$$

$$g(4+t)=\frac{g(t)+g(1+t)-1}{g(1+t)-1},$$

each for $0\leq t<1$, and by periodicity elsewhere.

Note that g will be continuous on $\mathbf{R}$ provided it is continuous on the *closed* interval $[0,2]$. The function f is given in turn by

$$f(x)=\frac{g(x-3)}{g(x-3)-1}.$$

Also solved by *Frank Adrian, P. J. Anderson (Canada), S. F. Barger, J. C. Binz (Switzerland), Nicholas Buck and Edward Dobrowolski (Canada), Con Amore Problem Group (Denmark), Ernie Croot (student), David Doster, Robert L. Doucette, A. J. Douglas (England), E. S. Freidkin, Jane Friedman and Marquis Griffith and Ryan Jackson and Mika Wheeler, Jiro Fukuta (Japan), Wee Teck Gan (student, England), Herbert Gintis, Jerrold W. Grossman, Randy Ho (student), Daryl W. Hochman, Hans Kappus (Switzerland), Brian S. King and Gerald Thompson, Murray S. Klamkin (Canada), Andrew Lazarus, Peter W. Lindstrom, Richard F. McCoart, David E. Manes, Reiner Martin (student), Leroy F. Meyers, Jean-Marie Monier (France), Valerian M. Nita, Ahmet Yaşar Özban (Turkey), R. Glenn Powers, Prestonsburg Community College Problem Solvers Group, Richard Quint, Rolf Richberg (Germany), Richard F. Ryan, Mark Sand, Harvey Schmidt, Jr., John S. Sumner, University of Wyoming Problem Circle, Michael Vowe (Switzerland), Edward T. H. Wang and Wan-Di Wei (Canada), Richard Weida, Chris Wildhagen (The Netherlands), Alan Wilson, Matt Wyneken, Yarmouk University Problem Group (Jordan), Paul J. Zwier, and the proposer.*

A sequence of means **June 1992**

1399. *Proposed by Ioan Sadoveanu, Ellensburg, Washington.*

Let a,b,c be positive numbers such that $a+b+c=1$. Let x_0,y_0,z_0 be positive numbers and for each $n\geq 0$, let

$$x_{n+1}=ax_n+by_n+cz_n,\qquad y_{n+1}=x_n^a y_n^b z_n^c,\qquad z_{n+1}=\left(\frac{a}{x_n}+\frac{b}{y_n}+\frac{c}{z_n}\right)^{-1}.$$

Show that the sequences $(x_n),(y_n)$, and (z_n) each converge, and that they converge to the same limit.

I. *Solution by Robert L. Doucette, McNeese State University, Lake Charles, Louisiana.*

Note that x_{n+1}, y_{n+1}, and z_{n+1} are the weighted arithmetic, geometric, and harmonic means (with weights a, b, c) respectively of x_n, y_n and z_n, and it is well known that

$$z_{n+1} \le y_{n+1} \le x_{n+1}. \qquad (*)$$

Hence

$$x_{n+2} = ax_{n+1} + by_{n+1} + cz_{n+1} \le ax_{n+1} + bx_{n+1} + cx_{n+1} = x_{n+1},$$

and

$$z_{n+2} = \left(\frac{a}{x_{n+1}} + \frac{b}{y_{n+1}} + \frac{c}{z_{n+1}}\right)^{-1} \ge \left(\frac{a}{z_{n+1}} + \frac{b}{z_{n+1}} + \frac{c}{z_{n+1}}\right)^{-1} = z_{n+1}.$$

This shows that (x_n) and (z_n) are convergent sequences, the former being a decreasing sequence bounded below by z_1, and the latter being an increasing sequence bounded above by x_1. Since $x_{n+1} = ax_n + by_n + cz_n$ and $b \neq 0$, the sequence (y_n) is convergent as well.

Let x, y, z be the limits of $(x_n), (y_n), (z_n)$, respectively. Taking the limit on each side of $x_{n+1} = ax_n + by_n + cz_n$ yields $x = ax + by + cz$, and this, together with $a + b + c = 1$, implies that $b(x-y) + c(x-z) = 0$. Since b and c are positive and $x \ge y \ge z$ from $(*)$, we must have $x = y = z$.

II. *Solution by the University of Wyoming Problem Circle, University of Wyoming, Laramie, Wyoming.*

We prove a more general result about sequences of means.

Let $\mathbf{R}_+$ denote the set of positive real numbers, and call a function $\Phi : \mathbf{R}_+^k \to \mathbf{R}_+$ a *mean* on $\mathbf{R}_+^k$ if it satisfies the following three properties:

(1) For all $(x_1, x_2, \ldots, x_k) \in \mathbf{R}_+^k$ and all $a \in \mathbf{R}_+$,

$$\Phi(ax_1, ax_2, \ldots, ax_k) = a\Phi(x_1, x_2, \ldots, x_k).$$

(2) For all $(x_1, x_2, \ldots, x_k) \in \mathbf{R}_+^k$,

$$\min\{x_1, x_2, \ldots, x_k\} \le \Phi(x_1, x_2, \ldots, x_k) \le \max\{x_1, x_2, \ldots, x_k\}.$$

(3) Φ is strictly increasing with respect to each variable.

We note that if a, b, c are positive numbers with $a + b + c = 1$, then the three functions

$$\Phi_1(x, y, z) = ax + by + cz, \qquad \Phi_2(x, y, z) = x^a y^b z^c,$$

$$\Phi_3(x, y, z) = \left(\frac{a}{x} + \frac{b}{y} + \frac{c}{z}\right)^{-1}$$

are all means on $\mathbf{R}_+^3$. Therefore, the following theorem proves the desired result.

THEOREM. *Suppose that $\Phi_1, \Phi_2, \ldots, \Phi_k$ are means on $\mathbf{R}_+^k$, and $x_1^{(0)}, x_2^{(0)}, \ldots, x_k^{(0)}$ are positive. For $i = 1, 2, \ldots, k$ and each $n \ge 0$, define*

$$x_i^{(n+1)} = \Phi_i\left(x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}\right).$$

Then the sequences $(x_i^{(n)})$ all converge, and they converge to the same limit.

Proof. For $n \geq 0$ define

$$m_n = \min\{x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}\} \quad \text{and} \quad M_n = \max\{x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}\}.$$

Then for $i = 1, 2, \ldots, k$, properties (3) and (2) imply that

$$x_i^{(n+1)} = \Phi_i\left(x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}\right)\begin{cases} \geq \Phi_i(m_n, m_n, \ldots, m_n) = m_n, \\ \leq \Phi_i(M_n, M_n, \ldots, M_n) = M_n. \end{cases}$$

It follows that $m_n \leq m_{n+1} \leq M_{n+1} \leq M_n$ so that the sequence (m_n) is increasing and bounded above while the sequence (M_n) is decreasing and bounded below. Therefore, the limits

$$m = \lim_{n \to \infty} m_n \quad \text{and} \quad M = \lim_{n \to \infty} M_n$$

exist. Moreover, they satisfy $0 < m \leq M$, since $0 < m_0 \leq m_n \leq M_n$. If $m = M$, then the inequalities $m_n \leq x_i^{(n)} \leq M_n$ imply the result by the squeeze principle.

So assume that $m < M$. We will reach a contradiction.

Now, for each $n \geq 0$, one of the numbers $x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}$ is equal to m_n, and the other numbers are at most M_n. Therefore, for each i,

$$\begin{aligned} x_i^{(n+1)} &= \Phi_i\left(x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}\right) \\ &\leq \max\{\Phi_i(m_n, M_n, \ldots, M_n), \Phi_i(M_n, m_n, \ldots, M_n), \ldots, \Phi_i(M_n, \ldots, M_n, m_n)\} \\ &= \max\left\{\Phi_i\left(\frac{m_n}{M_n}, 1, \ldots, 1\right), \Phi_i\left(1, \frac{m_n}{M_n}, \ldots, 1\right), \ldots, \Phi_i\left(1, \ldots, 1, \frac{m_n}{M_n}\right)\right\} M_n. \end{aligned}$$

However, $m_n \leq m$ and $M_n \geq M$ so that $m_n/M_n \leq m/M < 1$. Therefore,

$$x_i^{(n+1)} \leq \max\left\{\Phi_i\left(\frac{m}{M}, 1, \ldots, 1\right), \Phi_i\left(1, \frac{m}{M}, \ldots, 1\right), \ldots, \Phi_i\left(1, \ldots, 1, \frac{m}{M}\right)\right\} M_n = \alpha_i M_n$$

where $\alpha_i = \max\{\Phi_i((m/M), 1, \ldots, 1), \ \Phi_i(1, (m/M), \ldots, 1), \ldots, \Phi_i(1, \ldots, 1, (m/M))\}$. Then $\alpha_i < 1$ by property (3) (since $\phi_i(1, 1, \ldots, 1) = 1$). Consequently,

$$\begin{aligned} M_{n+1} &= \max\{x_1^{(n+1)}, x_2^{(n+1)}, \ldots, x_k^{(n+1)}\} \\ &\leq \max\{\alpha_1 M_n, \alpha_2 M_n, \ldots, \alpha_k M_n\} \\ &= \max\{\alpha_1, \alpha_2, \ldots, \alpha_k\} M_n = \alpha M_n \end{aligned}$$

where $\alpha = \max\{\alpha_1, \alpha_2, \ldots, \alpha_k\} < 1$. Iterating yields $M_n \leq \alpha^n M_0$ so that $M = \lim_{n \to \infty} M_n = 0$. But this contradicts the assumption that $M > m > 0$.

This completes the proof.

The Theorem implies that the means each can use different sets of numbers a, b, c. Property (2) can be replaced by $(2')\Phi(1, 1, \ldots, 1) = 1$. And (1), (2), (3) are equivalent to (1), (2′), (3).

Also solved by *P. J. Anderson (Canada), Michael H. Andreoli, S. F. Barger, Hongwei Chen, Con Amore Problem Group (Denmark), E. S. Freidkin, Wee Tack Gan (student, England), Herbert Gintis, H. Guggenheimer, Richard Holzsager, Liang J. Huang, Hans Kappus (Switzerland), Murray S. Klamkin (Canada), Peter W. Lindstrom, Beatriz Margolis (France), Jean-Marie Monier (France), Richard E. Pfiefer, Stew Roberts, John Sumner, Chris Wildhagen (The Netherlands), Xiaopu Yan, Paul Zwier, and the proposer.*

Other proofs were based on Jensen's Inequality and the power-mean inequality. A more general result can be found in J. M. Borwein and P. B. Borwein, *Pi and the AGM*, Wiley-Interscience Publication, 1987, pp. 267–8.

Condition for commutativity in semigroups **June 1992**

1400. *Proposed by Frances Barry (student) and Desmond MacHale, University College, Cork, Ireland.*

Let S be a semigroup. Suppose that n is a fixed positive integer such that $xy = y^n x^n$ for all $x, y \in S$. Prove that S is commutative.

Solution by Brian Borchers, New Mexico Institute of Mining and Technology, Socorro, New Mexico.

For any $x \in S$, $x^n = xx^{n-1} = (x^{n-1})^n x^n = (x^n)^n$. Then for any $a, b \in S$, $ab = b^n a^n = (a^n)^n (b^n)^n = a^n b^n = ba$. Thus S is commutative.

Also solved by *Kifah Al-Hami, P. J. Anderson* (Canada), *Michael H. Andreoli, S. F. Barger, J. C. Binz* (Switzerland), *Jeff Boersema, Nicholas Buck* (Canada), *Sydney Bulman-Fleming* (Canada), *David Callan, Centre College Math Problem Solving Group, Zhibo Chen, Con Amore Problem Group* (Denmark), *Richard A. Davis, Bill Doran, David Doster, Robert L. Doucette, A. J. Douglas* (England), *Kenneth Fogarty, Albert F. Gilman III, E. C. Greenspan and S. A. Greenspan, Jerrold W. Grossman, Rhett Guthrie* (student), *Lee O. Hagglund, Nick S. Hekster* (The Netherlands), *Russell Jay Hendel, Richard Hinthorn and Mohammad Fatehi, R. Daniel Hurwitz, Ebrahim Jahangard-Mahboob* (student), *Cristian Jansenson, Sharad Keny, Refik Keskin* (Turkey), *Brian King and Gerald Thompson, Murray S. Klamkin* (Canada), *David W. Koster, John Kurtzke and Lewis Lum, James Kuzmanovich, Kee-Wai Lau* (Hong Kong), *Gesing Leung* (Hong Kong), *Peter W. Lindstrom, Richard F. McCoart, Mihalis Maliakas, David E. Manes, Reiner Martin* (student), *Leroy F. Meyers, Kandasamy Muthuvel, David L. Neuhouser, Sofian Obeidat* (Jordan), *Teunis J. Ott, Leonard L. Palmer, Paul Peck, R. Glenn Powers, Richard Quint, James V. Rauff, Rolf Richberg* (Germany), *Richard F. Ryan, George Schillinger, R. P. Sealy* (Canada), *Robert W. Sheets, The Shippensburg University Mathematical Problem Solving Group, John S. Sumner, George D. Tintera, Trinity University Problem Solving Group, William P. Wardlaw, Richard Alan Winton, Edward T. Wong, University of Wyoming Problem Circle, Wayne D. Young, Jiafu Yu, Terry Zeanah, and the proposers.*

Consider the more general condition $(*)$: for each $x, y \in S$ there exist positive integers m and n (depending on x and y) such that $xy = y^m x^n$.

In general, $(*)$ does not imply commutativity; the quaternion group is a counterexample. However, with certain restrictions on m and n, which do not require m and n to be constant, $(*)$ does yield commutativity. See Kowol, G., "Conditions for the commutativity of semigroups", *Proceedings of the American Mathematical Society*, **56** (1976), pp. 85–88. Also see Q805 (this issue).

For ring semigroups, i.e., semigroups that are multiplicative semigroups of rings, the situation is different; condition $(*)$ does imply commutativity. See H. E. Bell, "A commutativity condition for rings", *Canadian Journal of Mathematics*, **28** (1976), pp. 986–991.

Inclusion-exclusion problem **June 1992**

1401. *Proposed by Edwards T. H. Wang, Wilfrid Laurier University, and Wan-Di Wei, University of Waterloo, Waterloo, Canada.*

Determine the number of permutations π of $1, 2, \ldots, 2n$ with the property that $|\pi(i+1) - \pi(i)| = n$ for some i, $1 \le i \le 2n-1$. Express your answer in the form $\sum_{k=1}^{n} A_k$.

Solution by the Con Amore Problem Group, Royal Danish School of Educational Studies, Copenhagen, Denmark.

For $j = 1, 2, \ldots, n$ we let S_j be the set of permutations π of $1, 2, \ldots, 2n$ for which the numbers j and $n+j$ are images of neighboring numbers, that is, $|\pi^{-1}(j) - \pi^{-1}(n+j)| = 1$.

The number x we are looking for can, according to the Principle of Inclusion and Exclusion, be expressed in the following way:

$$x = \sum_{1 \le j \le n} |S_j| - \sum_{1 \le j_i < j_2 \le n} |S_{j1} \cap S_{j2}| + \sum_{1 \le j_1 < j_2 < j_3 \le n} |S_{j_1} \cap S_{j_2} \cap S_{j_3}| - + \cdots, \quad (*)$$

where, as we shall see, the right-hand side consists of n terms, the first of them a sum of $\binom{n}{1}$ equal numbers, the second of them a sum of $\binom{n}{2}$ equal numbers, the third of them a sum of $\binom{n}{3}$ equal numbers, and so on. We calculate these n terms separately.

For each number j with $1 \leq n \leq j$ there are two ways in which j and $n+j$ can be images of given neighboring numbers, since we can have $\pi^{-1}(j) - \pi^{-1}(n+j)$ equal to either 1 or -1. For each of these two possibilities the pair $\{j, n+j\}$ and the $2n-2$ individual members of the set $\{1,2,\ldots,2n\} - \{j, n+j\}$ can be permuted in $(2n-1)!$ ways. So we have $|S_j| = 2 \cdot (2n-1)!$, and, since j can be chosen in $\binom{n}{1}$ ways, the first term of the right-hand side of $(*)$ is $\binom{n}{1} \cdot 2 \cdot (2n-1)!$.

For each pair $\{j_1, j_2\}$ with $1 \leq j_1 < j_2 \leq n$ there are 2^2 ways in which j_1 and $n+j_1$ can be images of given neighboring pairs and also j_2 and $n+j_2$ images of given neighboring pairs. For each of these 2^2 possibilities the two pairs $\{j_1, n+j_1\}$ and $\{j_2, n+j_2\}$ and the $2n-4$ individual members of the set $\{1,2,\ldots,2n\} - \{j_1, j_2, n+j_1, n+j_2\}$ can be permuted in $(2n-2)!$ ways. So we have $|S_{j_1} \cap S_{j_2}| = 2^2 \cdot (2n-2)!$ and, since the pair $\{j_1, j_2\}$ can be chosen in $\binom{n}{2}$ ways, the second term of the right-hand side of $(*)$ is $\binom{n}{2} \cdot 2^2 \cdot (2n-2)!$.

Continuing this way we find

$$x = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \cdot 2^k \cdot (2n-k)!.$$

Also solved by *J. C. Binz (Switzerland), David Callan, John W. Krussel, Peter W. Lindstrom, Richard F. McCoart, John S. Sumner and Kevin Dove, and the proposers.*

Callan showed that the proportion of permutations with the stated property approaches $1 - 1/e$ as n gets large.

Inequality in a circumcircle **June 1992**

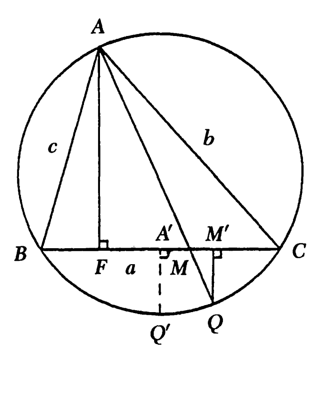
1402. *Proposed by Florin S. Pîrvănescu, Slatina, Romania.*

Let ABC be a given triangle, and M, N, and P be arbitrary points in the interiors of the line segments BC, CA, and AB respectively. Let lines AM, BN, and CP intersect the circumcircle of ABC in points Q, R, and S respectively. Prove that

$$\frac{AM}{MQ} + \frac{BN}{NR} + \frac{CP}{PS} \geq 9.$$

Solution by Richard E. Pfiefer, San Jose State University, San Jose, California.

Referring to the figure, we see that $\dfrac{AM}{MQ} = \dfrac{AF}{M'Q} \geq \dfrac{AF}{A'Q'}$, where A' is the midpoint of BC. Therefore the minimum value of $\dfrac{AM}{MQ} + \dfrac{BN}{NR} + \dfrac{CP}{PS}$ will be obtained (uniquely) when Q, R, S are the midpoints of the arcs BC, CA, AB respectively. So we will henceforth assume that Q, R, S are positioned so that AQ, BR, QS are the angle bisectors of angles A, B, C.



We know that the angle-bisector AQ cuts the side BC (of length a) in the ratio $c:b$, so $BM = ca/(b+c)$, and $MC = (ba)/(b+c)$. Therefore,

$$\frac{AM}{MQ} = \frac{AM^2}{AM \cdot MQ} = \frac{AM^2}{BM \cdot MC} = \frac{AM^2(b+c)^2}{a^2bc}. \qquad (*)$$

From Stewart's Theorem (see Howard Eves, *A Survey of Geometry*, Revised Edition, Allyn and Bacon, 1972, p. 58),

$$c^2MC + b^2BM - aBM \cdot MC - aAM^2 = 0,$$

it follows that

$$AM^2 = c^2\left(\frac{b}{b+c}\right) + b^2\left(\frac{c}{b+c}\right) - \left(\frac{ca}{b+c}\right)\left(\frac{ba}{b+c}\right)$$
$$= bc - \frac{ca^2b}{(b+c)^2}.$$

Substituting this into $(*)$ yields

$$\frac{AM}{MQ} = \left(\frac{b+c}{a}\right)^2 - 1.$$

Similarly, $\frac{BN}{NR} = \left(\frac{a+c}{b}\right)^2 - 1$, $\frac{CP}{PS} = \left(\frac{a+b}{c}\right)^2 - 1$.

Thus, we have

$$\frac{AM}{MQ} + \frac{BN}{NR} + \frac{CP}{PS} \geq \left(\frac{b+c}{a}\right)^2 + \left(\frac{a+c}{b}\right)^2 + \left(\frac{a+b}{c}\right)^2 - 3.$$

By the convexity of $f(x) = x^2$, we have

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

Also, by the arithmetic mean-geometric mean inequality, $x + 1/x \geqslant 2$ for $x > 0$. Thus,

$$\frac{AM}{MQ} + \frac{BN}{NR} + \frac{CP}{PS} \geq \frac{1}{3}\left(\left(\frac{b}{a} + \frac{a}{b}\right) + \left(\frac{c}{a} + \frac{a}{c}\right) + \left(\frac{c}{b} + \frac{b}{c}\right)\right)^2$$
$$\geq \frac{1}{3}(2+2+2)^2 - 3 = 9.$$

Equality holds if and only if $a = b = c$; that is to say, if and only if triangle ABC is equilateral and M, N, P are the midpoints of the sides.

Also solved by *P. J. Anderson (Canada), Con Amore Problem Group (Denmark), Jiro Fukuta (Japan), H. Guggenheimer, Hans Kappus (Switzerland), Murray S. Klamkin, Prestonsburg Community College Problem Solvers Group, John Sumner and Kevin Dove, László Szücs, Michael Vowe, and the proposer.*

Klamkin notes that the problem appeared in *Crux Mathematicorum* (Solution, **16** (1990), p. 158–159). He offers the conjecture that the corresponding inequality for an n-dimensional simplex is $\Sigma A_iM_i/M_iQ_i \geq (n+1)^2/(n-1)$.

Answers

Solutions to the Quickies on page 193.

A805. *Let S be the multiplicative semigroup with generators*

$$A=\begin{pmatrix}0 & a & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},\qquad B=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & b\\ 0 & 0 & 0\end{pmatrix},\qquad C=\begin{pmatrix}0 & 0 & c\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},$$

with a, b, c nonzero real numbers. Then for any matrix D in S, $D^r = 0$ for $r = 2, 3, \ldots$. However,

$$AB=\begin{pmatrix}0 & 0 & ab\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\neq\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}=BA.$$

A806. We will prove the more general inequality that

$$\frac{1}{x}\leq 2\cos\left(\frac{\pi x}{2x+1}\right),\qquad x\geq 1.$$

The change of variables $t=\dfrac{\pi}{2(2x+1)}$ leads to the equivalent inequality

$$\frac{2t}{\pi-2t}\leq\sin t\quad\text{for } 0<t\leq\pi/6.$$

This latter inequality follows by observing that in the interval $0\leq t\leq\pi/6$, the function $(2t)/(\pi-2t)$ is concave up, the function $\sin t$ is concave down, and these two functions agree at the end points of the interval.

Here is another proof of a slightly stronger result. Begin with

$$\cos\left(\frac{k\pi}{2k+1}\right)=\cos\left(\frac{\pi}{2}-\frac{\pi}{4k+2}\right)=\sin\frac{\pi}{4k+2}.$$

It is easy to show that $\sin x/x$ decreases for $0<x\leq\pi/2$. Thus

$$\cos\frac{k\pi}{2k+1}\geq\frac{\sin\pi/6}{\pi/6}\cdot\frac{\pi}{4k+2}=\frac{3}{4k+2}\geq\frac{1}{2k},$$

with equality if and only if $k=1$.

A807. We rewrite the partial sum in the form

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

The result, then, follows by applying the squeeze principle to the following inequality:

$$\ln 3=\int_{n}^{3n}\frac{dx}{x}>\sum_{k=n+1}^{2n}\frac{1}{k}>\int_{n}^{3n}\frac{dx}{x}-\frac{1}{n}+\frac{1}{3n}=\ln 3-\frac{2}{3n}.$$

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 the mathematics literature. Readers are invited to suggest items for review to the editors.

CORRECTION. In Vol. 66, No. 1 (February 1993), p. 68, the KDV (Korteweg-de Vries) equation in line −9 should read $u_t + 6uu_x + u_{xxx} = 0$.

Bartholdi, John J., III, and Kevin L. McCroan, Scheduling interviews for a job fair, *Operations Research* 38 (1990) 951–960.

The Employment Register at the winter U.S. Mathematics Meetings has long been a source of frustration to both employers and candidates, even before proving itself totally inadequate in the face of the recent imbalance of many more candidates than employers. With an underlying algorithm maximizing the total number of interviews, the Employment Register produced many low-priority matches for both employers and candidates, including—bizarrely—interviews that only one of the parties was interested in having. The recent crisis in the employment situation has provoked re-examination of the algorithm and what it should optimize. What is needed, however, is not just a new algorithm but a whole new, much better system. Maybe, for once, mathematicians should turn for advice to systems analysts, like the authors of this paper. They describe a program of theirs that schedules interviews for law firms and students at an Atlanta job fair. This fair's system allows both students and law firms to screen each other beforehand (based on job descriptions and resumés), provides each participant with an interview schedule several weeks in advance, schedules without conflicts *all* mutually requested meetings, and schedules "more interviews sooner" (so that participants tend to finish early). The program runs on a microcomputer, and updating schedules can take place between keystrokes of data entry.

Cipra, Barry A., The ubiquitous Reed-Solomon codes, *SIAM News* 26:1 (January 1993) 1, 11.

Audio compact discs, the picture transmissions from the Voyager II spacecraft to the outer planets, and the coming digital video revolution all owe their success to Reed-Solomon error-correcting codes. Already 33 years old, Reed-Solomon codes are an important contemporary application of finite fields; but curiously, they are found in very few textbooks of abstract algebra. Why, indeed, is mathematics so unlike the other sciences, whose discoveries of last year are in this year's textbook editions? Yet how many instructors of abstract algebra have bothered to learn the ins and outs of error-correcting codes, which they themselves weren't taught about in college? Why are we surprised that more than half of post-calculus mathematics courses are not taught in mathematics departments? Perhaps mathematics instructors need to abandon the philosophy that it's enough to teach just the same (sometimes very) old material that they were taught. Otherwise, today's students are likely to flock to other disciplines that are teaching today's science—and what is really our mathematics, if we will claim it and proclaim it.

Spackman, Kerry, and Sze Tan, When the turning gets tough..., *New Scientist* (13 March 1993) 26–31.

Choosing the right path and speed through turns is what distinguishes champion race car drivers from also-droves. The optimal path through a curve, for a particular vehicle and its performance boundaries, can be calculated through application of the cornu spiral and Fresnel integrals, followed by computer simulation. Not surprisingly, the paths driven by the champions correspond closely to the optimal ones.

Roberts, A. Wayne, project director, *Resources for Calculus Collection*, MAA, 1993. Vol. 1: Solow, Anita (ed.), *Learning by Discovery: A Lab Manual for Calculus*, xiii + 165 pp, \$22(P). ISBN 0–88385–083–4. Vol. 2: Fraga, Robert (ed.), *Calculus Problems for a New Century*, xvi + 427 pp, \$25(P). ISBN 0–88385–084–2. Vol. 3: *Applications of Calculus*, xiv + 262 pp, \$22(P). ISBN 0–88385–085–0. Vol. 4: Jackson, Michael B., and John R. Ramsay (eds.), *Problems for Student Investigation*, xiv + 206 pp, \$22(P). ISBN 0–88385–086–9. Vol. 5: Dudley, Underwood, *Readings for Calculus*, xiv + 196 pp, \$22(P). ISBN 0–88385–087–7. The entire set: \$85.

For the instructor who wants to ease into "calculus reform," or for one who wants to plunge in with both feet, here is a treasure trove. This comprehensive package, produced under NSF sponsorship by faculty at liberal arts colleges, includes 26 computer labs, hundreds of commented problems, 18 applications modules, 30 projects, and 36 readings. Teachers are expressly granted permission to reproduce and modify these materials for class use. Every mathematics department needs one or more copies of this series in its teaching resources library, and most instructors of calculus are going to want to have their own copies.

Swetz, Frank J., *The Sea Island Mathematical Manual: Surveying and Mathematics in Ancient China*, Penn State Press, 1992; xi + 73 pp, \$25, \$9.95(P). ISBN 0–271–00799–0

Composed in 236 A.D. by Liu Hui, the *Sea Island Mathematical Manual* describes accomplishments in surveying techniques that preceded those in the West by 1,000 years.

Barnsley, Michael F., and Lyman P. Hurd, *Fractal Image Compression*, A K Peters, Ltd. (289 Linden St., Wellesley, MA 02181); xi + 244 pp, \$49.50. ISBN 0–1–56881–000–8

In December 1992, Microsoft published "Microsoft Encarta," a compact disc including seven hours of sound, 100 animations, 800 color maps (with zoom ability), and 7,000 photographs, all encoded in less than 600 MB—thanks to fractals, specifically, the fractal image compression technique invented (and patented) by author Michael Barnsley. Photographs illustrate the compression algorithms, for which C code is given. Readers will need to be familiar with some real analysis (Borel measures, metric spaces).

Holden, Constance, Study flunks science and math tests, *Science* 258 (23 October 1992) 541. Tests flunk, study finds, *Science News* 142 (24 October 1992) 277.

A study by Boston College's Center for the Study of Testing, Evaluation, and Educational Policy, funded by the NSF, has found that school math tests (both standardized and textbook) rely overwhelmingly on testing "lower level thinking skills" (rote memorization and plugging into formulas) rather than "higher order" functions (problem-solving and reasoning). At schools that place high value on scores on standardized tests as a measure of success of the school (including many schools with high minority enrollments), teachers feel pressured to coach toward the tests and forego teaching thinking.

Katz, Victor J., *A History of Mathematics: An Introduction*, HarperCollins, 1993; xiv + 786 pp, $30. ISBN 0-673-38039-4

New histories on the history of mathematics appear rarely; unfortunately, few students take courses in the subject, unless they need to for teacher certification. This splendid new text is specifically designed for prospective teachers (both school and college) of mathematics: It concentrates on topics that occur in the curriculum, and also shows how they were treated in texts of the time (where some of the exercises come from!). Sidebars with biographies, notation, and further topics, plus illustrations of mathematics and mathematicians on postage stamps, enliven the book. Probably because so little of the curriculum involves 20th-century mathematics (unlike the situation in the other sciences), only the last chapter of this book treats it.

Stewart, Ian, *The Problems of Mathematics*, 2nd ed., Oxford University Press, 1992; x + 347 pp, $16.95(P). ISBN 0-19-286148-4

Prospective teachers who learn the history of mathematics from Katz's book above can join the public in supplementing their knowledge of 20th-century mathematics by reading this excellent compendium of contemporary mathematical discoveries. That the first edition of this popular work was out of date after *less than five years* will be a great surprise to the public, many of whom believe that mathematics was finished and inscribed on stone tablets hundreds (if not thousands) of years ago. This edition features new chapters on Kepler's sphere-packing problem, Laczkovich's solution to the circle-squaring problem, the Jones polynomial and its connection with quantum field theory, quantum cryptography, and *lots more*. A dozen or so references for each chapter, to works of varying levels of sophistication, point the reader to further details.

Holcomb, Zealure C., *Interpreting Basic Statistics: A Guide and Workbook Based on Excerpts from Journal Articles*, Pyrczak Publishing (P.O. Box 39731, Los Angeles, CA 90039), 1992; ix + 179 pp $18.50(P). ISBN 0-9623744-4-X

This workbook offers 36 brief excerpts and four complete short articles from research journals, together with exercises. The exercises test students' ability to locate information in the excerpt, perform simple calculations, critique reporting techniques, evaluate interpretations, and assess the design and execution of the experiment. For each excerpt, the book gives a brief note about the statistical topic involved. The skills that a student will develop from this workbook are an important complement to most introductory statistics texts.

Davenport, H., *The Higher Arithmetic: An Introduction to the Theory of Numbers*, 6th ed., Cambridge University Press, 1992; 217 pp, $44.95, $19.95(P).

A well-known classic introduction to number theory has been augmented in this new edition with a chapter (with four exercises) on computers and number theory, by J.H. Davenport. (The bibliography and index should have been revised, too.)

Pennisi, Elizabeth, Mathematics aids fingerprint detection, *Science News* 142 (21 November 1992) 350.

The FBI plans to use wavelets, particularly their data-compression property, to analyze, store and compare fingerprints electronically (and compactly).

NEWS AND LETTERS

TWENTYFIRST ANNUAL USA MATHEMATICAL OLYMPIAD PROBLEMS AND SOLUTIONS

1. Find, as a function of n, the sum of the digits of

$$9 \times 99 \times 9999 \times \cdots \times \left(10^{2^n} - 1\right),$$

where each factor has twice as many digits as the previous one.

Solution. Let $S(M)$ denote the sum of the base-10 digits of M. We shall prove that

$$S\left(9 \times 99 \times \cdots \times (10^{2^n} - 1)\right) = 9 \cdot 2^n.$$

Lemma. If m is a d-digit number and $M = (10^k - 1)m$ where $k \geq d$, then $S(M) = 9k$.

Proof. Write $M = (10^k - 1)m = p + q + r$, where $p = 10^k(m-1)$, $q = (10^d - m)$ and $r = (10^k - 10^d)$. Let us analyze the base-10 representations of p, q and r. The base-10 numeral representing p has the d digits of $m - 1$ followed by k zeros. Since $q = (10^d - 1) - (m-1) = 99\cdots9 - (m-1)$, the leading d digits of p can be paired with the corresponding digits of q so that the sum of each pair is 9. Thus $S(p) + S(q) = 9d$. The leading $k - d$ digits of r are nines and these are followed by d zeros, so $S(r) = 9(k-d)$. Since $k \geq d$, the contributions made by these three terms are non-overlapping in that no two terms have a nonzero digit in the same place. Hence

$$\begin{aligned} S(M) &= S(p) + S(q) + S(r) \\ &= 9d + 9(k-d) \\ &= 9k, \end{aligned}$$

and the lemma is proved.

Set $m = 9 \times 99 \times \cdots \times (10^{2^{n-1}} - 1)$ and $M = (10^{2^n} - 1)m$. Note that

$$m < 10^{1+2+\cdots+2^{n-1}} = 10^{2^n - 1},$$

so m has at most 2^n digits. It follows from the lemma that $S(M) = 9 \cdot 2^n$.

2. Prove

$$\frac{1}{\cos 0^\circ \cos 1^\circ} + \frac{1}{\cos 1^\circ \cos 2^\circ} + \cdots + \frac{1}{\cos 88^\circ \cos 89^\circ} = \frac{\cos 1^\circ}{\sin^2 1^\circ}.$$

Solution. In this discussion, all arguments of trigonometric functions will be in degrees.

We begin by noting that

$$\frac{\cos 1}{\sin^2 1} = \frac{\cot 1}{\sin 1} = \frac{\tan 89}{\sin 1},$$

so the identity to be proved may be written

$$\frac{\sin 1}{\cos 0 \cos 1} + \frac{\sin 1}{\cos 1 \cos 2} + \cdots + \frac{\sin 1}{\cos 88 \cos 89} = \tan 89. \quad (1)$$

Indeed, since

$$\sin 1 = \sin(x+1)\cos x - \cos(x+1)\sin x,$$

we have

$$\frac{\sin 1}{\cos x \cos(x+1)} = \tan(x+1) - \tan x,$$

so the left hand side of (1) is the telescoping sum

$$(\tan 1 - \tan 0) + (\tan 2 - \tan 1) + \cdots + (\tan 89 - \tan 88),$$

so the result is evident.

3. For a nonempty set S of integers, let $\sigma(S)$ be the sum of the elements of S. Suppose that $A = \{a_1, a_2, \ldots, a_{11}\}$ is a set of positive integers with $a_1 < a_2 < \cdots < a_{11}$ and that, for each positive integer $n \leq 1500$, there is a subset S of A for which $\sigma(S) = n$. What is the smallest possible value of a_{10}?

Solution. We shall prove that the smallest possible value of a_{10} is 248. For $1 \leq k \leq$

11, let $s_k = a_1 + a_2 + \cdots + a_k$ and let m signify the index for which $s_{m-1} < 1500 \leq s_m$. Note that s_{k-1} is the largest possible value of $\sigma(S)$ where S does not contain any of the elements $a_k, a_{k+1}, \ldots, a_n$ and a_k is the smallest possible value of $\sigma(S)$ where S does contain one of these elements. Now $a_k \leq s_{k-1} + 1$ for $2 \leq k \leq m$, since $s_{k-1} + 1 \leq 1500$ and if $a_k > s_{k-1} + 1$ there would be no subset S yielding $\sigma(S) = s_{k-1} + 1$. Adding s_{k-1} to both sides of $a_k \leq s_{k-1} + 1$, we find $s_k \leq 2s_{k-1} + 1$. Since $s_1 = 1$, it follows by induction that $s_k \leq 2^k - 1$ for $1 \leq k \leq m$. Since $1500 \leq s_m \leq 2^m - 1$, we see that $m = 11$. Hence

$$\left\lceil \frac{s_{k+1} - 1}{2} \right\rceil \leq s_k \leq 2^k - 1, \qquad (2)$$

for $k = 1, 2, \ldots, 10$, where $\lceil x \rceil$ denotes the least integer $\geq x$. In particular,

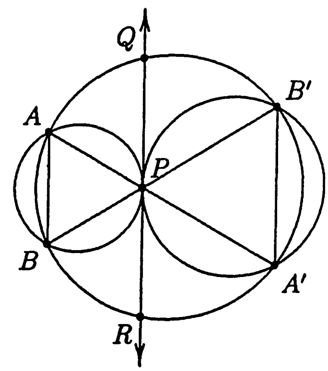
$$s_{10} \geq \left\lceil \frac{s_{11} - 1}{2} \right\rceil \geq \left\lceil \frac{1500 - 1}{2} \right\rceil = 750,$$

and $s_8 \leq 2^8 - 1 = 255$. Thus $a_9 + a_{10} = s_{10} - s_8 \geq 495$, from which we obtain $a_{10} \geq 248$. Looking back at the inequalities which lead to $a_{10} \geq 248$, we see that each can be satisfied with equality by taking $a_{10} = 248$, $a_9 = 247$, $s_{11} = 1500$, $s_{10} = 750$, and $s_8 = 255$. In this case, $a_{11} = s_{11} - s_{10} = 750$. In view of (2), $s_8 = 2^8 - 1$ requires $s_k = 2^k - 1$ and so $a_k = 2^{k-1}$ for $1 \leq k \leq 8$.

To complete the proof that the smallest possible value of a_{10} is 248, we must now check that the set $A = \{1, 2, 4, 8, 16, 32, 64, 128, 247, 248, 750\}$ has the property that for every $n \leq 1500$ there exist $S \subseteq A$ for which $\sigma(S) = n$. To see this, first note that from well-known facts concerning binary representation, every $n \leq 255$ is $\sigma(S)$ for some $S \subseteq \{1, 2, 4, 8, 16, 32, 64, 128\}$. Then by adding 247, 248, and 750 as necessary, we find for every $n \leq 1500$ a set $S \subseteq A$ such that $\sigma(S) = n$.

4. Chords $\overline{AA'}$, $\overline{BB'}$, $\overline{CC'}$ of a sphere meet at an interior point P but are not contained in a plane. The sphere through A, B, C, P is tangent to the sphere through A', B', C', P. Prove that $AA' = BB' = CC'$.

Solution. Intersecting everything with the plane through the four points A, A', B, B' results in a configuration in which $\overline{AA'}$, $\overline{BB'}$ are chords of a circle meeting at P such that the circumcircles of ABP and $A'B'P$ are tangent.



Let $\overleftrightarrow{QR}$ be the common tangent at P as shown above. Then $\angle ABP$ and $\angle APQ$ subtend the same arc in the circle through A, B, P. Similarly $\angle A'PR$ and $\angle A'B'P$ subtend the same arc in the circle through A', B', P. So we have

$$\begin{aligned} \angle ABP &= \angle APQ \\ &= \angle A'PR \qquad \text{(vertical angles)} \\ &= \angle A'B'P \\ &= \angle BAP, \end{aligned}$$

since $\angle A'B'P = \angle A'B'B$ and $\angle BAP = \angle BAA'$ subtend the same arc in the circle through A, A', B, B'. Hence triangle ABP is isosceles, and $AP = BP$. Adding the analogous equality $A'P = B'P$ yields $AA' = BB'$. Similarly $BB' = CC'$.

5. Let $P(z)$ be a polynomial with complex coefficients which is of degree 1992 and has distinct zeros. Prove that there exist complex numbers $a_1, a_2, \ldots, a_{1992}$ such that $P(z)$ divides the polynomial

$$\left(\cdots\left((z - a_1)^2 - a_2\right)^2 \cdots - a_{1991}\right)^2 - a_{1992}.$$

Solution. Set $k = 1992$ and let $r_1, r_2, \ldots, r_k$ be the zeros of $P(z)$. Inductively define nonempty sets S_i and complex numbers a_i as follows. Let

$$S_0 = \{r_1, r_2, \ldots, r_k\}.$$

For $i \geq 1$, let a_i be the average of two numbers in S_{i-1} unless S_{i-1} just has one element, in which case let a_i be that number. Then let

$$S_i = \{(z - a_i)^2 | \, z \in S_{i-1}\}.$$

If S_{i-1} has at least two elements, the two elements of which a_i is the average will yield the same value of $(z - a_i)^2$ and thus S_i will have fewer elements than S_{i-1}. Thus at most $k - 1$ steps are required before S_i contains just one element. It follows that $S_k = \{0\}$. This means that the polynomial

$$\left(\cdots\left((z-a_1)^2 - a_2\right)^2 \cdots - a_{k-1}\right)^2 - a_k$$

vanishes at $r_1, r_2, \ldots, r_k$, so it is a multiple of $P(z)$.

Note. The same argument handles the case in which the zeros of $P(z)$ are not necessarily distinct.

THIRTYTHIRD ANNUAL INTERNATIONAL MATHEMATICAL OLYMPIAD PROBLEMS

1. Find all integers a, b, c with $1 < a < b < c$ such that $(a-1)(b-1)(c-1)$ is a divisor of $abc - 1$.

2. Let $\mathbb{R}$ denote the set of all real numbers. Find all functions $f : \mathbb{R} \to \mathbb{R}$ such that

$$f\left(x^2 + f(y)\right) = y + (f(x))^2$$

for all $x, y \in R$.

3. Consider nine points in space, no four of which are coplanar. Each pair of points is joined by an edge (that is, a line segment) and each edge is either colored blue or red or left uncolored. Find the smallest value of n such that whenever exactly n edges are colored, the set of colored edges necessarily contains a triangle all of whose edges have the same color.

4. In the plane let C be a circle, L a line tangent to the circle C, and M a point on L. Find the locus of all points P with the following property: there exists two points Q, R on L such that M is the midpoint of QR and C is the inscribed circle of triangle PQR.

5. Let S be a finite set of points in three-dimensional space. Let S_x, S_y, S_z be the sets consisting of the orthogonal projections of the points of S onto the yz-plane, zx-plane, xy-plane, respectively. Prove that

$$|S|^2 \leq |S_x| \cdot |S_y| \cdot |S_z|,$$

where $|A|$ denotes the number of elements in the finite set $|A|$. (Note: The orthogonal projection of a point onto a plane is the foot of the perpendicular from that point to the plane.)

6. For each positive integer n, $S(n)$ is defined to be the greatest integer such that, for every positive integer $k \leq S(n)$, n^2 can be written as the sum of k positive squares. (a) Prove that $S(n) \leq n^2 - 14$ for each $n \geq 4$. (b) Find an integer n such that $S(n) = n^2 - 14$. (c) Prove that there are infinitely many integers n such that $S(n) = n^2 - 14$.

The USA placed second in the 33rd International Mathematical Olympiad, held July 10-21, 1992 in Moscow, Russia. Members of the USA team were: Wei-Hwa Huang, Kiran Kedlaya, Robert Kleinberg, Sergey Levin, Leonhard Ng and Andrew Schultz.

The booklet *Mathematical Olympiads 1992* is available from:

Dr. Walter Mientka
Department of Mathematics
University of Nebraska
Lincoln, NE 68588-0658.

For further information about this sequence of examinations, contact the Executive Director of the Committee, Professor Mientka, at the above address. This report was prepared by Cecil Rousseau.

NEW, Industrial Strength

DERIVE® XM

for 386 & 486 based PCs with eXtended Memory

DERIVE® XM (eXtended Memory) is a computer algebra system designed especially for power users who need to do large symbolic math problems as well as heavy duty number crunching.

"Industrial strength" *DERIVE XM* operates identically to the original, menu-driven ***DERIVE, A Mathematical Assistant*** program. *DERIVE XM* provides the same features and functions, intelligently applying the rules of mathematics to solve symbolic and numeric problems in algebra, trigonometry, calculus, and matrix algebra. In addition, *DERIVE XM* can make use of up to 4 gigabytes (4 billion bytes) of extended memory.

DERIVE XM is a powerful tool for mathematicians, engineers, and scientists. It's easy to use, and the size of problems you can solve is limited only by your computer's memory and speed.

DERIVE XM requires a 386 or 486 based PC compatible computer with 2 or more megabytes of memory. The suggested retail price is $375. Contact Soft Warehouse for a list of dealers.

Soft Warehouse INC
HONOLULU • HAWAII

2000 Years of Mathematical Knowledge on a Disk

3660 Waialae Ave., Suite 304 • Honolulu, Hawaii, USA 96816-3236
Phone: (808) 734-5801 • Fax: (808) 735-1105

DERIVE is a registered trademark of Soft Warehouse, Inc.

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

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