INDEX TO VOLUME 67

AUTHORS

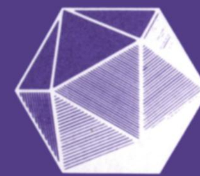
TITLES

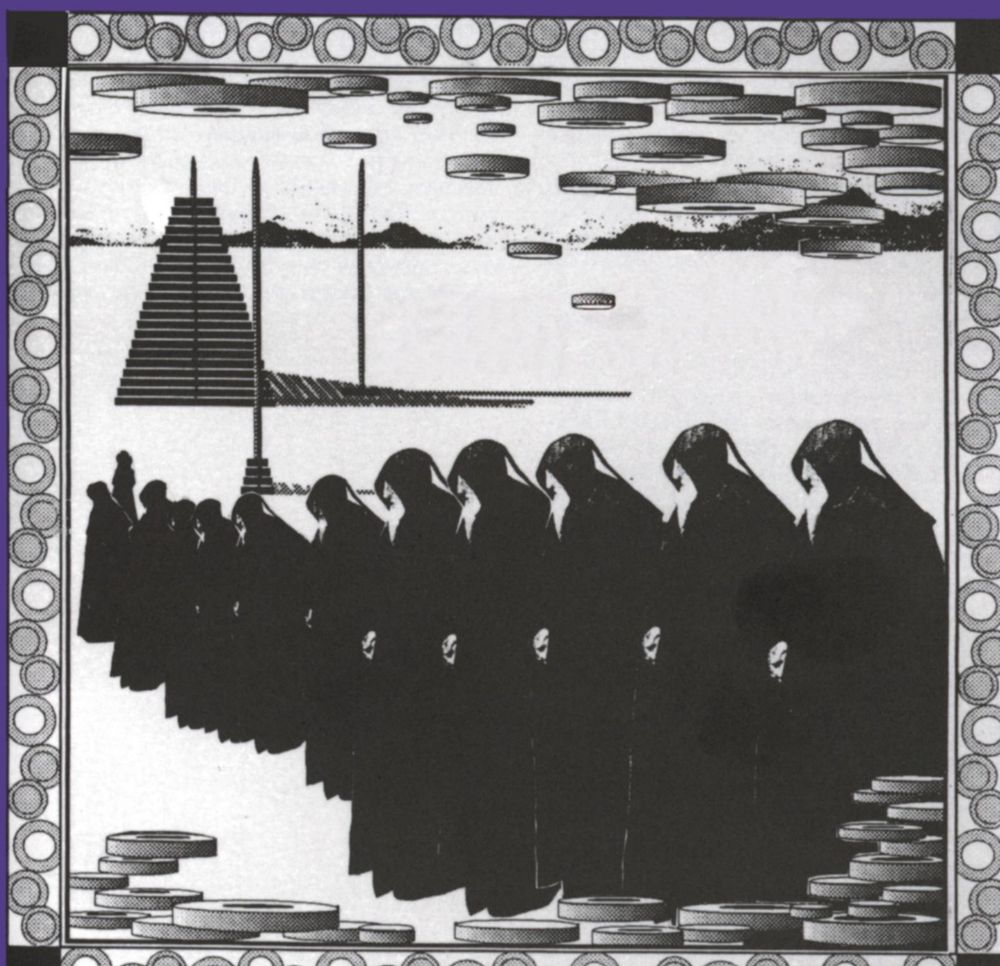
PROBLEMS

P, S, Q, *and* C *refers to Proposals, Solutions, Quickies, and Comments, respectively. Page numbers are given in parantheses. Thus,* P1338(56) *refers to proposal number* 1338, *which appears on page 56.*

Vol. 67, No. 5 December 1994

MATHEMATICS MAGAZINE





- The Towers and Triangles of Professor Claus (or, Pascal Knows Hanoi)
- Bernoulli Trials and Mahonian Statistics: A Tale of Two q's
- Minimum Integral Drawings of the Platonic Groups

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: Paul Zorn, Editor-Elect, Department of Mathematics, St. Olaf College, 1520 St. Olaf Avenue, Northfield, MN 55057-1098. 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 illustration: Carolyn Westbrook, with the monks borrowed from Pieter Bruegel the Elder's *The Misanthrope,* 1568.

AUTHOR

David Poole received a B.Sc. from Acadia University and a Ph.D. from McMaster University in 1984. His research areas are noncommutative ring theory and discrete mathematics. He is also interested in mathematics education and serves on the Education Committee of the Canadian Mathematics Society. For the past decade, he has taught at Trent University, where he is currently helping to develop a project in mathematics and science education for prospective elementary school teachers. His interest in the Tower of Hanoi can be traced to a problem in MATHEMATICS MAGAZINE. One attempt to solve this problem eventually took on a life of its own, and after several metamorphoses, became the present article.

Vol. 67, No. 5 December 1994

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

ARTICLES

The Towers and Triangles of Professor Claus (or, Pascal Knows Hanoi)

DAVID G. POOLE
Trent University
Peterborough, Ontario, Canada K9J 7B8

Introduction

It is a tribute to its inventor that, more than a century after its introduction, the famous Tower of Hanoi puzzle still holds some challenges and surprises for puzzle enthusiasts, mathematicians, and computer scientists. The first version was produced in Paris in 1883 by a certain "Professor Claus." It consisted of three wooden pegs affixed to a base and a tower of eight disks of different diameters on one of the pegs, no disk resting on a smaller one. The object of the puzzle is to transfer the tower to a different peg, moving one topmost disk at a time in such a way that no disk is ever placed on a smaller one.

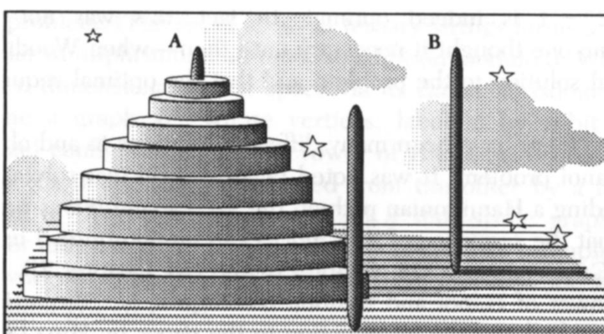


FIGURE 1
The Tower of Hanoi

In 1884, it was revealed by de Parville [13] that "Claus" was an anagrammatic pseudonym for "Lucas"—Edouard Lucas, a well-known number theorist [44] who had a keen interest in recreational mathematics [43]. The popularity of the puzzle was no doubt enhanced by de Parville's account of the mythical Tower of Brahma on which, it was claimed, the Tower of Hanoi[1] was based. In this version, we find three diamond needles, a tower of 64 golden disks, and a group of Brahmin monks attempting to transfer the tower according to the rules given above. We are told that

[1]The Tower of Brahma is without doubt just a colourful legend, but the Tower of Hanoi is real in more ways than one. Although introduced as a puzzle in Paris, the object we know as the "Tower of Hanoi" exists as an ancient relic on the grounds of a Buddist shrine near Hanoi [32]. One can only conjecture about the connection between this artifact and Lucas' puzzle.

once the monks complete their task the world will end in a thunderclap. This apocryphal (and apocalyptic!) tale was recounted in English by Rouse Ball [4] and later by Gardner [28].

The first solution to the Tower of Hanoi puzzle published in the mathematical literature appeared in 1884 in an article by Allardice and Fraser [2]. In fact, if we formulate a mathematical abstraction of the problem to a tower of n disks, it is not difficult to determine the number of moves required to transfer the tower from one peg to another. This problem frequently appears as an elementary exercise in discrete mathematics textbooks; see, for example, [30] or [47]. If we let h_n denote the number of moves required to transfer a tower of n disks, it is easy to see that we first move the top $n-1$ disks from the initial peg to an intermediate peg, then the largest disk to the target peg, and finally the $n-1$ disks from the intermediate peg to the target peg, giving us the recurrence relation

$$h_n = 2h_{n-1} + 1 \tag{1}$$

with initial condition $h_1 = 1$ (or $h_0 = 0$). By solving the recurrence (1) directly or by an easy application of mathematical induction we find the well-known solution

$$h_n = 2^n - 1. \tag{2}$$

(Thus, even if the Brahmins could move one disk per second, we would have h_{64} exceeding five billion centuries; presumably, the world is safe for a while yet!) It is interesting to note that although it is implicit in the problem that h_n should be the *minimum* number of moves required, there is nothing in the solution given above to establish that $2^n - 1$ is indeed optimal. In fact, this was not published until 1981—perhaps no one thought it necessary until then—when Wood [64] showed that (2) is the optimal solution to the problem and that the optimal sequence of moves is unique.

Over the years, there have been many different solutions to and observations about the Tower of Hanoi problem. It was noted by Crowe [9] that solving the problem is equivalent to finding a Hamiltonian path on the n-cube and it was probably known to Lucas himself that the sequence of disk moves can be determined using what is now called a binary Gray code (see [28, Chapter 6] and [29, Chapter 2]). For more on the history of the Tower of Hanoi and solution techniques, see the articles by Gardner [28] and Dewdney [14] and the introduction to the paper by Hinz [33].

In recent years there has been a resurgence of interest in the Tower of Hanoi in the computer science literature. Much of this interest revolves around a lively debate over the relative merits of iterative versus recursive algorithms for solving the problem ([7], [11], [14], [15, Chapter 51], [17], [31], [38], [61]). Computer science textbooks frequently include the Tower of Hanoi as an exercise in comparative programming methodologies ([1], [37]).

It is hard to see how such a simple problem with a complete and well-known solution could possibly hold anything more for us. But as we shall see, Professor "Claus," perhaps unintentionally, built so much into his little puzzle that the solution (2) given above is just the beginning of the story.

In this paper, we will take yet another approach to the Tower of Hanoi problem, providing an essentially algebraic method for solving the classical Tower of Hanoi problem and several of its variants. Using only elementary ideas from graph theory, linear algebra, number theory, and group theory, we will establish a correspondence between configurations of disks arising in the Tower of Hanoi puzzle and the odd binomial coefficients, viewed as entries in Pascal's triangle. We will then exploit this correspondence to solve a variety of Tower of Hanoi problems.

The Hanoi graph

The first time someone is confronted with the Tower of Hanoi puzzle, it is quite likely that he or she will make an error in moving the disks, the result being that the tower is transferred to the target peg in a number of moves that is not minimal. Consideration of this situation raises a number of interesting questions: Which configurations of disks can be obtained from the initial configuration via a sequence of legal moves? Is there an efficient way to recognize whether a given configuration occurs in a minimal move sequence from the initial to the target configurations? Is there an algorithm that will determine a minimal move sequence needed to move from an arbitrary legal configuration of disks to a single tower on a specified peg or to another legal configuration? The notion of error detection and correction has been considered by Walsh [61] and Scarioni and Speranza [53]. The problem involving an arbitrary initial configuration has been called the generalized Tower of Hanoi problem and has been considered by Er in [18]–[20].

These problems are essentially the same and are most easily dealt with by first setting up an appropriate mathematical framework within which we can analyze a variety of Tower of Hanoi problems. To begin, we label the pegs 0, 1, and 2 and label the n disks from smallest to largest $0, 1, \ldots, n-1$. A configuration of disks is called *legal* if no disk is on top of a smaller one. Then observe that a legal configuration of disks corresponds to an n-bit ternary string $a_{n-1} \cdots a_1 a_0$ where $a_i \in \{0,1,2\}$ and $a_i = j$ if disk i lies on peg j. (Note that the largest disk corresponds to the leftmost bit.) Since any set of disks can be placed on a peg in exactly one way and since there are 3^n such ternary strings, there are also 3^n legal configurations on n disks in the Tower of Hanoi problem. (The set of all n-bit ternary strings forms a vector space, V_n, over $\mathbf{Z}_3$ with scalar multiplication and bit-wise addition modulo 3. We will exploit the structure of this n-dimensional vector space via its canonical isomorphism with $\mathbf{Z}_3^n$.)

We now define a graph H_n whose vertices, labelled by n-bit ternary strings, correspond to legal configurations in the Tower of Hanoi on n disks and where two vertices are adjacent if one can be obtained from the other by a legal move. This graph, introduced in [54], is sometimes called a state-space graph and has been mentioned but not fully exploited by several authors (see, for example, [12], [19], [33], and [56]). Following Lu [39], we will refer to these graphs—one for each positive integer n—as *Hanoi graphs*.

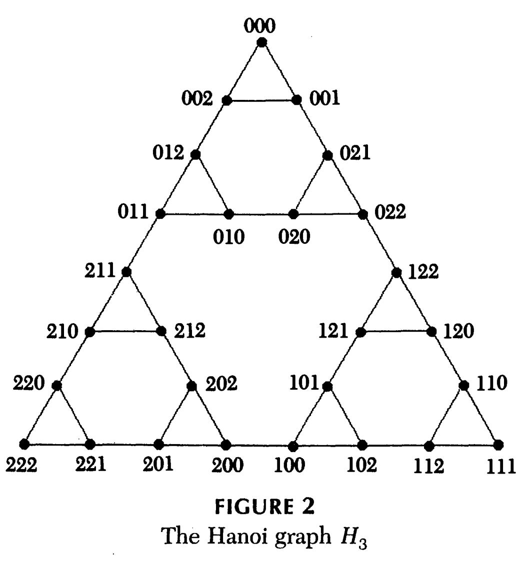


FIGURE 2
The Hanoi graph H_3

Before discussing some of the properties of the Hanoi graph H_n we need to establish some notation and conventions. To avoid unnecessarily cluttered notation we will identify a vertex with its label. The constant strings $00 \cdots 0$, $11 \cdots 1$, and $22 \cdots 2$ with k identical bits will be denoted $\mathbf{0}_k$, $\mathbf{1}_k$, and $\mathbf{2}_k$ respectively; if $k = n$ we will simply write $\mathbf{0}$, $\mathbf{1}$, and $\mathbf{2}$ and we will refer to the corresponding vertices as *corner vertices*. (These vertices correspond to *perfect* Tower of Hanoi configurations: All of the disks are on a single peg.) It will sometimes be convenient to drop the leading zeros of a string so that, for example, 00012 will also be written simply as 12 if it is understood that $n = 5$. For $i \in \{0, 1, 2\}$, $[i]$ denotes the set of all vertices/strings whose leading bit is i; that is, the set of all vertices whose corresponding Tower of Hanoi configuration has the largest disk on peg i. We will refer to [0], [1], and [2] as the *blocks* of H_n and let us establish the convention that we will draw these blocks clockwise as in FIGURE 1.

Now each block of H_n is clearly isomorphic to H_{n-1} and there is exactly one edge between distinct blocks; for example, the unique edge connecting [1] and [2] is between $1\mathbf{0}_{n-1}$ and $2\mathbf{0}_{n-1}$. This corresponds to the fact that the only way to move the largest disk is for all of the smaller disks to be on another peg. This observation shows that H_n can be recursively constructed and allows us to establish many of its properties by elementary mathematical induction. Let us begin by recording some of the properties of the *unlabelled* version of these graphs. (Recall that a graph is *biconnected* if there are at least two distinct paths between any two distinct vertices. The *diameter* of a graph is the maximum distance between two vertices.)

PROPOSITION. *For all positive integers n, H_n is a planar, biconnected, Hamiltonian graph of order 3^n and diameter $2^n - 1$ with exactly 3 vertices of degree 2 and all other vertices of degree 3.*

The proofs of most of these properties follow from the remarks above using mathematical induction and we leave the details as exercises. Note that the diameter of H_n is equal to the maximum number of moves in a minimal move sequence between any two legal Tower of Hanoi configurations with n disks. That this upper bound is $2^n - 1$ is proved in [33, Theorem 1] and stated (but proved incorrectly) in [**65**, Theorem 3]; (2) shows that the bound is sharp and hence is the diameter of H_n. Later we will discuss in more detail the problem of finding the distance between an arbitrary pair of vertices of H_n.

To see that H_n is Hamiltonian, we require the following lemma.

LEMMA 0. *For all positive integers n, there exists a Hamiltonian path in H_n between any pair of distinct corner vertices.*

Proof. Without loss of generality, we may assume that the vertices are $\mathbf{0}$ and $\mathbf{1}$. For $n = 1$, the result is trivial. Assuming the result to hold for H_{n-1}, we consider H_n. Since each block of H_n is a copy of H_{n-1}, by induction we can find Hamiltonian paths from $\mathbf{0} = 0\mathbf{0}_{n-1}$ to $0\mathbf{1}_{n-1}$ in [0], $2\mathbf{1}_{n-1}$ to $2\mathbf{0}_{n-1}$ in [2] and $1\mathbf{0}_{n-1}$ to $1\mathbf{1}_{n-1} = \mathbf{1}$ in [1]. Splicing these three paths together, we obtain the desired Hamiltonian path in H_n. (See FIGURE 3.)

The proof that H_n is in fact Hamiltonian is analogous: By the lemma we can construct Hamiltonian paths from $0\mathbf{1}_{n-1}$ to $0\mathbf{2}_{n-1}$ in [0], $1\mathbf{2}_{n-1}$ to $1\mathbf{0}_{n-1}$ in [1], and $2\mathbf{0}_{n-1}$ to $2\mathbf{1}_{n-1}$ in [2], and then splice these together to get a Hamiltonian cycle in H_n. This property has also been noted in [**12**] and in [**39**] where, in addition, it is shown that the Hamiltonian cycle in H_n is unique up to the direction of traversal.

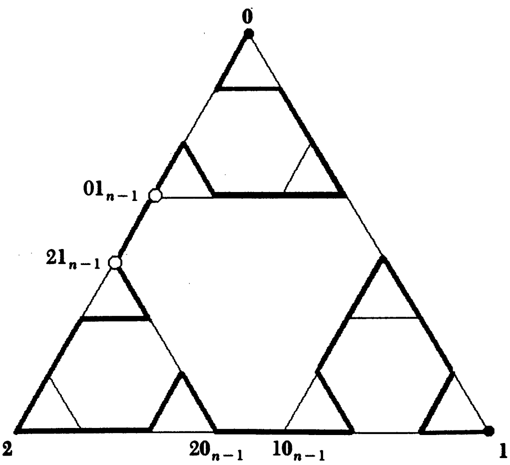


FIGURE 3

Automorphisms and symmetry

The Hanoi graph H_n clearly possesses a high degree of symmetry both as a graph and (as we have drawn it) a geometric object. We can describe this explicitly in terms of the labelled vertices of the graph.

The automorphisms of H_n are the adjacency-preserving bijections of the graph onto itself. It is easy to see that such a mapping is completely determined by its action on the vertices $\mathbf{0}$ and $\mathbf{1} = \mathbf{0}_{n-1}1$. Since there are three possible images of $\mathbf{0}$, namely $\mathbf{0}$, $\mathbf{1}$, and $\mathbf{2}$, and for each of these two possible images of $\mathbf{1}$, we see that there are six distinct automorphisms of H_n. Defining R by $R(\mathbf{0}) = \mathbf{1}$, $R(\mathbf{0}_{n-1}1) = \mathbf{1}_{n-1}2$ and F by $F(\mathbf{0}) = \mathbf{0}$, $F(\mathbf{1}) = \mathbf{2}$, it can easily be checked that $FR \neq RF$. Thus $A(H_n)$, the group of automorphisms of H_n, is nonabelian and so $A(H_n) \cong D_6$, the dihedral group of order 6.

If we imagine H_n to be a plane figure, drawn on an equilateral triangular grid, then D_6 acts on H_n as its group of symmetries with R a rotation and F a reflection. Explicitly, if x is a vertex of H_n and we identify x with its label, we find that $R(x) = \mathbf{1} + x$ and $F(x) = 2x$. (All operations are bitwise modulo 3 or, equivalently, vector operations over $\mathbf{Z}_3$.) Note that, in terms of configurations in the Tower of Hanoi, R corresponds to a cyclic relabelling of the three pegs while F corresponds to interchanging the labels on the target and intermediate pegs.

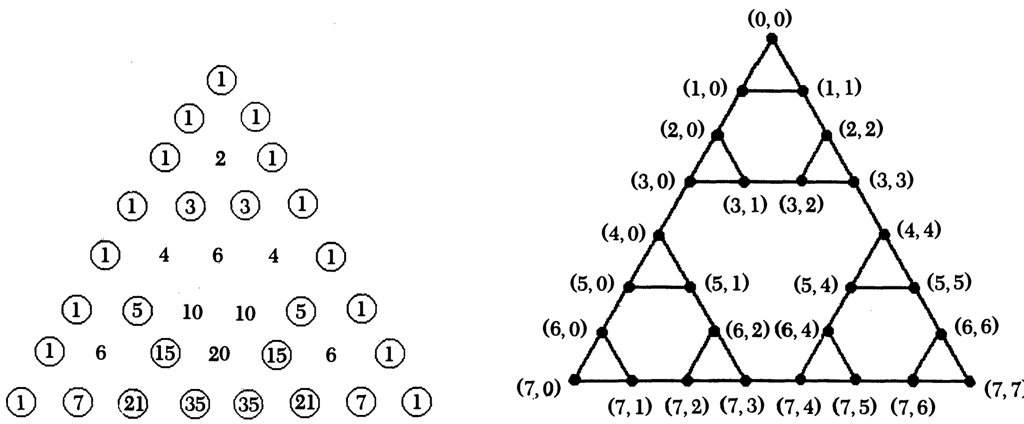
The Hanoi graph exhibits a type of "partial symmetry" that will be of use to us as well: *translational symmetry* on the blocks. Since each block of H_n is a copy of H_{n-1}, we may map a block onto any other block via translation. Clockwise translation is obtained by the mapping $T(x) = x + 12_{n-1}$ while the inverse mapping $T^{-1}(x) = x + 21_{n-1}$ gives anticlockwise translation. We will prove that T (and hence T^{-1}) is adjacency-preserving on the blocks in Lemma 2 below.

So we now have a tool for representing Tower of Hanoi problems. But how can we use Hanoi graphs to *solve*, as well as visualize, these problems?

The Lucas Correspondence

As n becomes large, Stewart ([55], [56]) has observed that H_n becomes a fractal with a striking resemblance to the Sierpiński gasket [46] and Pascal's triangle modulo 2 ([59], [63]), all of which have fractal dimension $\log_2 3$ [63]. Let us define another family of graphs as follows: P_n is the graph whose vertices correspond to the *odd* binomial coefficients $\binom{r}{k}$, $0 \le r < 2^n$, $0 \le k \le r$ with adjacency determined by adjacency (horizontal or diagonal) in Pascal's triangle. Label the vertices with ordered pairs: (r, k) is a vertex if $\binom{r}{k} \equiv 1 \pmod 2$. (See FIGURE 4.)

P_n is constructed recursively from three copies of P_{n-1} (cf. [59]) that we will call the *blocks* of P_n. Formally, the three blocks are the subgraphs of P_n induced by the following sets of vertices: $\{(r,k) | 0 \le r < 2^{n-1}, 0 \le k \le r\}$, $\{(r,k) | 2^{n-1} \le r < 2^n, 2^{n-1} \le k \le r\}$, and $\{(r,k) | 2^{n-1} \le r < 2^n, 0 \le k \le r - 2^{n-1}\}$. By analogy with H_n, we will denote these three blocks by [0], [1], and [2], respectively; although we are using the same notation for the blocks of both graphs, the intended one will usually be clear from context.



(a) (b)

FIGURE 4

We can also define a translation mapping $\hat{T}$ on the blocks of P_n. The mapping translates a vertex to the corresponding vertex in the next block in a clockwise direction. It is defined by

$$\hat{T}((r,k)) = \begin{cases} (r + 2^{n-1}, k + 2^{n-1}) & \text{if } (r,k) \in [0] \\ (r, k - 2^{n-1}) & \text{if } (r,k) \in [1] \\ (r - 2^{n-1}, k) & \text{if } (r,k) \in [2]. \end{cases} \tag{3}$$

It is easy to check that the action of $\hat{T}$ is as described, that $\hat{T}$ has an inverse given by $\hat{T}^{-1} = \hat{T}^2$ and that $\hat{T}$ is adjacency-preserving; we leave the details as an exercise.

From the graphs of H_n and P_n for small values of n and the fact that, for all positive integers n, H_n and P_n have 3^n vertices [49], it is tempting to conjecture that H_n and P_n are isomorphic as graphs. This is indeed the case, as we shall soon prove, but first we need some notation.

For $i \ge 0$, let α_i denote the ternary string 12_i, so that $\{\alpha_i | i \ge 0\} = \{1, 12, 122, 1222, \ldots\}$. For a fixed positive integer n, set $\beta_i = 2^{n+i}\alpha_i \pmod 3$ and $\beta = \{\beta_i | 0 \le i < n\}$. Thus

$$\beta = \begin{cases} \{1, 21, 122, \ldots, 21 \cdots 1\}, & n \text{ even} \\ \{2, 12, 211, \ldots, 21 \cdots 1\}, & n \text{ odd}. \end{cases}$$

Note that β is a basis for the vector space V_n.

We now define mappings ρ and δ from V_n into $\{0,1,2,\dots,2^n-1\}$ and a mapping λ: $V_n \to V_n$. By identifying a vertex of H_n with its label in V_n, we may take these mappings to be mappings of $V(H_n)$, the vertex set of H_n. Technically, we should write ρ_n, λ_n, and δ_n but we will omit the subscripts as this is unlikely to cause confusion.

Let $a = a_{n-1}\beta_{n-1} + \cdots + a_1\beta_1 + a_0\beta_0$, $a_i \in \mathbf{Z}_3$. Define

$$\rho(a) = a_{n-1}^2 2^{n-1} + \cdots + a_1^2 2 + a_0^2.$$

Observe that, modulo 3, $a_i^2 = 0$ if, and only if, $a_i = 0$ and otherwise $a_i^2 = 1$; hence the effect of ρ is to replace ternary strings by binary strings (1 and 2 by 1, 12 and 21 by 10, 122 and 211 by 100, and so on) and then evaluate the result as a decimal integer. Mnemonically, $\rho(a)$ gives the *row* number of the vertex a in H_n as measured from row 0 containing the corner vertex $\mathbf{0}$.

Define

$$\lambda(a) = a_{n-1}^2\beta_{n-1} + \cdots + a_1^2\beta_1 + a_1^2\beta_0.$$

Observe that $\lambda(a)$ returns the *leftmost* vertex in the row containing a.

Now define $\delta(a) = \rho(a - \lambda(a))$ where the subtraction, by convention, is performed in V_n. If we superimpose H_n on a triangular grid, $\delta(a)$ represents the *displacement* of a from the leftmost vertex in its row.

Finally we define $\varphi(a) = (\rho(a), \delta(a))$. We are going to prove that φ is a bijective, adjacency-preserving mapping from H_n onto P_n; in other words, we will establish the following:

THEOREM. *For all positive integers n, φ: $H_n \to P_n$ is a graph isomorphism.*

We will refer to φ (together with its inverse) as the *Lucas Correspondence*. Our proof of this result involves a happy coincidence: To deal with P_n, we will make use of Lucas' Theorem on binomial coefficients modulo a prime [42] while the graph H_n comes to us via Lucas' alter ego, "Professor Claus."

LEMMA 1 (Lucas' Theorem). *Let p be a prime and let*

$$r = r_m p^m + \cdots r_1 p + r_0 \qquad (0 \le r_i < p),$$
$$k = k_m p^m + \cdots k_1 p + k_0 \qquad (0 \le k_i < p).$$

Then

$$\binom{r}{k} \equiv \prod_{i=0}^{m} \binom{r_i}{k_i} \pmod{p}.$$

For a short proof of this result, see Fine's note [26].

In our setting, we are only interested in the case $p = 2$ so that $r_i, k_i \in \{0,1\}$ for all i. As an immediate consequence, we see that $\binom{r}{k} \equiv 1 \pmod 2$ if, and only if, $k_i \le r_i$ for all i. Thus we have the well-known result that, for a fixed r, the number of odd binomial coefficients $\binom{r}{k}$ is equal to $2^{b(r)}$ where $b(r)$ is the number of ones in the binary expansion of r.

We require three additional lemmas. Recall that T is the mapping that translates vertices of H_n one block in the clockwise direction.

LEMMA 2. *T is adjacency-preserving on the blocks.*

Proof. To see this, let $a = a_{n-1} \cdots a_1 a_0 \in V(H_n)$. Now the leading bit a_{n-1} corresponds to the largest disk and determines the block to which a belongs. Since

$T(a) = a + 12_{n-1}$, T acts as the rotation R^{-1} on $a_{n-2} \cdots a_1 a_0$. But R^{-1} is adjacency-preserving on H_{n-1} and, within a block in H_n, adjacent vertices have the same leading bit. (In terms of Tower of Hanoi configurations, T moves the largest disk one peg clockwise and all other disks one peg counterclockwise. So one configuration can be obtained from another by moving a single disk, other than the largest one, if, and only if, the same is true of the configurations that result from applying T.) Hence T is adjacency-preserving on the blocks.

Let us define a mapping $\sigma\colon H_n \to P_n$ to be *block-preserving* if for all $i \in \{0,1,2\}$, $a \in [i]$ in H_n if, and only if, $\sigma(a) \in [i]$ in P_n.

LEMMA 3. *φ is block-preserving.*

Proof. Let $a = \sum_{i=0}^{n-1} a_i \beta_i \in V(H_n)$.

Case 0. $a \in [0]$ if, and only if, $a_{n-1} = 0$ in which case

$$\rho(a) = \sum_{i=0}^{n-1} a_i^2 2^i \leq 1 + 2 + 2^2 + \cdots + 2^{n-2} = 2^{n-1} - 1,$$

and so by definition $\varphi(a) = (\rho(a), \delta(a)) \in [0]$.

Case 1. $a \in [1]$ if, and only if, $a_{n-1} = 2$ whence $2^{n-1} \leq \rho(a) \leq 2^n - 1$. We also have $\lambda(a)_{n-1} = 2^2 = 1$ and so $(a - \lambda(a))_{n-1} = 2 - 1 = 1$. It follows that $\delta(a) = \rho(a - \lambda(a)) \geq 2^{n-1}$ and therefore $\varphi(a) \in [1]$.

Case 2. $a \in [2]$ if, and only if, $a_{n-1} = 1$ and so again we have $2^{n-1} \leq \rho(a) \leq 2^n - 1$. However, here $\lambda(a)_{n-1} = 1^2 = 1$ and so $(a - \lambda(a))_{n-1} = 0$. Thus $\delta(a) = \rho(a - \lambda(a)) < 2^{n-1}$ and therefore $\varphi(a) \in [2]$.

The three converse implications follow immediately.

This result will be considerably strengthened in our main theorem where we will prove that φ is in fact adjacency-preserving. The next lemma shows that the Lucas correspondence φ is compatible with the translation mappings T and $\hat{T}$ of H_n and P_n, respectively. The formal definition of $\hat{T}$ was given in (3).

LEMMA 4. $\varphi \circ T = \hat{T} \circ \varphi$.

Proof. Let $a = \sum_{i=0}^{n-1} a_i \beta_i \in V(H_n)$, $a_i \in \mathbf{Z}_3$, and suppose $a \in [0]$. Then by Lemma 3, $\varphi(a) \in [0]$ in P_n and we have

$$\begin{aligned}\hat{T} \circ \varphi(a) &= \hat{T}((\rho(a), \delta(a))) \\ &= ((\rho(a) + 2^{n-1}, \delta(a) + 2^{n-1})).\end{aligned}$$

On the other hand, $a_{n-1} = 0$ so $a + 12_{n-1} = a + 2\beta_{n-1} = \sum_{i=0}^{n-2} a_i \beta_i + 2\beta_{n-1}$. Hence

$$\rho(a + 12_{n-1}) = \sum_{i=0}^{n-1} a_i^2 2^i = \sum_{i=0}^{n-2} a_i^2 2^i + 2^{n-1} = \rho(a) + 2^{n-1}.$$

We also have

$$\lambda(a + 12_{n-1}) = \sum_{i=0}^{n-1} a_i^2 \beta_i = \sum_{i=0}^{n-2} a_i^2 \beta_i + \beta_{n-1} = \lambda(a) + \beta_{n-1}$$

so that

$$\begin{aligned}\delta(a+1 2_{n-1}) &= \rho((a+1 2_{n-1}) - \lambda(a+1 2_{n-1}))\\ &= \rho(a+2\beta_{n-1} - (\lambda(a)+\beta_{n-1}))\\ &= \rho(a-\lambda(a)+\beta_{n-1})\\ &= \rho(a-\lambda(a)) + 2^{n-1} \text{ (as above)}\\ &= \delta(a) + 2^{n-1}.\end{aligned}$$

Therefore

$$\begin{aligned}\varphi\circ T(a) &= \varphi(a+1 2_{n-1})\\ &= (\rho(a+1 2_{n-1}), \delta(a+1 2_{n-1}))\\ &= ((\rho(a)+2^{n-1}, \delta(a)+2^{n-1}))\\ &= \hat{T}\circ\varphi(a).\end{aligned}$$

The cases for $a \in [1]$ or $[2]$ are handled similarly.

Proof of the theorem. Let β be the basis of V_n determined above, let $a \in V(H_n)$, and write $a = a_{n-1}\beta_{n-1} + \cdots + a_1\beta_1 + a_0\beta_0$, $a_i \in \mathbf{Z}_3$. We first must show that we indeed have $\varphi(a) \in V(P_n)$.

Now from the definition of λ we have

$$\lambda(a)_i = a_i^2 = \begin{cases}0\\1\\1\end{cases} \text{ if } a_i = \begin{cases}0\\1\\2\end{cases} \tag{4}$$

whence

$$(a-\lambda(a))_i = a_i - \lambda(a)_i = \begin{cases}0\\0\\1\end{cases} \text{ if } a_i = \begin{cases}0\\1\\2\end{cases}. \tag{5}$$

Now

$$\rho(a) = a_{n-1}^2 2^{n-1} + \cdots + a_1^2 2 + a_0^2$$

and

$$\delta(a) = \rho(a-\lambda(a)) = \left(a_{n-1} - a_{n-1}^2\right)^2 2^{n-1} + \cdots + \left(a_1 - a_1^2\right)^2 2 + \left(a_0 - a_0^2\right)^2$$

so that, by Lucas' Theorem, we have

$$\begin{aligned}\binom{\rho(a)}{\delta(a)} &\equiv \prod_{i=0}^{n-1} \binom{a_i^2}{\left(a_i - a_i^2\right)^2} \pmod 2\\ &\equiv 1 \pmod 2\end{aligned}$$

using (4) and (5). Thus $\varphi(a) = (\rho(a), \delta(a)) \in V(P_n)$ as required.

To show that φ is $1-1$, suppose that $a, b \in V(H_n)$, $a \neq b$. Let $a = \sum_{i=0}^{n-1} a_i\beta_i$ and $b = \sum_{i=0}^{n-1} b_i\beta_i$. If $\rho(a) \neq \rho(b)$ then $\varphi(a) \neq \varphi(b)$ and we are done. Suppose therefore that $\rho(a) = \rho(b)$. Then

$$\sum_{i=0}^{n-1} a_i^2 2^i = \sum_{i=0}^{n-1} b_i^2 2^i$$

or, in terms of the binary representations of $\rho(a)$ and $\rho(b)$,

$$\left(a_{n-1}^2 \cdots a_1^2 a_0^2\right) = \left(b_{n-1}^2 \cdots b_1^2 b_0^2\right).$$

It follows that $a_i^2 = b_i^2$ for $0 \le i \le n-1$. Therefore, for each i either $a_i = b_i$ or $a_i = 2b_i$ in $\mathbf{Z}_3$. Since $a \neq b$, there exists some j such that $a_j \neq b_j$. So, interchanging a and b if necessary, we may assume that $a_j = 1$ and $b_j = 2$. Now $\lambda(a)_j = a_j^2 = 1$ and $\lambda(b)_j = b_j^2 = 1$ so $(a - \lambda(a))_j = a_j - a_j^2 = 0$ while $(b - \lambda(b))_j = b_j - b_j^2 = 1$. But $(a - \lambda(a))_j^2 = 0$ and $(b - \lambda(b))_j^2 = 1$ are just the jth bits in the binary representations of $\delta(a)$ and $\delta(b)$, respectively. Therefore $\delta(a) \neq \delta(b)$ and consequently $\varphi(a) \neq \varphi(b)$. Hence φ is one-to-one.

The fact that φ is onto is an immediate consequence of the fact that $|V(H_n)| = |V(P_n)| = 3^n$, which we noted above. However, we can easily give a direct proof that φ is onto and thereby obtain this equality of cardinalities as a corollary.

To this end, let $(r, k) \in P_n$ (that is, $\binom{r}{k} \equiv 1 \pmod 2$) and write

$$r = \sum_{i=0}^{n-1} r_i 2^i, \quad k = \sum_{i=0}^{n-1} k_i 2^i$$

where $r_i, k_i \in \{0, 1\}$ for $0 \le i \le n-1$ and, by Lucas' Theorem,

$$\binom{r_i}{k_i} \equiv 1 \pmod 2, \; 0 \le i \le n-1.$$

For a nonnegative integer m with binary representation $m_t \cdots m_1 m_0$ define

$$\tau(m) = \sum_{i=0}^{t} m_i \beta_i.$$

Set

$$a = \tau(r) + \tau(k) = \sum_{i=0}^{n-1} (r_i + k_i)\beta_i$$

where, as usual, we perform the additions bitwise modulo 3.

We claim that $\varphi(a) = (r, k)$. First observe that, by Lucas' Theorem,

$$k_i = \begin{cases} 0 & \text{if } r_i = 0 \\ 0 \text{ or } 1 & \text{if } r_i = 1 \end{cases}$$

so that

$$a_i = (r_i + k_i) \pmod 3 = \begin{cases} 0 & \text{if } r_i = 0 \\ 1 \text{ or } 2 & \text{if } r_i = 1 \end{cases}$$

and thus

$$a_i^2 = \left\{ \begin{matrix} 0 & \text{if } r_i = 0 \\ 1 & \text{if } r_i = 1 \end{matrix} \right\} = r_i.$$

Therefore

$$\rho(a) = \sum_{i=0}^{n-1} a_i^2 2^i = \sum_{i=0}^{n-1} r_i 2^i = r.$$

Next observe that

$$\lambda(a) = \sum_{i=0}^{n-1} a_i^2 \beta_i = \sum_{i=0}^{n-1} r_i \beta_i = \tau(r)$$

so that, using the definition of a,

$$\rho(a-\lambda(a))=\rho(a-\tau(r))=\rho(\tau(k))=k,$$

where the last equality follows from the observation that the composition $\rho\circ\tau$ is the identity mapping on $\{0,1,2,\dots,2^n-1\}$. Putting all of this together, we see that

$$\varphi(a)=(\rho(a),\delta(a))=(\rho(a),\rho(a-\lambda(a)))=(r,k)$$

as claimed. Hence φ is onto.

Finally, we must prove that φ is adjacency-preserving; that is, a is adjacent to b in H_n if, and only if, $\varphi(a)$ is adjacent to $\varphi(b)$ in P_n. We will use induction on n. For $n=1$ the result is clear. Assume the result for $n<m$ and consider the case $n=m$.

Let a be adjacent to b in H_m. If a and b lie in the same block then they are adjacent vertices in an isomorphic copy of H_{m-1}. More precisely, we have $T^i(a)$ adjacent to $T^i(b)$ for some $i\in\{0,1,2\}$. By the induction hypothesis, $\varphi(T^i(a))$ is adjacent to $\varphi(T^i(b))$ in P_{m-1}. But by Lemma 4, this implies that $\hat{T}^i(\varphi(a))$ is adjacent to $\hat{T}^i(\varphi(b))$ which, since $\hat{T}$ is adjacency-preserving, in turn implies that $\varphi(a)$ is adjacent to $\varphi(b)$.

If a and b lie in distinct blocks but are adjacent, there are only three possibilities: $\{a,b\}$ is $\{01_{m-1},21_{m-1}\}$, $\{02_{m-1},12_{m-1}\}$ or $\{10_{m-1},20_{m-1}\}$. But it is straightforward to check that then $\{\varphi(a),\varphi(b)\}$ is $\{(2^{m-1}-1,0),(2^{m-1},0)\}$, $\{(2^{m-1}-1,2^{m-1}-1),(2^{m-1},2^{m-1})\}$ or $\{(2^m-1,2^{m-1}-1),(2^m-1,2^{m-1})\}$, respectively. Clearly $\varphi(a)$ and $\varphi(b)$ are adjacent in each case.

Now suppose a and b are not adjacent. If they lie in the same block then, by an inductive argument similar to that used above, $\varphi(a)$ and $\varphi(b)$ are not adjacent. On the other hand, if a and b lie in distinct blocks of H_m, then $\varphi(a)$ and $\varphi(b)$ lie in distinct blocks of P_m by Lemma 3. Since φ is one-to-one, $\{\varphi(a),\varphi(b)\}$ cannot be any of the three pairs corresponding to edges between blocks. Hence $\varphi(a)$ and $\varphi(b)$ are not adjacent in this case either.

This completes the proof of the theorem.

We should note at this point that if one simply wishes to know that H_n and P_n *are* isomorphic, this can easily be established by induction; we leave this as an exercise or refer the interested reader to [35]. However, to be able to solve problems related to the Tower of Hanoi we need to know the Lucas Correspondence explicitly and hence we have elected to give a constructive proof of the main theorem.

The Lucas Correspondence allows us to pass information between H_n and P_n. Some immediate consequences are summarized in the next two corollaries while in the next section we look at a variety of Tower of Hanoi problems.

COROLLARY 1. *For all positive integers n, $|V(P_n)|=3^n$.*

(It is an amusing coincidence that the preceding corollary was posed as Problem 1222 in this MAGAZINE [58] and 1222 appears as one of the elements of the basis β in our description of the Lucas Correspondence.)

COROLLARY 2. *The number of legal Tower of Hanoi configurations whose distance from a perfect state is r is $2^{b(r)}$ where $b(r)$ is the number of ones in the binary expansion of r.*

Proof. This follows immediately from the comments after Lucas' Theorem and from symmetry. (See also Problem 1 in the next section.)

In [61], Walsh gave an elementary method for detecting errors in a move sequence for the standard Tower of Hanoi problem (that is, transferring disks from one perfect configuration to another). In other words, we would like to be able to determine, by inspection, whether a given configuration occurs in the unique minimal move sequence between perfect configurations. This is just the problem of determining whether a given vertex occurs on the "boundary" of H_n. Using the Lucas Correspondence, this is almost a triviality.

COROLLARY 3. *In H_n, vertex a occurs on the shortest path between* **0** *and* **2**, **0** *and* **1**, *or* **1** *and* **2** *if, and only if,* $\delta(a) = 0$, $\delta(a) = \rho(a)$, *or* $\rho(a) = 2^n - 1$, *respectively.*

Solving Tower of Hanoi problems

A great many Tower of Hanoi problems are simply shortest-path problems for the graph H_n. Many of these can be handled efficiently using the Lucas Correspondence. We can also exploit the symmetry of the Hanoi graphs and we have a formal mechanism for doing so. In the problems that follow, if a and b are vertices of a graph G, $\mathrm{d}(a, b)$ will denote the distance from a to b (that is, the length of a shortest path from a to b in G).

PROBLEM 1. *The generalized Tower of Hanoi problem.*

This is the rubric used by Er [20] to refer to the problem of finding the minimum number of moves required to transfer the disks from an arbitrary legal configuration to a single peg using the standard rules. In our model, this translates into the problem of finding the distances $\mathrm{d}(a, \mathbf{0})$, $\mathrm{d}(a, \mathbf{1})$, and/or $\mathrm{d}(a, \mathbf{2})$ given an arbitrary vertex a of H_n.

Now an easy induction shows that, in P_n, every vertex other than $(0, 0)$ is adjacent to a vertex in the row above. It follows that the distance to $(0, 0)$ from any vertex in row r is r. So, using the Lucas Correspondence, we see that

$$\mathrm{d}(a, \mathbf{0}) = \rho(a)$$

for any vertex a of H_n and, using symmetry,

$$\mathrm{d}(a, \mathbf{1}) = \mathrm{d}(a + \mathbf{2}, \mathbf{0}) = \rho(a + \mathbf{2}) \text{ and } \mathrm{d}(a, \mathbf{2}) = \mathrm{d}(a + \mathbf{1}, \mathbf{0}) = \rho(a + \mathbf{1}).$$

We can also obtain the corresponding shortest path in each case (that is, the sequence of intermediate Tower of Hanoi positions between the initial and target configurations). On the graph H_n, start at **0**. Let i go from $n - 1$ to 0. Examine the binary representations of $\rho(a)$ and $\delta(a)$. If the ith digit of $\rho(a)$ is a 1, then we know by Lucas' Theorem that the corresponding digit of $\delta(a)$ is either a 0 or a 1: if it is 0, move 2^i vertices to the left, and if it is 1, move 2^i vertices to the right (horizontal edges are not used). This gives the shortest path from **0** to a; reversing it gives the solution to the generalized Tower of Hanoi problem.

In the Appendix, we give a *Mathematica* procedure that implements this implicit algorithm.

Example 1. Consider the case $n = 5$ and $a = 20201$. We compute

$$20201 = 21111 + 2111 + 12 \pmod 3$$

so that

$$20201 = \beta_4 + 2\beta_3 + \beta_1$$

and so the binary representation of $\rho(20201)$ is 11010. Converting to decimal, we have $\rho(20201) = 26$, which is the distance from 20201 to 00000. The binary representation of $\delta(20201)$ is obtained by subtracting the ith binary digit of $\rho(20201)$ from the coefficient of β_i (and then squaring—but this does not change the result since $0^2 = 0$ and $1^2 = 1$). Thus the binary representation of $\delta(20201)$ is 01000 so that $\delta(20201) = 8$.

The shortest path from $\mathbf{0}$ to a then travels 16 vertices left, 8 vertices right and 2 vertices left (see FIGURE 5). The path obtained by the reversal of this is given in the Appendix.

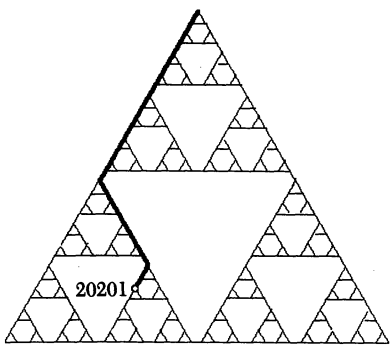


FIGURE 5

This method is also given in [5] but without proof and only in the case of the standard Tower of Hanoi problem of moving from one perfect configuration to another.

Example 2. A more general example that can be handled by this technique involves colouring the rings. Consider a standard set of $3n$ disks, initially stacked on peg 0. Starting from the smallest disk, colour the disks red, white, blue, red, white, blue, and so on, ending with the largest disk coloured blue and n disks of each colour. We wish to find the minimum number of moves needed to obtain the red disks stacked legally on peg 0, white on peg 1, and blue on peg 2.

This is just the problem of determining, as a function of n, $\mathrm{d}(\mathbf{0}, a)$ where $a = 210210 \cdots 210$. We compute

$$\begin{aligned} &210210 \cdots 210 \\ =\ &211111 \cdots 111 \\ &+2111 \cdots 111 \\ &+21 \cdots 111 \\ &\quad\vdots \\ &+21, \end{aligned}$$

if n is even and

$$\begin{aligned} &210210 \cdots 210 \\ =\ &211111 \cdots 111 \\ &+2111 \cdots 111 \\ &+21 \cdots 111 \\ &\quad\vdots \\ &+2, \end{aligned}$$

if n is odd. Hence $\rho(a)$ has binary representation $1010\cdots 10$ in the first case and $1010\cdots 101$ in the second. Converting to decimal we obtain

$$\rho(a)=\begin{cases}2+8+\cdots+2^{3n-1}=\frac{2}{3}(2^{3n}-1), & n \text{ even}\\ 1+4+\cdots+2^{3n-1}=\frac{1}{3}(2^{3n+1}-1), & n \text{ odd,}\end{cases}$$

which answers the question.

(As an exercise show that, in the last example, there is a configuration with the red disks on one peg, white disks on another peg and blue disks on a third peg whose distance from the initial configuration is $2^{3n}-1$ moves—the maximum possible—and that this is independent of the parity of n.)

In **[50]** there is another example along these lines that can easily be handled by the technique given here.

PROBLEM 2. *Average number of moves.*

This problem was considered by Er **[19]** in his analysis of Problem 1 and by Scarioni and Speranza **[53]** as part of their analysis of an error-correcting algorithm for the standard Tower of Hanoi problem. We can obtain this result directly from the graph H_n. Letting E_n denote the expected, or average, distance of a vertex from $\mathbf{0}$ and $S_n=\Sigma\{\mathrm{d}(a,\mathbf{0})\,|\,a\in H_n\}$, we obtain the following recurrence relation:

$$S_n=3S_{n-1}+2\cdot 3^{n-1}\cdot 2^{n-1}$$

(where the last term arises from translating each vertex in blocks [1] and [2] 2^{n-1} vertices upwards so it corresponds with a vertex in block [0]). We then have

$$\begin{aligned}E_n&=\frac{S_n}{3^n}\\ &=\frac{S_{n-1}}{3^{n-1}}+\frac{2}{3}2^{n-1}\\ &=E_{n-1}+\tfrac{2}{3}2^{n-1},\end{aligned}$$

which, with the initial condition $E_0=0$, is easily seen to have the solution

$$E_n=\tfrac{2}{3}(2^n-1).$$

In other words, the average number of moves required to (optimally) transfer the disks from a randomly chosen initial configuration to a perfect configuration on a specified target peg is 2/3 the number of moves required by the optimal solution (2) to the standard problem.

Alternatively, fix a vertex/configuration a and ask for the average distance/number of moves $A_n(a)$ to a randomly chosen corner vertex/perfect configuration. But clearly

$$\begin{aligned}A_n(a)&=\tfrac{1}{3}(\mathrm{d}(a,\mathbf{0})+\mathrm{d}(a,\mathbf{1})+\mathrm{d}(a,\mathbf{2}))\\ &=\tfrac{1}{3}(\mathrm{d}(a,\mathbf{0})+\mathrm{d}(a+\mathbf{2},\mathbf{0})+\mathrm{d}(a+\mathbf{1},\mathbf{0}))\\ &=\tfrac{1}{3}(\rho(a)+\rho(a+\mathbf{2})+\rho(a+\mathbf{1})).\end{aligned}\tag{6}$$

Now $\mathbf{1}=2\Sigma_{i=0}^{n-1}\beta_i$ and $\mathbf{2}=\Sigma_{i=0}^{n-1}\beta_i$ in V_n so (6) is equal to

$$\frac{1}{3}\left(\sum_{i=0}^{n-1}a_i^2 2^i+\sum_{i=0}^{n-1}(a_i+1)^2 2^i+\sum_{i=0}^{n-1}(a_i+2)^2 2^i\right).$$

But for all $x \in \mathbf{Z}_3$, $x^2 + (x+1)^2 + (x+2)^2 = 2$; this in turn gives

$$A_n(a) = \frac{2}{3} \sum_{i=0}^{n-1} 2^i$$
$$= \tfrac{2}{3}(2^n - 1).$$

Thus $A_n(a)$ is independent of a, can be computed without solving a recurrence relation, and is equal to E_n.

PROBLEM 3. *The all-pairs shortest path problem.*

This problem, a further generalization of Problem 1, involves finding the minimum number of moves needed to go from one legal Tower of Hanoi configuration to another. In terms of the Hanoi graph, this is just the problem of finding a shortest path between all possible pairs of vertices—the so-called all-pairs shortest path problem for H_n. Of course, there are efficient algorithms for solving this problem for an arbitrary graph—for example, Dijkstra's Algorithm or Floyd's Algorithm can be used—but H_n is such a specialized graph that it ought to be possible to avoid "using a cannon to shoot a fly."

Several authors have approached this problem, directly and indirectly, with mixed success. The argument that has been used often goes something like this: Consider two configurations, a and b, and suppose the largest disk is on peg i in a and peg j in b, $i \neq j$. Then from a we must first move all disks but the largest to peg $k \neq i, j$, then move the largest disk from i to j, and finish by redistributing the disks from k to obtain the target configuration b. This argument, which appears in **[12]** and **[65]**, implicitly assumes that the largest disk moves at most once, which, if correct, leads to an inductive proof that the shortest path from a to b is unique. Unfortunately this is false! A glance at H_n shows why; for $n = 3$, there are two shortest paths from 001 to 010, namely {001, 002, 012, 010} and {001, 021, 020, 010} and in the unique shortest path from 011 to 100—{011, 211, 212, 202, 200, 100}—the largest disk moves twice (see FIGURE 2). It is not difficult to prove that, between any pair of vertices of H_n, there are at most two shortest paths. But, as Hinz notes in **[33]** and **[36]**, it does not seem to be an easy matter to decide whether the largest disk that moves is in fact moved once or twice. In terms of the graph, if a and b lie in different blocks of H_n, the problem is to decide whether a shortest path includes vertices in just these two blocks or passes through the third block as well. Once this is known, the actual path can be found using any of the existing algorithms for solving Problem 1.

We can use the Lucas Correspondence to resolve this problem by providing a straightforward means of determining which of the two possible routes a shortest path in H_n will take. Let a_1 and a_2 be distinct vertices of H_n. Without loss of generality, we may take them to lie in different blocks, say $a_1 \in [0]$ and $a_2 \in [1]$. Let $\varphi(a_1) = (r_1, k_1)$ and $\varphi(a_2) = (r_2, k_2)$. Let $\mathrm{d}_1 = \mathrm{d}_1(a_1, a_2)$ denote the length of the shortest path from a_1 to a_2 that does not pass through [2] and let $\mathrm{d}_2 = \mathrm{d}_2(a_1, a_2)$ denote the length of the shortest path from a_1 to a_2 that does pass through [2]. Then we compute

$$\begin{aligned} \mathrm{d}_1 &= \mathrm{d}\big((r_1, k_1), (2^{n-1} - 1, 2^{n-1} - 1)\big) + \mathrm{d}\big((r_2, k_2), (2^{n-1}, 2^{n-1})\big) + 1 \\ &= \big((2^{n-1} - 1 - r_1) + (r_1 - k_1)\big) + \big((r_2 - 2^{n-1})\big) + 1 \\ &= r_2 - k_1 \end{aligned} \qquad (7)$$

and

$$\begin{aligned} d_2 &= d((r_1,k_1),(2^{n-1}-1,0)) + d((r_2,k_2),(2^n-1,2^{n-1})) + 2^{n-1}+1 \\ &= ((2^{n-1}-1-r_1)+k_1) + ((2^n-1-r_2)+(k_2-2^{n-1})) + 2^{n-1}+1 \\ &= 3\cdot 2^{n-1} - 1 - r_1 - r_2 + k_1 + k_2. \end{aligned} \tag{8}$$

Note that what we are counting here are the numbers of horizontal and diagonal edges in each of the two paths. (See FIGURE 6 for a schematic representation of this.)

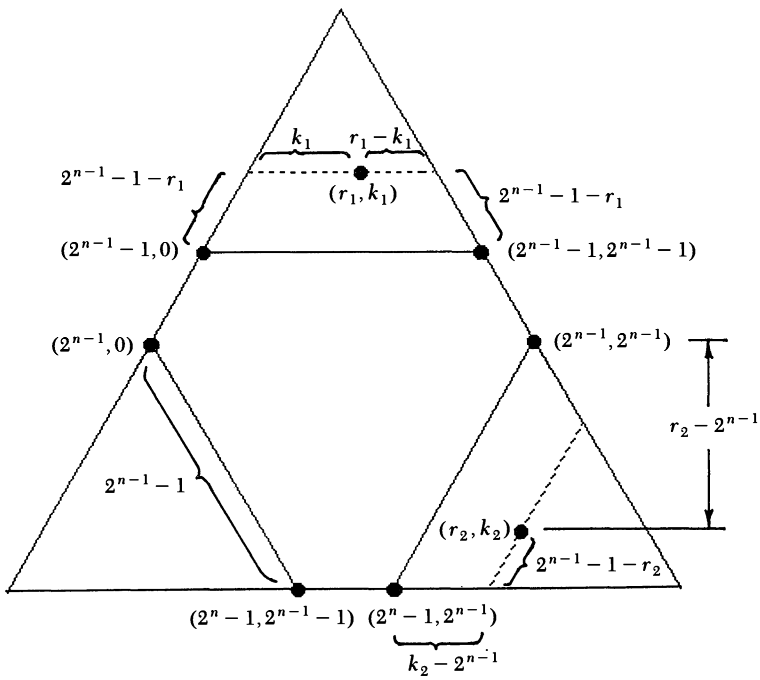


FIGURE 6

Thus we see that

$$d_2 \le d_1 \text{ if, and only if, } r_1 + 2r_2 \ge 2k_1 + k_2 + 3\cdot 2^{n-1} - 1.$$

Using symmetry or by direct calculation, similar results can be obtained for a_1 and a_2 in [0] and [2] or in [1] and [2].

Example 3. Find d(011, 102) in H_3. The Lucas Correspondence gives

$$\varphi(011) = (3,0) \text{ and } \varphi(102) = (7,5)$$

from which, using (7) and (8), we calculate

$$d_1(011,102) = 7-0 = 7 \text{ and } d_2(011,102) = 11-3-7+0+5 = 6.$$

Hence d(011, 102) = 6 and the shortest path passes through all three blocks of H_3 (see FIGURE 2 and the Appendix).

It should be noted that, even without the decision procedure we have given here, much statistical information can be obtained about shortest paths in H_n. For example, it can be shown that the average length of a shortest path between vertices in H_n is, asymptotically, $466/885$ of the diameter of the graph ([8], [33]).

Variations and some open problems

One of the factors that has undoubtedly contributed to the longevity of the Tower of Hanoi problem is that there is a seemingly endless number of variations on the same theme. The rules governing the movement and placement of the disks can be either strengthened or relaxed, the number of disks of each size can be increased and we can view these multiple disks as distinguishable or indistinguishable, the number of pegs can be increased, and so on.

1. *Straightline Hanoi.* One easy variation is to modify the usual rules by adding the restriction that a disk may only be moved to or from peg 1 **[54]**, a variation that is called the Straightline Tower of Hanoi in **[47]**. It is not difficult to prove that the minimum number of moves needed to transfer a tower of n disks from peg 0 to peg 2 is then $3^n - 1$ and so corresponds to a Hamiltonian path from $\mathbf{0}$ to $\mathbf{2}$ in H_n the existence of which was established in Lemma 0. (In fact, this gives an alternative proof of Lemma 0.)

2. *Cyclic Hanoi.* In **[3]**, Atkinson introduced the cyclic Tower of Hanoi problem in which the pegs are arranged in a circle and the disks are restricted to move one peg at a time in a clockwise direction. He presented elementary recursive solutions for c_n, the number of moves needed to transfer a tower of n disks one peg in the clockwise direction, and a_n, the number of moves needed to transfer the same tower one peg in the counter clockwise direction. It is not difficult to prove that

$$c_n = \frac{1}{2\sqrt{3}}\left((1+\sqrt{3})^{n+1} - (1-\sqrt{3})^{n+1}\right) - 1$$

and

$$a_n = \frac{1}{4\sqrt{3}}\left((1+\sqrt{3})^{n+2} - (1-\sqrt{3})^{n+2}\right) - 1.$$

(See also **[21]**, **[23]**, **[24]**, and **[62]**.)

3. *Rainbow Hanoi.* Other variants of the Tower of Hanoi involve colouring the disks and, unlike our Example 2 above, using the colours to restrict the movement of the disks. Having disks of different colours move in different directions is treated by Er in **[22]** and forbidding disks of the same colour to come into contact with one another is discussed by Minsker in **[48]** where it is called the Rainbow Tower of Hanoi problem.

4. *Multi-disk Hanoi.* In **[64]**, Wood posed two variations. In one, the restriction that a disk may not be placed on top of a smaller one is relaxed to allow disk i to be placed on top of disk j if, and only if, $i - j \le b$ for some fixed integer $b \le n-1$. In the other, there are d copies of each disk and the usual rules apply. (In fact we get two versions of this depending on whether we treat the copies as distinguishable or indistinguishable.) A generalization of this in which there are d_i copies of disk i is given in **[30**, Exercise 1.12**]**.

5. *Multi-peg Hanoi.* Perhaps the greatest challenge is posed by increasing the number of pegs. In **[16]**, Dudeney gave a version with four pegs (actually he used cheeses and stools instead of disks and pegs) and called it "The Reve's Puzzle." Dudeney himself undoubtedly knew the answer in this case and, since then, several other authors have published "solutions" (**[6]**, **[27]**, **[52]**, **[57]**) or given algorithms for

"solving" the problem ([25], [34], [40], [41], [51]). However, no one has yet succeeded in proving the *minimality* of these solutions and this is apparently a very difficult problem (cf. [10], the Afterword to the revised edition of [28], [30, Exercise 1.25], [45], and the Editorial Note following [57]).

Open problems:

(1) Is the Lucas Correspondence of any help in variants 1–3?
(2) Construct and analyze analogues of the Hanoi graph H_n in variants 4 and 5. Is there a generalization of the Lucas Correspondence that works here?

Conclusion

Edouard Lucas could hardly have foreseen the attention his Tower of Hanoi puzzle would receive in the century or so since he introduced it. Mathematicians and computer scientists, both amateur and professional, have been kept quite busy by the problems it has spawned. The number of papers that have been written about the Tower of Hanoi and its offspring is testament to the fact that this little puzzle holds a seemingly neverending supply of challenges.

As we have seen here, the Hanoi graph H_n provides a convenient model for visualizing and analyzing Tower of Hanoi problems. The Lucas Correspondence of H_n with the graph P_n of the odd binomial coefficients then gives us an effective computational tool for dealing with shortest path problems in H_n. It comes as a pleasant surprise that Lucas' investigation of binomial coefficients modulo a prime can be used to develop and prove this useful correspondence.

While it is highly unlikely that he knew of this connection between H_n and P_n, we would like to think that it would have pleased "Professor Claus."

Appendix

The following *Mathematica* procedures implement many of the ideas contained in this paper.

We first generate, in order, all 2^n n-bit binary strings that we view as the coefficients with respect to β (in *reverse* order) of the leftmost vertices of H_n. This function is taken from [60].

```
binstrings[n_] := Flatten[Apply[Outer,
      Prepend[Table[{0, 1}, {n}], List]], n - 1]
```

Next we construct the change-of-basis matrix **M** from β to the standard basis for V_n. (Exercise: Verify that, for all n, this gives the correct matrix and that $\mathbf{M} = \mathbf{M}^{-1}$ over $\mathbf{Z}_3$.)

```
basismat[n_] := Block[{M},
  M = Table[0, {n}, {n}];
  Do[ M[[j, j]] = PowerMod[2, j, 3];
      M[[i, j]] = PowerMod[2, j - 1, 3],
      {j, 1, n - 1}, {i, j + 1, n}];
  M[[n, n]] = PowerMod[2, n, 3];
  M]
```

Now we use **M** to generate n-bit ternary strings as needed.

```
ternrep[a_] := Block[{n, M},
n = Dimensions[a][[1]];
M = basismat[n];
Mod[M.a, 3] ]
```

Here is the Lucas Correspondence: $r = \rho(a)$ and $k = \delta(a)$.

```
LucasMap[a_] := Block[{n, t, r, k},
   n = Dimensions[a][[1]];
   t = ternrep[a];
   r = Sum[PowerMod[t[[n - i]], 2, 3] 2^i, {i, 0, n - 1}];
   k = Sum[t[[n - i]] 2^i, {i, 0, n - 1}] - r;
   {r, k} ]
```

The next procedure generates the unique path from a to $\mathbf{0}$. The path is actually generated in reverse order by iteratively adding to the string labelling the "top" vertex of a triangle a sequence of 2^i consecutive ternary strings from either the left or right sides of H_n: If the ith bit t_i of the ternary representation of a with respect to β is nonzero, move 2^i vertices left if $t_i = 1$, otherwise move 2^i vertices right.

```
movetower[a_] := Block[{n, top, m, t, b, moveseq},
   n = Dimensions[a][[1]];
   top = Table[0, {n}];

moveseq = {top};
M = basismat[n];
b = binstrings[n];
t = ternrep[a];

Do[ If[t[[i]] != 0,
    (Do[AppendTo[moveseq,
            Mod[top + t[[i]] M.b[[j + 1]], 3]],
        {j, 1, 2^(n - i)}];
    top = Last[moveseq])],
{i, n}];
Reverse[moveseq] ]
```

Using symmetry, we now obtain a procedure that solves the generalized Tower of Hanoi problem as described in Problem 1.

```
HanoiMoveSequence[a_, p_] :=
   Mod[movetower[Mod[a - p, 3]]+p, 3]
```

We now develop a procedure for solving the All-Pairs Tower of Hanoi problem, as described in Problem 3. We begin with a procedure that returns the shortest path between vertices a_1 and a_2 in blocks [0] and [1], respectively. Using symmetry, we then extend this to handle an arbitrary pair of vertices a_1 and a_2. Both procedures use a modified version of the built-in *Mathematica* function **Prepend**.

```
preprend[l_, i_] := Map[Prepend[#, i]&, l]

HanoiPath01[a1_, a2_] := Block[
   {n, r1, r2, k1, k2, d1, d2, one, path },
   n = Dimensions[a1][[1]];
   r1 = LucasMap[a1][[1]]; k1 = LucasMap[a1][[2]];
```

```
    r2 = LucasMap[a2][[1]]; k2 = LucasMap[a2][[2]];
    one = Table[1, {n - 1}];
    d1 = r2 - k1;
    d2 = 3 2^(n - 1) - 1 - r1 - r2 + k1 + k2;
    path = If[d1 <= d2,
    Join[prepend[HanoiMoveSequence[Rest[a1], 2], 0],
      Reverse[prepend[HanoiMoveSequence[Rest[a2], 2], 1]] ],
    Join[prepend[HanoiMoveSequence[Rest[a1], 1], 0],
      prepend[HanoiMoveSequence[one, 0], 2],
      Reverse[prepend[HanoiMoveSequence[Rest[a2], 0], 1]]]];
    path]
HanoiPath[a1_, a2_] := Block[
    {u1, u2, v1, v2, rotate, flip, path},
    u = a1; v = a2;
    If[u[[1]] == v[[1]],
        path = prepend[
        HanoiPath[Rest[u], Rest[v]], a1[[1]]],
        (rotate = a1[[1]];
        u1 = Mod[a1 - rotate, 3];
        u2 = Mod[a2 - rotate, 3];
        flip = u2[[1]];
        v1 = flip u1; v2 = flip u2;
        path = Mod[flip HanoiPath01[v1, v2] + rotate, 3])];
        path ]
```

Here is *Mathematica's* solution to Example 1.

```
LucasMap[{2, 0, 2, 0, 1}]
```

$\{26, 8\}$

```
HanoiMoveSequence[{2, 0, 2, 0, 1}, 0]
```

$\{\{2,0,2,0,1\},\{2,0,2,2,1\},\{2,0,2,2,2\},\{2,1,2,2,2\},\{2,1,2,2,1\},\{2,1,2,0,1\},$
$\{2,1,2,0,0\},\{2,1,1,0,0\},\{2,1,1,0,2\},\{2,1,1,1,2\},\{2,1,1,1,1\},\{0,1,1,1,1\},$
$\{0,1,1,1,2\},\{0,1,1,0,2\},\{0,1,1,0,0\},\{0,1,2,0,0\},\{0,1,2,0,1\},\{0,1,2,2,1\},$
$\{0,1,2,2,2\},\{0,0,2,2,2\},\{0,0,2,2,0\},\{0,0,2,1,0\},\{0,0,2,1,1\},\{0,0,0,1,1\},$
$\{0,0,0,1,2\},\{0,0,0,0,2\},\{0,0,0,0,0\}\}$

As a check that our formula from Problem 2 for the average path length to a corner vertex is correct, let's compute $A_5(20201)$:

```
a = {2, 0, 2, 0, 1}
Sum[LucasMap[Mod[a + k, 3]][[1]], {k, 0, 2}]/3
```

$\frac{62}{3}$,

which equals $\frac{2}{3}(2^5 - 1)$ as expected.

Finally, here is the shortest path from Example 3 in our discussion of the All-Pairs problem:

```
HanoiPath[{0, 1, 1}, {1, 0, 2}]
```

$\{\{0,1,1\},\{2,1,1\},\{2,1,2\},\{2,0,2\},\{2,0,0\},\{1,0,0\},\{1,0,2\}\}$

Acknowledgements. I would like to thank Andreas Hinz for his thoughtful and detailed comments on this paper and for drawing my attention to several references of which I was otherwise unaware. In his paper [35], Hinz gives a non-constructive proof of the Lucas Correspondence and uses it to derive results about P_n from knowledge of H_n, thereby complementing the results we have given here. I am also indebted to an anonymous referee whose comments greatly improved the presentation of this article.

REFERENCES

1. A. Aho, J. Hopcroft and J. Ullman, *The Design and Analysis of Computer Algorithms*, Addison-Wesley Publishing Co., Reading, MA, 1974.
2. R. E. Allardice and A. Y. Fraser, La Tour d'Hanoï, *Proc. Edinburgh Math. Soc.* 2 (1884), 50–53.
3. M. D. Atkinson, The cyclic Towers of Hanoi, *Inform. Process. Lett.* 13 (1981), 118–119.
4. W. W. Rouse Ball and H. S. M. Coxeter, *Mathematical Recreations and Essays*, 12th edition, University of Toronto Press, Toronto, 1974.
5. E. R. Berlekamp. J. H. Conway and R. K. Guy, *Winning Ways for Your Mathematical Plays*, Academic Press, London, 1982.
6. A. Brousseau, Tower of Hanoi with more pegs, *J. Recreational Math.* 8 (1976), 169–178.
7. P. Buneman and L. Levy, The Towers of Hanoi Problem, *Inform. Process. Lett.* 10 (1980), 243–244.
8. T.-H. Chan, A statistical analysis of the Towers of Hanoi problem, *Internat. J. Comput. Math.* 28 (1989), 57–65.
9. D. W. Crowe, The n-dimensional cube and the Tower of Hanoi, *Amer. Math. Monthly* 63 (1956), 29–30.
10. P. Cull and E. F. Ecklund, Jr., On the Towers of Hanoi and generalized Towers of Hanoi problems, *Congr. Numer.* 35 (1982), 229–238.
11. ______, Towers of Hanoi and analysis of algorithms, *Amer. Math. Monthly* 92 (1985), 407–420.
12. P. Cull and C. Gerety, Is Towers of Hanoi really hard? *Congr. Numer.* 47 (1985), 237–242.
13. H. de Parville, La Tour d'Hanoï et la question du Tonkin, *La Nature* 12 (1884), 285–286.
14. A. K. Dewdney, *The Armchair Universe*, W. H. Freeman, New York, 1986.
15. ______, *The Turing Omnibus*, Computer Science Press, New York, 1989.
16. H. E. Dudeney, *The Canterbury Puzzles and Other Curious Problems*, E. P. Dutton, New York, 1908; 4th edition, Dover Publications, Inc., Mineola, NY, 1958.
17. M. C. Er, A representation approach to the Tower of Hanoi Problem, *Comput. J.* 25 (1982), 442–447.
18. ______, An iterative solution to the generalized Towers of Hanoi problem, *BIT* 23 (1983), 295–302.
19. ______, An analysis of the generalized Towers of Hanoi problem, *BIT* 23 (1983), 429–435.
20. ______, The generalized Towers of Hanoi problem, *J. Inform. Optim. Sci.* 5 (1984), 89–94.
21. ______, The cyclic Towers of Hanoi: A representation approach, *Comput. J.* 27 (1984), 171–175.
22. ______, The generalized colour Towers of Hanoi: An iterative algorithm, *Comput. J.* 27 (1984), 278–282.
23. ______, A generalization of the cyclic Towers of Hanoi: An iterative solution, *Internat. J. Comput. Math.* 15 (1984), 129–140.
24. ______, The complexity of the generalised cyclic Towers of Hanoi problem, *J. Algorithms* 6 (1985), 351–358.
25. ______, An optimal algorithm for Reve's puzzle, *Inform. Sci.* 45 (1988), 39–49.
26. N. J. Fine, Binomial coefficients modulo a prime, *Amer. Math. Monthly* 54 (1947), 589–592.
27. J. S. Frame, Solution to Problem 3918, *Amer. Math. Monthly* 48 (1941), 216–217.
28. M. Gardner, *The Scientific American Book of Mathematical Puzzles and Diversions*, Simon and Schuster, New York, 1959; revised edition, *Hexaflexagons and Other Mathematical Diversions*, University of Chicago Press, Chicago, 1988.
29. ______, *Knotted Doughnuts and Other Mathematical Entertainments*, W. H. Freeman, New York, 1986.
30. R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete Mathematics*, Addison-Wesley Publishing Co., Reading, MA, 1989.
31. P. J. Hayes, A note on the Towers of Hanoi problem, *Comput. J.* 20 (1980), 243–244.
32. P. J. Hilton, private communication.
33. A. M. Hinz, The Tower of Hanoi, *Enseign. Math.* 35 (1989), 289–321.
34. ______, An iterative algorithm for the Tower of Hanoi with four pegs, *Computing* 42 (1989), 133–140.
35. ______, Pascal's Triangle and the Tower of Hanoi, *Amer. Math. Monthly* 99(1992), 538–544.
36. ______, Shortest paths between regular states of the Tower of Hanoi, *Inform. Sci.* 63(1992), 173–181.
37. R. L. Kruse. *Data Structures and Program Design*, Prentice-Hall, Englewood Cliffs, NJ, 1984.
38. I. Lavallée, Note sur le problème des Tours de Hanoï, *Rev. Roumaine Math. Pures Appl.* 30 (1985), 433–438.
39. X. M. Lu, Towers of Hanoi graphs, *Internat. J. Comput. Math.* 19 (1986), 23–38.
40. ______, Towers of Hanoi with arbitrary $k \geq 3$ pegs, *Internat. J. Comput. Math.* 24 (1988), 39–54.
41. ______, An iterative solution for the 4-peg Towers of Hanoi, *Comput. J.* 32 (1989), 187–189.
42. E. Lucas, Sur les congruences des nombres eulériens et les coefficients différentiels des fonctions trigonométriques suivant un module prémier, *Bull. Soc. Math. France* 6 (1878), 49–54.

43. ______, *Récréations Mathématiques*, Gauthier-Villars, Paris, 1891–1894.
44. ______, *Théorie des Nombres*, Gauthier-Villars, Paris, 1891.
45. W. F. Lunnon, The Reve's Puzzle, *Comput. J.* 29 (1986), 478.
46. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, San Francisco, 1982.
47. S. B. Maurer and A. Ralston, *Discrete Algorithmic Mathematics*, Addison-Wesley Publishing Co., Reading, MA, 1990.
48. S. Minsker, The Towers of Hanoi rainbow problem: Coloring the rings, *J. Algorithms* 10 (1989), 1–19.
49. R. B. Nelson, E. T. H. Wang and B. Poonen, Solution to Problem 1222, this MAGAZINE 59 (1986), 241–242.
50. H. Noland, problem 1350, this MAGAZINE 63 (1990), 189.
51. J. S. Rohl and T. D. Gedeon, The Reve's Puzzle, *Comput. J.* 29 (1986), 187–188; Corrigendum, *Ibid.* 31 (1988), 190.
52. T. Roth, The Tower of Brahma revisited, *J. Recreational Math.* 7 (1974), 116–119.
53. F. Scarioni and H. G. Speranza, A probabilistic analysis of an error-correcting algorithm for the Towers of Hanoi puzzle, *Inform. Process. Lett.* 18 (1984), 99–103.
54. R. S. Scorer, P. M. Grundy and C. A. B. Smith, Some binary games, *Math. Gaz.* 28 (1944), 96–103.
55. I. Stewart, Les fractals de Pascal, *Pour la Science* 129 (1988), 100–104.
56. ______, Le lion, la lama et la laitue, *Pour la Science* 142 (1989), 102–107.
57. B. M. Stewart, Solution to Problem 3918, *Amer. Math. Monthly* 48 (1941), 217–219.
58. M. Sved, Problem 1222, this MAGAZINE 58 (1985), 237.
59. ______, Divisibility—with visibility, *Math. Intelligencer* 10 (1988), 56–64.
60. S. Wagon, *Mathematica in Action*, W. H. Freeman, New York, 1991.
61. T. R. Walsh, The Towers of Hanoi revisited: Moving the rings by counting the moves, *Inform. Process. Lett.* 15 (1982), 64–67.
62. ______, Iteration strikes back—at the cyclic Towers of Hanoi, *Inform. Process. Lett.* 16 (1983), 91–93.
63. S. Wolfram, Geometry of binomial coefficients, *Amer. Math. Monthly* 91 (1984), 566–571.
64. D. Wood, The Towers of Brahma and Hanoi revisited, *J. Recreational Math.* 14 (1981), 17–24.
65. ______, Adjudicating a Towers of Hanoi contest, *Internat. J. Comput. Math.* 14 (1983), 199–207.

Math Bite: A Simple Proof That Every Sequence Has a Monotone Subsequence

Known proofs that every sequence of real numbers has a monotone subsequence seem to be, as one recent text [1] put it, "tricky." Here is a simple proof.

Obviously a sequence with no greatest member has an increasing subsequence.

Let s denote the sequence whose nth member is s_n.

If there is a number k such that the "tail" $s_k, s_{k+1}, s_{k+2}, \dots$ has no greatest member then this tail has an increasing, and consequently monotone, subsequence, which will be a subsequence of s. If there no such k, let s_p be a greatest member of s, s_q a greatest member of the tail $s_{p+1} s_{p+2}, \dots,$ let s_r be a greatest member of the tail $s_{q+1}, s_{q+2}, \dots$ and so on. Then $s_p \geq s_q \geq s_r \dots$ and we have a monotone subsequence.

REFERENCE

1. R. P. Burn, *Numbers and Functions: Steps into Analysis*, Cambridge University Press, Cambridge, 1992, p. 43.

—HUGH THURSTON
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, CANADA V6T 1Z2

43. ______, *Récréations Mathématiques*, Gauthier-Villars, Paris, 1891–1894.
44. ______, *Théorie des Nombres*, Gauthier-Villars, Paris, 1891.
45. W. F. Lunnon, The Reve's Puzzle, *Comput. J.* 29 (1986), 478.
46. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, San Francisco, 1982.
47. S. B. Maurer and A. Ralston, *Discrete Algorithmic Mathematics*, Addison-Wesley Publishing Co., Reading, MA, 1990.
48. S. Minsker, The Towers of Hanoi rainbow problem: Coloring the rings, *J. Algorithms* 10 (1989), 1–19.
49. R. B. Nelson, E. T. H. Wang and B. Poonen, Solution to Problem 1222, this MAGAZINE 59 (1986), 241–242.
50. H. Noland, problem 1350, this MAGAZINE 63 (1990), 189.
51. J. S. Rohl and T. D. Gedeon, The Reve's Puzzle, *Comput. J.* 29 (1986), 187–188; Corrigendum, *Ibid.* 31 (1988), 190.
52. T. Roth, The Tower of Brahma revisited, *J. Recreational Math.* 7 (1974), 116–119.
53. F. Scarioni and H. G. Speranza, A probabilistic analysis of an error-correcting algorithm for the Towers of Hanoi puzzle, *Inform. Process. Lett.* 18 (1984), 99–103.
54. R. S. Scorer, P. M. Grundy and C. A. B. Smith, Some binary games, *Math. Gaz.* 28 (1944), 96–103.
55. I. Stewart, Les fractals de Pascal, *Pour la Science* 129 (1988), 100–104.
56. ______, Le lion, la lama et la laitue, *Pour la Science* 142 (1989), 102–107.
57. B. M. Stewart, Solution to Problem 3918, *Amer. Math. Monthly* 48 (1941), 217–219.
58. M. Sved, Problem 1222, this MAGAZINE 58 (1985), 237.
59. ______, Divisibility—with visibility, *Math. Intelligencer* 10 (1988), 56–64.
60. S. Wagon, *Mathematica in Action*, W. H. Freeman, New York, 1991.
61. T. R. Walsh, The Towers of Hanoi revisited: Moving the rings by counting the moves, *Inform. Process. Lett.* 15 (1982), 64–67.
62. ______, Iteration strikes back—at the cyclic Towers of Hanoi, *Inform. Process. Lett.* 16 (1983), 91–93.
63. S. Wolfram, Geometry of binomial coefficients, *Amer. Math. Monthly* 91 (1984), 566–571.
64. D. Wood, The Towers of Brahma and Hanoi revisited, *J. Recreational Math.* 14 (1981), 17–24.
65. ______, Adjudicating a Towers of Hanoi contest, *Internat. J. Comput. Math.* 14 (1983), 199–207.

Math Bite: A Simple Proof That Every Sequence Has a Monotone Subsequence

Known proofs that every sequence of real numbers has a monotone subsequence seem to be, as one recent text [1] put it, "tricky." Here is a simple proof.

Obviously a sequence with no greatest member has an increasing subsequence.

Let s denote the sequence whose nth member is s_n.

If there is a number k such that the "tail" $s_k, s_{k+1}, s_{k+2}, \dots$ has no greatest member then this tail has an increasing, and consequently monotone, subsequence, which will be a subsequence of s. If there no such k, let s_p be a greatest member of s, s_q a greatest member of the tail $s_{p+1} s_{p+2}, \dots$, let s_r be a greatest member of the tail $s_{q+1}, s_{q+2}, \dots$ and so on. Then $s_p \geq s_q \geq s_r \dots$ and we have a monotone subsequence.

REFERENCE

1. R. P. Burn, *Numbers and Functions: Steps into Analysis*, Cambridge University Press, Cambridge, 1992, p. 43.

—HUGH THURSTON
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, CANADA V6T 1Z2

Bernoulli Trials and Mahonian Statistics: A Tale of Two q's

JENNIFER TREADWAY
DON RAWLINGS
California Polytechnic State University
San Luis Obispo, CA 93407

1. Introduction

Any casual comparison of the topic of Bernoulli trials with that of Mahonian statistics would lead one to the superficial observation that the letter q is common to both subjects: As is well known, the q of the Bernoulli coin-tossing scheme denotes the probability that tails occurs. Somewhat less well known is the fact that Mahonian statistics give rise to a branch of combinatorics that in some circles is referred to as q-combinatorics.

Of course, the use of the same symbol in these two contexts is mere coincidence. However, as will be explained, there is a surprising natural connection between these two famous q's of mathematics.

Our discovery of this connection was sparked by a comment made by Professor Jim Delany. Following a colloquium talk on permutation statistics [8], Delany quizzically remarked that there were some striking similarities between the talk and an article by Moritz and Williams [6] that had just appeared in this MAGAZINE. In brief, Moritz and Williams considered the problem of determining the probability that a given permutation would be the result of a certain coin-tossing game. It is their game that forms the link between the Bernoulli coin-tossing scheme and the subject of Mahonian statistics.

Beyond being just a curiosity, there are a few practical aspects of our discovery. For starters, the relationship between Bernoulli trials and Mahonian statistics immediately provides the answers to some of the questions posed by Moritz and Williams. We will also make use of the combinatorics of a certain Mahonian statistic, known as the comajor index, to give alternate proofs to some of the results in [6].

We begin our "tale of two q's" with a short review of Moritz' and Williams' work. Before proceeding, the reader may find it helpful to peruse their article.

2. Synopsis of Moritz' and Williams' article

To set the stage, we briefly present some terminology and notation. As is usual, any linear rearrangement of the integers $1, 2, \ldots, n$ will be referred to as a permutation on the set $\{1, 2, \ldots, n\}$. For instance, $\sigma = 4, 3, 1, 2$ is a permutation of the integers $1, 2, 3, 4$. We will let $S(n)$ denote the set of permutations on $\{1, 2, \ldots, n\}$.

When there is no possibility of confusion, the commas appearing in a permutation will be deleted. Instead of writing $\sigma = 4, 3, 1, 2$, we will henceforth write $\sigma = 4312$. Furthermore, the symbol $\sigma(i)$ will be used to denote the i^{th} element from the left in a permutation $\sigma \in S(n)$. For $\sigma = 4312$, $\sigma(2) = 3$. Thus, a typical permutation $\sigma \in S(n)$ will be expressed in the form $\sigma = \sigma(1)\sigma(2)\ldots\sigma(n)$.

The principal problem considered by Moritz and Williams may now be paraphrased as follows:

Principal Problem. Players $1, 2, \ldots, n$ in turn toss a coin that comes up heads with probability p and tails with probability $q = (1 - p)$. A player, upon tossing heads, goes out of the game and passes the coin to the next player still in the game. The remaining players continue to toss until all have gone out. The problem is to determine, for any permutation $\sigma = \sigma(1)\sigma(2)\ldots\sigma(n) \in S(n)$, the probability that the players go out in the order specified by σ (that is, for $1 \leq i \leq n$, $\sigma(i)$ is the i^{th} player to go out).

To solve this problem, Moritz and Williams introduced the notion of the norm of a permutation. Denoted by $|\sigma|$, the *norm* of $\sigma \in S(n)$ is the number of tails in the shortest sequence of coin tosses for which the game ends in σ.

As an illustration, consider the permutation $\sigma = 4\,7\,3\,6\,8\,2\,1\,5 \in S(8)$. Given that players 1 through 8 toss in natural order, the "minimal flipping sequence" that produces σ must begin with players 1, 2, and 3 each tossing tails and player 4 tossing heads. Continuing through the first pass of the natural tossing order, players 5, 6, 7, and 8 must respectively toss tails, tails, heads, and tails. For convenience, the outcomes of the first and subsequent passes through the tossing order that together comprise the minimal flipping sequence of σ are tabulated in Table 1. By adding the number of tails in the final column of Table 1, the norm of $\sigma = 4\,7\,3\,6\,8\,2\,1\,5$ is seen to be $|\sigma| = 6 + 3 + 2 + 0 = 11$.

TABLE 1 MINIMAL FLIPPING SEQUENCE PRODUCING $\sigma = 4\,7\,3\,6\,8\,2\,1\,5$.

Pass Number	Player 1	2	3	4	5	6	7	8	Number of Tails
1	T	T	T	H	T	T	H	T	6
2	T	T	H	—	T	H	—	H	3
3	T	H	—	—	T	—	—	—	2
4	H	—	—	—	H	—	—	—	0

In order to state Moritz' and Williams' elegant solution to this coin-tossing game, it is convenient to introduce some standard, though not widely known, terminology and notation. The *q-analog* and *q-factorial* of a nonnegative integer n are respectively defined to be the polynomials

$$[n] \equiv 1 + q + q^2 + \ldots + q^{n-1} \quad \text{and} \quad [n]! \equiv [1][2]\ldots[n]$$

where, by convention, $[0] \equiv 0$ and $[0]! \equiv 1$. To gain some familiarity with these polynomials, the reader will find it helpful to verify the results listed in Table 2.

TABLE 2 THE q-ANALOG OF $n!$ FOR $1 \leq n \leq 4$.

n	$[n]!$
1	$[1]! = 1$
2	$[2]! = [1]![2] = (1)(1 + q) = 1 + q$
3	$[3]! = [2]![3] = (1 + q)(1 + q + q^2) = 1 + 2q + 2q^2 + q^3$
4	$[4]! = [3]![4] = (1 + 2q + 2q^2 + q^3)(1 + q + q^2 + q^3) = 1 + 3q + 5q^2 + 6q^3 + 5q^4 + 3q^5 + q^6$

It is also instructive to consider the value of $[n]!$ for some particular values of q. For instance, if $q = \frac{1}{2}$, then $[3]! = 1 + 2(1/2) + 2(1/4) + (1/8) = 21/8$. In the case $q = 1$, it is not difficult to see that $[n] = n$ and that $[n]! = n!$. This latter case provides the rational for referring to $[n]$ and $[n]!$ respectively as q-analogs of n and its factorial.

Using the notation of q-analogs, Moritz' and Williams' main result may then be stated in the following form:

THEOREM 1. *For $\sigma \in S(n)$, let $P(\sigma)$ denote the probability that players $1, 2, \ldots, n$ go out in the order specified by σ. Then $P(\sigma) = q^{|\sigma|}/[n]!$.*

An interesting point to note is that Theorem 1 provides a continuum of measures on $S(n)$ that varies from the unit measure concentrated on a single permutation to the equiprobable measure: If $q = 0$, then $P(\sigma)$ is equal to 1 if $\sigma = 12\ldots n$ and is equal to 0 otherwise. If $q = 1$, then $P(\sigma) = 1/n!$ for all $\sigma \in S(n)$.

Besides giving an explicit formula for $P(\sigma)$, Moritz and Williams also considered a number of peripheral questions that arise in connection with the partitioning of $S(n)$ "according to norm." Let $S(n,k)$ denote the set of permutations contained in $S(n)$ that have norm k, that is,

$$S(n,k) \equiv \{\sigma \in S(n) : |\sigma| = k\}.$$

This partitioning of $S(n)$ is very natural in the sense that the permutations in $S(n,k)$ occur with equal probability, namely $q^k/[n]!$. Thus, partitioning $S(n)$ according to norm is equivalent to partitioning according to probability.

As the computation of the norm is indirect (involving the minimal flipping sequence), Moritz and Williams posed the following natural questions:

Query 1. Is there a direct method for computing $|\sigma|$ that would thereby provide a simpler characterization of $S(n,k)$?

Query 2. Is there a closed formula for the cardinality of $S(n,k)$? Have these numbers arisen in other contexts?

Although Moritz and Williams did not answer these questions, they did derive a number of significant identities pertaining to the cardinality of $S(n,k)$. As these identities are relevant to the discourse to follow, we present them in Theorem 2.

THEOREM 2. *Suppose $\left\langle {n \atop k} \right\rangle$ denotes the cardinality of $S(n,k)$ and that $m(n)$ denotes the binomial coefficient $\binom{n}{2} = n(n-1)/2$. Then*

(i) *Symmetry Property*: $\left\langle {n \atop k} \right\rangle = \left\langle {n \atop m(n)-k} \right\rangle$

(ii) *Row Sum Property*: $\sum_{k=0}^{m(n)} \left\langle {n \atop k} \right\rangle = n!$

(iii) *Additive Recursive Property*: $\left\langle {n \atop k} \right\rangle = \sum_{j=k-n+1}^{k} \left\langle {n-1 \atop j} \right\rangle$

(iv) *Generating Function Property*: $\sum_{\sigma \in S(n)} q^{|\sigma|} = \sum_{k=0}^{m(n)} \left\langle {n \atop k} \right\rangle q^k = [n]!$.

At this point, it is convenient to briefly outline the rest of this paper. In section 3, we will give a short introduction to the subject of Mahonian statistics. We will then see that part (iv) of Theorem 2 says that the norm is a Mahonian statistic. This fact alone will partially answer Moritz' and Williams' Query 2. In section 4, we will show that the norm is not a new Mahonian statistic. In fact, the norm turns out to be equal to the statistic known as the comajor index. This observation answers Query 1 by providing a simpler way of computing the norm. In section 5, we will use the combinatorics of the comajor index to give alternate proofs of the results of Theorem 2. Finally, in section 6, we indicate how a slight change in Moritz' and Williams' game leads to a measure that is "induced" by the well-known inversion number of a permutation.

3. Mahonian statistics

So what exactly is a Mahonian statistic? Well, as coined in 1975 by Dominique Foata [3], the term refers to any statistic on $S(n)$ that has the q-factorial of n as its generating function. More explicitly, a statistic Φ: $S(n) \to \left\{0, 1, \ldots, \binom{n}{2}\right\}$ is said to be *Mahonian* if

$$\sum_{\sigma \in S(n)} q^{\Phi(\sigma)} = [n]!.$$

For instance, consider the function Φ: $S(3) \to \{0, 1, 2, 3\}$ defined by

σ	123	132	213	231	312	321
$\Phi(\sigma)$	2	0	1	2	3	1

Since $q^2 + q^0 + q^1 + q^2 + q^3 + q^1 = 1 + 2q + 2q^2 + q^3 = [3]!$, Φ is Mahonian on $S(3)$.

The connection of Moritz' and Williams' work with the subject of Mahonian statistics should now be self-evident: Part (iv) of Theorem 2 merely states that the norm is Mahonian! This observation has a number of interesting implications, which we now proceed to bring out within the context of a brief digression through the history of Mahonian statistics.

The inversion number, which qualifies as the oldest known Mahonian statistic, perhaps dates back to the origin of determinants. In preparation for defining it, an ordered pair (i, j) is said to be an *inversion* in a permutation $\sigma \in S(n)$ if $1 \leq i < j \leq n$ and $\sigma(i) > \sigma(j)$. We then define the *inversion number* of σ, denoted by *inv* σ, to be the number of inversions in σ, that is,

$$inv\, \sigma \equiv |\{(i, j) : 1 \leq i < j \leq n, \sigma(i) > \sigma(j)\}|.$$

As an example, note that the permutation $\sigma = 31254 \in S(5)$ has 3 inversions; namely (1,2), (1,3), and (4,5). Thus, $inv\, \sigma = 3$.

The connection of the inversion number with the determinant should be clear: One of the standard definitions of the determinant of a matrix $A = (a_{ij})_{n \times n}$ of real numbers is given by

$$\det A \equiv \sum_{\sigma \in S(n)} (-1)^{inv\, \sigma} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

However, as the theory of determinants is only incidental to our story here, no more will be said of this connection.

The first of a series of results of direct significance for our purposes came in 1839 when Rodriguez [**10**] showed that the inversion number is Mahonian:

THEOREM 3. *If $a(n, k)$ denotes the number of permutations in $S(n)$ with k inversions, then*

$$\sum_{\sigma \in S(n)} q^{inv\, \sigma} = \sum_{k=0}^{m(n)} a(n, k) q^k = [n]!$$

where, as before, $m(n) = \binom{n}{2}$.

The second development as chronicled in Comtet [1, 236–240], occurred in 1871 with Bourget's proof of the recurrence relationship given in Theorem 4.

THEOREM 4. *The numbers* $a(n,k)$ *satisfy, for* $n \geq 1$, *the recurrence relationship*

$$a(n,k) = \sum_{j=k-n+1}^{k} a(n-1,j)$$

with initial conditions $a(n,0) \equiv 1$ *and* $a(0,k) \equiv 0$ *for* $k \geq 0$.

In addition to Bourget's result, Comtet goes on to catalog and prove a host of facts concerning the numbers $a(n,k)$. For later reference, we record some of these facts below.

THEOREM 5. *The numbers* $a(n,k)$ *satisfy the following identities*:
(i) $a(n,k) = a(n, m(n) - k)$

(ii) $a(n,k) = a(n,k-1) + a(n-1,k)$ *for* $k < n$

(iii) $\sum_{k=0}^{m(n)} a(n,k) = n!$

(iv) $\sum_{k=0}^{m(n)} (-1)^k a(n,k) = 0$

(v) $\sum_{\sigma \in S(n)} inv\, \sigma = \sum_{k=0}^{m(n)} k a(n,k) = \frac{1}{2}\binom{n}{2} n!$.

We are now in a position to reap the first benefits from our observation that the norm is Mahonian. To facilitate this "harvest," we begin by stating and proving Theorem 6.

THEOREM 6. *Suppose that* Φ *is Mahonian and that* $b(n,k) \equiv |\{\sigma \in S(n) \colon \Phi(\sigma) = k\}|$. *Then* $b(n,k)$ *is equal to the number of permutations in* $S(n)$ *with* k *inversions, that is*, $b(n,k) = a(n,k)$.

The proof is trivial: Since the coefficient of q^k in the generating function $[n]!$ is unique, it follows from Theorem 3 that $b(n,k) = a(n,k)$.

Although partitioning $S(n)$ according to norm is certainly not equivalent to partitioning according to inversion number, the upshot of Theorem 6 and part (iv) of Theorem 2 is that $\left\langle {n \atop k} \right\rangle = a(n,k)$. As fallout, we first note that this explains the obvious overlap between the identities satisfied by the numbers $\left\langle {n \atop k} \right\rangle$ and those satisfied by the $a(n,k)$. Explicitly, we have the following correspondence of results:

(i) of Theorem 2 $\leftrightarrow$ (i) of Theorem 5
(ii) of Theorem 2 $\leftrightarrow$ (iii) of Theorem 5
(iii) of Theorem 2 $\leftrightarrow$ Theorem 4
(iv) of Theorem 2 $\leftrightarrow$ Theorem 3.

Parts (ii), (iv) and (v) of Theorem 5 provide new identities for $\left\langle {n \atop k} \right\rangle$.

Furthermore, the equality $\left\langle {n \atop k} \right\rangle = a(n,k)$ answers the second half of Moritz' and Williams' Query 2 in the affirmative: Yes, the numbers $\left\langle {n \atop k} \right\rangle$ have occurred in other contexts.

In concluding this section, we present a second classic Mahonian statistic, namely the *major index*, which is both historically significant and apropos to the next section. Also known as the greater index, the major index was conceived of and used in the study of ordered partitions by Major Percy A. MacMahon in the early twentieth century. Working in a more general setting than that of permutations, MacMahon **[4, 5]** was the first to establish that the major index has the same generating function as does the inversion number. It is this very result for which the class of Mahonian statistics have been named.

To define the major index, it is convenient to first introduce the notion of a descent. A permutation $\sigma \in S(n)$ is said to have a *descent* at index i, $1 \leq i \leq n-1$, if $\sigma(i) > \sigma(i+1)$. Furthermore, the *descent set* of $\sigma \in S(n)$ is defined to be

$$\operatorname{Des} \sigma \equiv \{i: 1 \leq i \leq n-1, \sigma(i) > \sigma(i+1)\}.$$

The *major index* of $\sigma \in S(n)$ is then defined to be

$$maj\, \sigma \equiv \sum_{i \in \operatorname{Des} \sigma} i$$

where, by convention, a sum over an empty set is defined to be 0. As an illustration, consider $\sigma = 4\,7\,3\,6\,8\,2\,1\,5 \in S(8)$. By inserting asterisks to highlight descents in σ, we have $\sigma = 4\,7 * 3\,6\,8 * 2 * 1\,5$. Thus, $\operatorname{Des} \sigma = \{2, 5, 6\}$ and $maj\, \sigma = 2 + 5 + 6 = 13$.

4. The comajor index: a simple interpretation of the norm

It turns out that the norm ($|\sigma|$) is not a new Mahonian statistic. In fact, by closely examining the examples in Moritz' and Williams' article, it is not too difficult for one knowledgeable of permutation statistics to guess that the norm is no more than a slightly disguised variation of the major index. This slight variation has come to be known as the comajor index.

As apparently first considered by Désarménien and Foata [2], the *comajor index* of a permutation $\sigma \in S(n)$ is defined to be

$$comaj\, \sigma \equiv \sum_{i \in \operatorname{Des} \sigma} (n-i).$$

Thus, whereas the major index sums the indices of descents "relative to the left-hand side of σ," the comajor index sums descent indices "relative to the right-hand side of σ."

Consider the permutation $\sigma = 4\,7\,3\,6\,8\,2\,1\,5 \in S(8)$. From Table 1, we have that $|\sigma| = 11$. By inspection, we see that $\operatorname{Des} \sigma = \{2, 5, 6\}$, which implies that $comaj\, \sigma = (8-2) + (8-5) + (8-6) = 11$. Thus, $|\sigma| = comaj\, \sigma$.

We now state our main result:

THEOREM 7. *For any $\sigma \in S(n)$, we have $|\sigma| = comaj\, \sigma$.*

Rather than a formal proof of Theorem 7, we give an enlightening example. We begin by reconsidering the minimal flipping sequence (MFS for short) of Table 1 and its associated permutation σ in the compact form

$$\begin{array}{lcccccccccccccccccccccc}
\text{MFS:} & T & T & T & H & T & T & H & T & | & T & T & H & T & H & H & | & T & H & T & | & H & H \\
\sigma: & & & & 4 & & & 7 & & * & & & 3 & & 6 & 8 & * & & 2 & & * & 1 & 5
\end{array}$$

where bars have been inserted in the MFS to indicate points in the game at which the coin is passed to the beginning of the tossing order and where, as before, asterisks have been inserted in σ to highlight descents. The important observations to make here are listed below:

(1) There is a natural one-to-one correspondence between the bars and the asterisks.
(2) The contribution to $comaj\, \sigma$ made by the first descent from the left in σ is equal to the number of tails that precede the first bar in the MFS.
(3) For $2 \leq i \leq des\, \sigma$, the contribution to $comaj\, \sigma$ made by the ith descent in σ is equal to the number of tails between the $(i-1)$st and ith bars in the MFS.

Together, these observations imply that $|\sigma| = comaj\ \sigma$.

Of course, Theorem 7 provides an answer to Query 1 of Moritz and Williams. The comajor index may be directly calculated from σ. Moreover, the characterization of $S(n,k)$ in terms of the comajor index, $S(n,k) = \{\sigma \in S(n)\colon comaj\ \sigma = k\}$, is easier to use.

5. Alternate proofs to the results of Theorem 2

Using the combinatorics associated with the comajor index, it is possible to prove all of the identities of Theorem 2. As the Row Sum Property (part (ii) of Theorem 2) is the most straightforward, we begin with it.

First, it is clear that the comajor index is minimized on $S(n)$ at $\sigma = 12\dots n$. Note that $comaj(12\dots n) = 0$. Also easy to see is that the comajor index achieves its maximum on $S(n)$ at $\sigma = n\dots.21$. Thus, the maximum is equal to $(n-1)+(n-2)+\cdots+1 = n(n-1)/2$. Recalling that $m(n) = n(n-1)/2$, it then follows that

$$\bigcup_{k=0}^{m(n)} \{\sigma \in S(n)\colon comaj\ \sigma = k\} = S(n).$$

As the sets in the above union are pairwise disjoint, part (ii) of Theorem 2 is immediate.

In order to establish the remaining properties listed in Theorem 2, we will make use of two bijections. The first bijection, which we will denote by C, is sometimes referred to as complementation on $S(n)$. For $\sigma \in S(n)$, it is the permutation defined by

$$C(\sigma) \equiv (n+1-\sigma(1))(n+1-\sigma(2))\dots(n+1-\sigma(n)).$$

For instance, $C(47368215) = 52631784$.

A proof of the Symmetry Property for the numbers $\left\langle {n \atop k} \right\rangle$ may now be given by simply observing the change in the comajor index that results when the complementation operation is applied to a permutation. Just note that relative to $\{1,2,\dots,n-1\}$, the sets Des $C(\sigma)$ and Des σ are complementary, that is,

$$\begin{aligned}\operatorname{Des} C(\sigma) &\equiv \{i\colon 1 \le i \le n-1,\ n+1-\sigma(i) > n+1-\sigma(i+1)\}\\ &= \{i\colon 1 \le i \le n-1,\ \sigma(i) < \sigma(i+1)\}\\ &= \{1,2,\dots,n-1\} \setminus \operatorname{Des} \sigma.\end{aligned}$$

We therefore have that

$$comaj\, C(\sigma) = \sum_{i \in \operatorname{Des} C(\sigma)} (n-i) = \sum_{i=1}^{n-1} (n-i) - \sum_{i \in \operatorname{Des}(\sigma)} (n-i) = m(n) - comaj\ \sigma.$$

Thus, when restricted to permutations of $S(n)$ having comajor index k, C is a bijection from $S(n,k)$ to $S(n, m(n)-k)$, thereby establishing the Symmetry Property of Theorem 2.

Our second bijection, which we will denote by Γ, is a variation of one presented in [7] and is based on observing the effect that "inserting n" into a permutation $\theta \in S(n-1)$ has on the comajor index. To observe this effect, the n possible "insertion positions" in $\theta = \theta(1)\theta(2)\dots\theta(n-1)$ are labeled as follows: Using the labels $0, 1, \ldots, n-1$ in order, first scan θ from left to right and label the positions that, upon insertion of n, will *not* result in the "creation" of a new descent. Then scanning back from right to left, label the remaining positions. For instance, for $\theta = 4\,7\,3\,6\,2\,1\,5$, the top and bottom rows of the display

$$\begin{array}{ccccccccccccccc}
 & \cdot & & \cdot & & 0 & & \cdot & & 1 & & 2 & & \cdot & & 3 \\
 & & & & & \downarrow & & & & \downarrow & & \downarrow & & \downarrow & & \\
\theta = & & 4 & & 7 & & 3 & & 6 & & 2 & & 1 & & 5 & \\
 & \uparrow & & \uparrow & & & & \uparrow & & & & & & \uparrow & & \\
 & 7 & & 6 & & \cdot & & 5 & & \cdot & & \cdot & & 4 & & \cdot
\end{array}$$

respectively indicate the labels as distributed by the two scans.

Now, in terms of the just described insertion procedure, for ordered pairs $(\theta; l) \in S(n-1) \times \{0, 1, \ldots, n-1\}$ we define a map Γ by

$$\Gamma(\theta; l) \equiv \sigma,$$

where σ is the permutation that results when n is inserted into position l of θ. As an example, from the preceding display, we see that $\Gamma(4\,7\,3\,6\,2\,1\,5; 4) = 4\,7\,3\,6\,2\,1\,8\,5$. Two important facts concerning Γ are summed up in Theorem 8.

THEOREM 8. *For $n \geq 1$, Γ is a bijection from $S(n-1) \times \{0, 1, \ldots, n-1\}$ to $S(n)$. Moreover, if $\Gamma(\theta; l) = \sigma$, then comaj $\sigma = l +$ comaj θ.*

It is the final property of Theorem 8 that explains the method behind our seemingly arbitrary labeling procedure. Leaving the proof of Theorem 8 to the diligent reader, we move on to some applications.

We first use Γ to establish the fact that the comajor index is Mahonian. The proof is elementary and proceeds by induction. For $n = 1$, we have that

$$\sum_{\sigma \in S(1)} q^{comaj\,\sigma} = q^{comaj\,1} = q^0 = 1 = [1]!.$$

Then, if we inductively assume that

$$\sum_{\sigma \in S(n-1)} q^{comaj\,\sigma} = [n-1]! \quad \text{for} \quad n \geq 2,$$

it follows from Theorem 8 that

$$\begin{aligned}
\sum_{\sigma \in S(n)} q^{comaj\,\sigma} &= \sum_{l=0}^{n-1} \sum_{\theta \in S(n-1)} q^{l + comaj\,\theta} \\
&= \sum_{l=0}^{n-1} q^l \sum_{\theta \in S(n-1)} q^{comaj\,\theta} = [n] \cdot [n-1]! = [n]!.
\end{aligned}$$

Thus, the comajor index is Mahonian. In view of Theorem 7, this establishes part (iv) of Theorem 2.

As a second application of Γ, we begin by noting that, if $\theta \in S(n-1, k-l)$, then

$$comaj\,\Gamma(\theta; l) = l + comaj\,\theta = l + (k-l) = k.$$

Then, by restriction, Γ may be viewed as a bijection from the union

$$\bigcup_{l=0}^{n-1} S(n-1, k-l) \times \{l\}$$

to the set $S(n, k)$. As the sets in this union are mutually disjoint, it follows that

$$\left\langle {n \atop k} \right\rangle = \sum_{l=0}^{n-1} \left\langle {n-1 \atop k-l} \right\rangle = \sum_{j=k-n+1}^{k} \left\langle {n-1 \atop j} \right\rangle.$$

Thus, we have established Bourget's Additive Recursive Property as given in Theorems 2(iii) and 4.

We conclude this section with two remarks. First, bijections analogous to Γ, that is, based on insertion, have long been used to study the classical Eulerian numbers (for instance, see [7]). Second, all of the other identities for the numbers $\left\langle {n \atop k} \right\rangle$ listed in Theorem 5 may be derived from their generating function $[n]!$. For the interested reader, these derivations are outlined by Comtet [1, 236–240].

6. An inversion game

In view of the connection between the comajor index (or norm) and Moritz' and Williams' game, it is only natural to ask whether or not there are games that lead to measures induced by the major index or the inversion number. In other words, are there games for which the probability that $\sigma \in S(n)$ occurs is given by $q^{inv\,\sigma}/[n]!$ or $q^{maj\,\sigma}/[n]!$?

The answer to this question is "yes." We will present a game that leads to the measure induced by the inversion number. It is stated below.

Inv Game. Players $1, 2, \ldots, n$ in turn toss a coin that comes up heads with probability p and tails with probability $q = (1 - p)$. A player, upon tossing heads, goes out of the game and passes the coin to the beginning of the tossing order. The remaining players continue to toss until all have gone out.

As is easily seen, the Inv Game is essentially the same as the game of Moritz and Williams. There is just one exception: Whenever heads is tossed, instead of the coin being passed to the next player remaining in the game, the coin is passed back to the beginning of the tossing order.

By appropriately adapting the proof of Theorem 1 as given by Moritz and Williams, it is not difficult to verify the following theorem.

THEOREM 9. *For $\sigma \in S(n)$, the probability that players $1, 2, \ldots, n$ go out of the Inv Game in the order specified by σ is given by $P(\sigma) = q^{inv\,\sigma}/[n]!$.*

The details are left as an exercise.

7. Epilogue

By exploiting the fact that the norm is Mahonian, we have been able to answer all but one of Moritz' and Williams' questions: We do not know whether or not there is a simple closed formula for $\left\langle {n \atop k} \right\rangle$. However, given that no such formula is listed in Comtet's book, which is a veritable encyclopedia of combinatorial results, it is doubtful that one is known.

Beyond this unsolved problem, an unexplored landscape is emerging. There is a host of intriguing questions that pertain to the interplay between probability and permutation statistics. A further discussion of this topic is given in [9].

REFERENCES

1. L. Comtet, *Advanced Combinatorics*, D. Reidel Publishing Co., Dordrecht, 1974.
2. J. Désarménien and D. Foata, Functions symmétriques et séries hypogéométriques basiques multivariées, *Bull. Soc. Math. France* 113 (1985), 3–22.
3. D. Foata, Distributions Euleriennes et Mahoniennes sur le groupe des permutations, *Higher Combinatorics* (M. Aigner, ed.), D. Reidel Publishing Co., Dordrecht, 1977.
4. P. A. MacMahon, The indices of permutations and the derivation therefrom of functions of a single variable associated with the permutations of any assemblage of objects, *Amer. J. Math.* 35 (1913), 281–322.
5. P. A. MacMahon, *Combinatory Analysis*, Cambridge University Press, London/New York, 1915 (reprinted by Chelsea Publishing Co., New York, 1960).
6. R. H. Moritz and R. C. Williams. A coin-tossing problem and some related combinatorics, this Magazine (1) 61 (1988), 24–29.
7. D. Rawlings, The r-major index, *J. Combin. Theory* Ser. A 25 (1981), 175–183.
8. D. Rawlings, The ABCs of classical enumeration, colloquium talk presented at Cal Poly State University, San Luis Obispo, CA, February 1988.
9. D. Rawlings, Bernoulli trials and permutation statistics, *Inter. J. of Math. and Math. Sciences* (2) 15 (1992), 291–312.
10. O. Rodriguez, Note sur les inversions ou dérangements produits dans les permutations, *J. de Mathématique* 4 (1839), 236–240.

Proof with Words: An Efficient Trisection of a Line Segment

Many texts give a construction of the trisection of a line segment based upon Euclid [Elements, VI, 9] that requires 9 elements (6 circles and 3 lines). The following "efficient" construction uses only 6 elements (2 circles and 4 lines).

Proof. Draw $\overline{CB}$. Then $\overline{AB}$ and $\overline{CE}$ are medians of $\triangle BCD$.

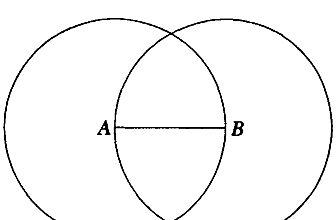


FIGURE 1

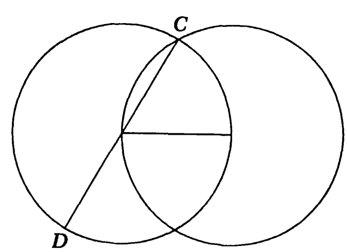


FIGURE 2

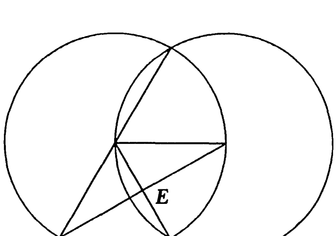


FIGURE 3

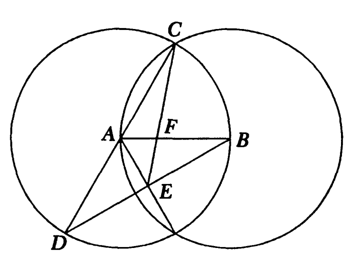


FIGURE 4

—Scott Coble
Gonzaga University
Spokane, WA 99258

REFERENCES

1. L. Comtet, *Advanced Combinatorics*, D. Reidel Publishing Co., Dordrecht, 1974.
2. J. Désarménien and D. Foata, Functions symmétriques et séries hypogéométriques basiques multivariées, *Bull. Soc. Math. France* 113 (1985), 3–22.
3. D. Foata, Distributions Euleriennes et Mahoniennes sur le groupe des permutations, *Higher Combinatorics* (M. Aigner, ed.), D. Reidel Publishing Co., Dordrecht, 1977.
4. P. A. MacMahon, The indices of permutations and the derivation therefrom of functions of a single variable associated with the permutations of any assemblage of objects, *Amer. J. Math.* 35 (1913), 281–322.
5. P. A. MacMahon, *Combinatory Analysis*, Cambridge University Press, London/New York, 1915 (reprinted by Chelsea Publishing Co., New York, 1960).
6. R. H. Moritz and R. C. Williams. A coin-tossing problem and some related combinatorics, this MAGAZINE (1) 61 (1988), 24–29.
7. D. Rawlings, The r-major index, *J. Combin. Theory* Ser. A 25 (1981), 175–183.
8. D. Rawlings, The ABCs of classical enumeration, colloquium talk presented at Cal Poly State University, San Luis Obispo, CA, February 1988.
9. D. Rawlings, Bernoulli trials and permutation statistics, *Inter. J. of Math. and Math. Sciences* (2) 15 (1992), 291–312.
10. O. Rodriguez, Note sur les inversions ou dérangements produits dans les permutations, *J. de Mathématique* 4 (1839), 236–240.

Proof with Words: An Efficient Trisection of a Line Segment

Many texts give a construction of the trisection of a line segment based upon Euclid [Elements, VI, 9] that requires 9 elements (6 circles and 3 lines). The following "efficient" construction uses only 6 elements (2 circles and 4 lines).

Proof. Draw $\overline{CB}$. Then $\overline{AB}$ and $\overline{CE}$ are medians of $\triangle BCD$.

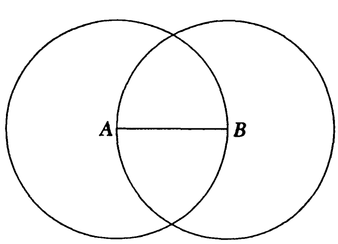


FIGURE 1

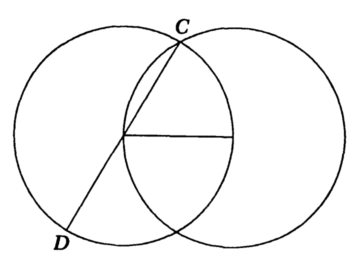


FIGURE 2

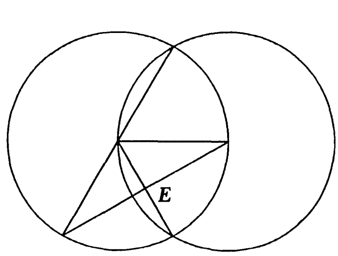


FIGURE 3

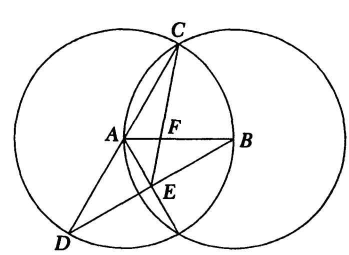


FIGURE 4

—SCOTT COBLE
GONZAGA UNIVERSITY
SPOKANE, WA 99258

NOTES

Minimum Integral Drawings of the Platonic Graphs

HEIKO HARBORTH
MEINHARD MÖLLER
Technische Universität Braunschweig
D-38106 Braunschweig, Germany

In memory of Torrence D. Parsons

The five platonic solids sustained continued interest for more than 2000 years in arts, science, and different fields of mathematics. Here we will consider the five platonic graphs that correspond to the vertices and edges of the five regular solids.

An *integral drawing* of a graph G is a mapping of the vertices of G into different points of the plane, and of the edges of G into straight line segments of integer length that connect corresponding endpoints in such a way that two line segments have at most one point in common, either a common endpoint or a crossing.

The existence of integral drawings for any graph G is guaranteed, for example, by the construction in [7] where all points are on a circle. Integral drawings of graphs might be of interest in science, for example, in the context of integral multiples of wave lengths or energy quanta.

A *planar graph* is a graph that can be embedded in the plane in such a way that no two edges intersect geometrically except at a vertex to which they are both incident. A graph embedded in the plane in this way is called a *plane graph*. Then a *plane integral drawing* is an integral drawing that is a plane graph. It is not known if it is always possible to produce a plane integral drawing of a planar graph.

CONJECTURE. *For every planar graph there exists a plane integral drawing.*

For special graphs G one can ask for the smallest integral drawing in the sense that the diameter $d(G)$ of an integral drawing is a minimum, where the *diameter of a drawing* is the largest integer in the set of lengths of all its straight line segments.

If we denote the minimum diameter of a plane integral drawing of a graph G by $d_p(G)$, these values for the five platonic graphs: tetrahedron (T), octahedron (O), cube (C), dodecahedron (D), and icosahedron (I), are determined in [2] to be

$$d_P(T) = 17, \qquad d_P(O) = 13, \qquad d_P(C) = 2, \qquad d_P(D) = 2, \qquad d_P(I) = 159.$$

See also [6] for two of these drawings.

In this note we will present the minimum diameters $d(G)$ for integral drawings of the platonic graphs G in the case where edge intersections are allowed.

RESULT. *The minimum diameters $d(G)$ of integral drawings of the platonic graphs G are*

$$d(T) = 4, \qquad d(O) = 7, \qquad d(C) = 1, \qquad d(D) = 1, \qquad d(I) = 8.$$

There exist exactly 1, 3, *and* 2 *different integral drawings with minimum diameter for the three integral drawings of* T, O, *and* I, *respectively.*

For C and D we obtain $d(C) = d(D) = 1$ immediately from FIGURES 1 and 2. For the remaining three triangulations T, O, and I, a computer was used to check successively whether for the first natural numbers n an integral drawing with diameter n can be constructed. For $d(T) = 4$ proofs without computer can be found in [3, 4, 5], and the unique smallest drawing of T is given in FIGURE 3.

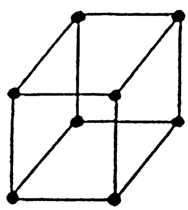


FIGURE 1
Integral drawing of the cube graph with all edges of length 1.

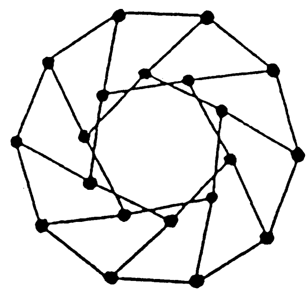


FIGURE 2
Integral drawing of the dodecahedron graph with all edges of length 1.

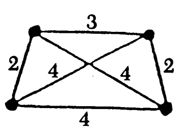


FIGURE 3
Unique integral drawing of the tetrahedron graph with minimum diameter 4.

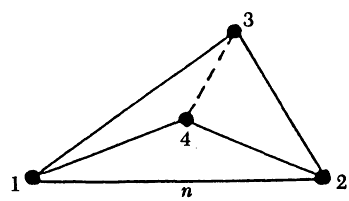


FIGURE 4
Order of constructed points for the tetrahedron graph.

The calculations can be done as follows. We choose 3, 4, or 5 integral triangles with side lengths $\leq n$ such that these can be triangles around a vertex v of T, O, or I with consecutive triangles having one side in common. A subsequent triangle may be added either clockwise or counterclockwise to the preceding one. In all possible cases the difference between the sums of the angles α_i at the vertex v of all triangles added clockwise and all triangles added counterclockwise has to be an integral multiple of 2π. This has to be checked by computation using the cosine rule for the integral triangles. If $\cos(\alpha_1 \pm \alpha_2 \pm \cdots) = 1$ is checked, all computations can be done by integer operations, that is, exactly. However, this method seems to use too much computing time, even for the octahedron. Therefore we used the following straightforward method that, although it requires consideration of computational errors, is possible on most personal computers.

Starting with points 1 and 2 at distance n we construct points $3, 4, \ldots,$ in the order shown in FIGURES 4, 5, and 6 as intersection points of circles with integral radii in all possible ways. The calculations of the coordinates of these points causes computational errors. Moreover, we determine the lengths of the dotted lines in FIGURES 4 to

6 from the coordinates of their endpoints. This causes further computational errors. For each dotted line determined in the figures, we check to see whether its length differs from an integer by less than a certain error E; if not, the next case is considered. If ε is the error for any simple computing operation of the computer used, for example, $\varepsilon = 10^{-17}$, then for this error E the estimate $E < 10^4 n^3 \varepsilon$ can be deduced. Since only $n \leq 11$ is used, we obtain $E < 10^8 \varepsilon$, that is, $E < 10^{-9}$ in the example. By this method we get sets of integral side lengths for all three graphs, and it remains in every case to check by the exact method described earlier whether these really give an integer solution.

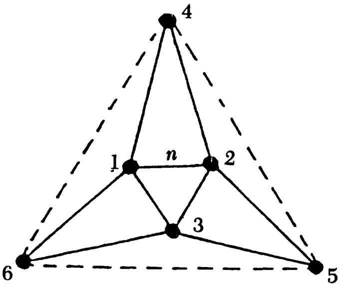


FIGURE 5
Order of constructed points for the octahedron graph.

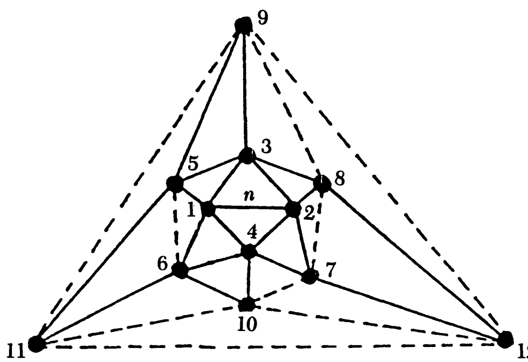


FIGURE 6
Order of constructed points for the icosahedron graph.

The minimum diameters $d(T)$, $d(O)$, and $d(I)$ obtained are as asserted above, and the computations yield only the 1, 3, and 2 minimum drawings shown in FIGURES 3, 7, and 8 for T, O, and I, respectively.

Finally, minimum integral drawings of T and O exist that use only three different edge lengths (see FIGURES 3 and 7). The computational search for integral drawings of I with diameters up to 11 yields only the two further examples in FIGURE 9. These also have only three different edge lengths. For T, drawings with just two lengths do exist. This follows from the results in [5], or by the fact that each of all six possible two-distance sets with four points that are determined in [1] contains at least one irrational distance. For O and I we do not know whether integral drawings can have just two different lengths.

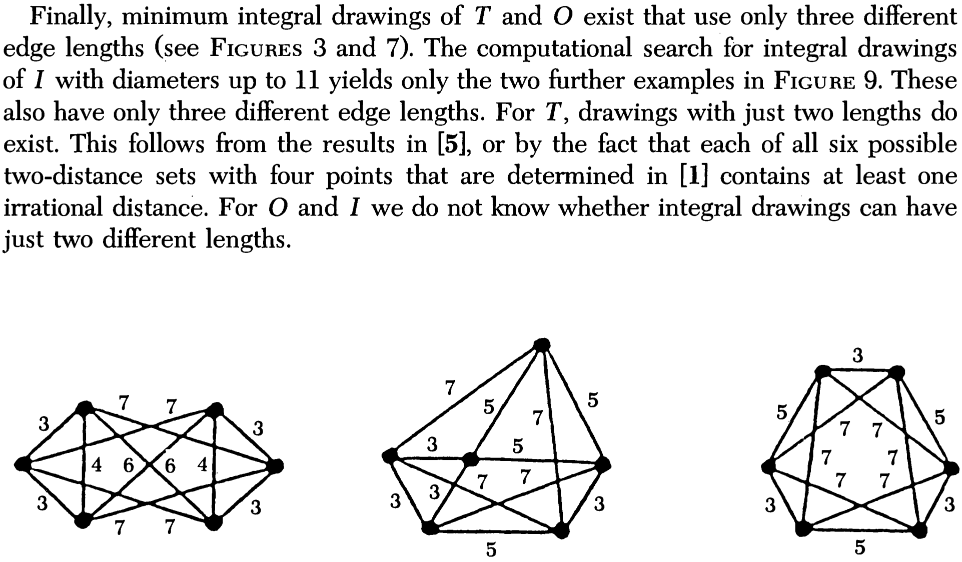


FIGURE 7
All three minimum integral drawings of the octahedron graph ($d(O) = 7$).

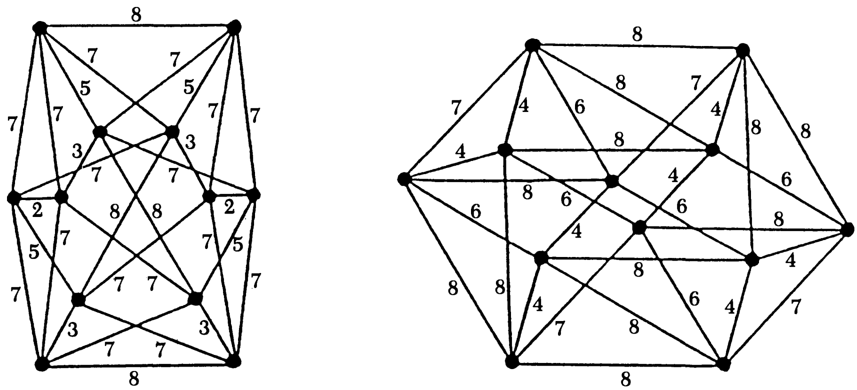


FIGURE 8
All two minimum integral drawings of the icosahedron graph ($d(I) = 8$).

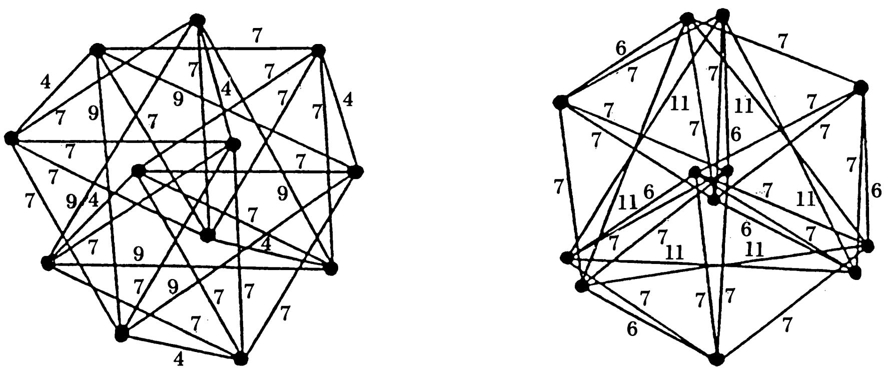


FIGURE 9
All two possible integral drawings of the icosahedron graph with diameters 9, 10, or 11.

REFERENCES

1. S. J. Einhorn and I. J. Schoenberg, On Euclidean sets having only two distances between points, *Indag. Math.* 28 (1966), 479–504.
2. H. Harborth, A. Kemnitz, M. Möller, and A. Süssenbach, Ganzzahlige planare Darstellungen der platonischen Körper, *Elem. Math.* 42 (1987), 118–122.
3. H. Harborth and A. Kemnitz, Diameters of integral point sets, *Colloquia Math. Soc. János Bolyai* 48, *Intuitive Geometry*, Siófok (Hungary) 1985, North-Holland, Amsterdam, 1987, 255–266.
4. H. Harborth, Aufgabe 950, *Elem. Math.* 42 (1987), 172–173.
5. H. Harborth and L. Piepmeyer, Point sets with small integral distances, In: *Applied Geometry and Discrete Mathematics, The Victor Klee Festschrift*, P. Gritzmann and B. Sturmfels (editors), DIMACS 4 (1991), 319–324.
6. N. Hartsfield and G. Ringel, *Pearls in Graph Theory*, Academic Press, San Diego, 1990, p. 173.
7. A. Müller, Auf einem Kreis liegende Punktmengen ganzzahliger Entfernungen, *Elem. Math.* 8 (1953), 37–38.

A Comparison of Two Elementary Approximation Methods

HARVEY DIAMOND
West Virginia University
Morgantown, WV 26506

LOUISE RAPHAEL
Howard University
Washington, DC 20059

1. Introduction There are two familiar methods one learns in calculus for approximating the value of a differentiable function near one or more data points where the function is known and/or easy to evaluate. The first method is linear interpolation and the second is the differential approximation. A natural question about the the two methods is: "Which is better?" This question was posed by Leon Henkin during a visit to a seminar on curve fitting given by one of the authors in the Summer Mathematics Institute for talented minority students at Berkeley.

In what follows, we study this question from several viewpoints, using only elementary methods. What we find interesting is that, starting with an intuitive but mathematically ill-formulated question, we are lead to a fairly rich investigation touching on several mathematical approaches, including a result on existence/uniqueness that applies to convex functions, an exact/local analysis based on the quadratic case, and a discussion of numerical methods.

2. Formulation and preliminary observations The function f is assumed to be differentiable. We begin by setting out the two alternatives:
Differential approximation: Given $f(a), f'(a)$ we approximate

$$f(x) \cong f(a) + f'(a)(x-a) \equiv f_D(x). \tag{1}$$

Linear interpolation: Given $f(a), f(b)$, we approximate

$$f(x) \cong f(a) + \frac{f(b)-f(a)}{b-a}(x-a) \equiv f_L(x). \tag{2}$$

FIGURE 1 illustrates $f(x)$ and its two approximations.

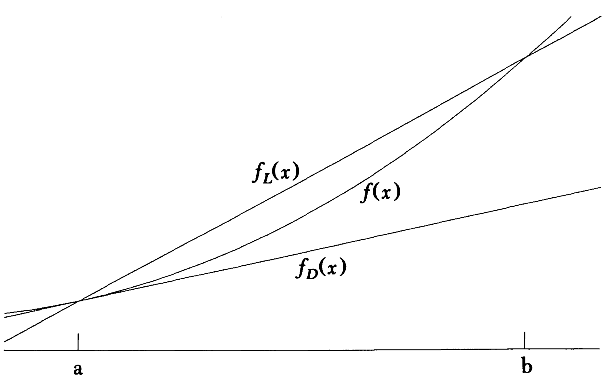


FIGURE 1

The *approximation error* of an approximation method at a particular value of x is simply the difference of $f(x)$ and its approximate value, in this case either $f(x)-f_D(x)$ or $f(x)-f_L(x)$. The *size* of the approximation error is its absolute value. We would be entitled to say that one method is better than the other if for all $a \leq x \leq b$ its approximation error had smaller size than that of the other method. We will see shortly, however, that this situation never occurs in our problem. The approximation method having the error of smaller size depends on the value of x, and one must try to say something about *where* (i.e. for which values of x) one method is better than another, and for which functions.

We begin with some simple observations. If x is near a, f_D is better, while if x is near b, f_L is better. In fact, the only way this could not be true is if the two approximations are identical. To explain this, note first that both approximations are the same at $x=a$. Near $x=b$, if $f_D(b) \neq f(b)$ then f_L is better near b, since $f_L(b)=f(b)$. On the other hand, if $f_D(b)=f(b)$ then both approximations agree at $x=a$ and $x=b$ and hence are identical since they are linear functions. Next, using the definition of the derivative, one can show that $f_D(x)$ is the best linear approximation of $f(x)$ near $x=a$. Consider all lines through the point $(a, f(a))$ that have the form $y=f(a)+m(x-a)$. The error in approximating $f(x)$ using such a line, at $x \neq a$, is then

$$f(x)-(f(a)+m(x-a))=(x-a)\left\{\frac{f(x)-f(a)}{(x-a)}-m\right\},$$

so that the term in brackets goes to zero as x goes to a (and hence gives the smallest size error) if, and only if, m is equal to $f'(a)$. Thus either f_D is better than f_L near a or else f_L also has derivative $f'(a)$, in which case it is identically equal to f_D.

Having shown that f_D is better near $x=a$ and f_L is better near $x=b$, the question of which approximation method is better becomes, "For what values of x is f_D better than f_L and for what values is the opposite true?" In turn, the answer to this question hinges on the answer to the question "At what points are the approximation errors of the two methods the same size?" We note that at any point where the approximation errors have the same size, the approximation errors must be of opposite sign or else (again) the approximations are identical. This occurs because if the approximation errors are the same at some point then $f_D=f_L$ at that point and, along with $f_D(a)=f_L(a)$, which is always true, we see that both approximations correspond to the same line. Thus if the functions f_D and f_L are not identical, the set of points at which the approximation errors have the same size is therefore precisely where the approximation errors sum to zero, which is precisely the solution set of the equation

$$f(x)-f_D(x)+f(x)-f_L(x)=0 \quad \text{or} \tag{3}$$

$$f(x)=[f_D(x)+f_L(x)]/2. \tag{4}$$

A geometric interpretation of (4) is shown in FIGURE 2, namely the solution set of (4) is the set of points where the line $y=[f_D(x)+f_L(x)]/2$ intersects the curve $y=f(x)$. The label c in FIGURE 2 is the location of x at the point of intersection in the interval (a, b).

3. Qualitative results Continuing with the graph in FIGURE 2, we observe that $f(x)$ is concave up and there is exactly one value of x, $x=c$ in the figure, where the sizes of the two approximation errors are the same. The following theorem shows that this is true for any function that is strictly concave up or concave down and has two continuous derivatives.

THEOREM 1. *Let $f''(x)$ be continuous and either strictly positive or strictly negative for all x in some open interval I. Then for any a, b in I, f_D and f_L cannot be identical, and, aside from the point $x = a$, there is exactly one other point in I, lying between a and b, at which the approximation errors have the same size.*

Proof. Motivated by equation (3), we define

$$g(x) = f(x) - f_D(x) + f(x) - f_L(x).$$

First, f_D and f_L cannot be identical for if they are, $g(a) = 0$, $g(b) = 0$, and $g'(a) = 0$. The first two facts (by Rolle's theorem) imply the existence of a point α between a and b for which $g'(\alpha) = 0$. Since $g'(a) = 0$ there exists another point β between a and α at which $g''(\beta) = 0$. But $g''(x) = 2f''(x)$ since f_D and f_L are linear, so $f''(\beta) = 0$. This contradicts the hypothesis that $f''(x)$ is either strictly positive or strictly negative in I.

Since f_D and f_L are not identical, the approximation errors have the same size if, and only if, (3) is satisfied, i.e., $g(x) = 0$. Now $g(a) = 0$ is always true. If g has two other distinct roots in I there are, by Rolle's theorem, at least two distinct points in I where $g'(x)$ is zero, and so at least one point in I where $g''(x) = 0$, again contradicting the hypothesis on $f''(x)$. Thus g has at most one other root in I, but since we showed that f_D and f_L are not identical, by previous arguments there is at least one point between a and b where $g(x) = 0$. Thus, aside from $x = a$, there is a unique root of g and it lies between a and b. This proves the theorem.

We might ask what can happen if the hypotheses of the theorem do not hold and f has one inflection point, say $x = c$. In this case f_D and f_L can be identical, but if not, there is, interestingly, still only one point where the approximation errors have the same size.

THEOREM 2. *Let $f''(x)$ be continuous and suppose that the closed interval $I = [\alpha, \beta]$ contains an interior point c with $f''(c) = 0$ and $f''(x) > 0$ for $x > c$, $f''(x) < 0$, for $x < c$, where $x \in I$ is assumed in each case. Then the following holds:*

i) There are infinitely many pairs a, b, with $a < c < b$, for which $f_D(x)$ and $f_L(x)$ are identical.

ii) If f_D and f_L are not identical for a particular choice of $a, b \in I$ then there is exactly one point between a and b at which the approximation errors have the same size.

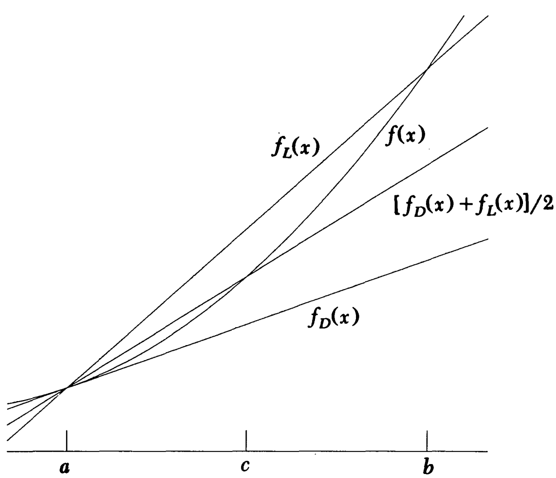


FIGURE 2

Proof. We sketch a proof of part i) with the aid of FIGURE 3. The idea is that, using values of a close to but less than c, the graph of $f_D(x)$ (the tangent from $(a, f(a))$), will intersect the graph of $f(x)$ at some $x = b$ with $b > c$ and $b \in I$. Then f_L will be identical to f_D. To effect this idea as a proof, we begin by explicitly exhibiting the dependence of f_D on the point a by using the notation $f_D(x; a)$. Now we can show, using Rolle's theorem, that

$$f_D(x;c) < f(x) \quad \text{for } x > c, \tag{5}$$

i.e., the tangent through the turning point is below the right, concave-up, portion of the curve. Similarly, we can show that

$$\text{if } a < c,\ f_D(x;a) > f(x) \quad \text{for } x \leq c, \tag{6}$$

i.e., a tangent constructed on the left, concave-down, portion of the curve is above the left portion of the curve. Now consider $f_D(\beta; a)$ as a function of a; it is clearly a continuous function of a. When $a = c$, we have $f_D(\beta; c) < f(\beta)$, by virtue of (5), so for $a < c$ and $|a - c|$ sufficiently small, $f_D(\beta; a) < f(\beta)$ holds by continuity. On the other hand, $f_D(c; a) > f(c)$ by virtue of (6). The last two inequalities imply, by the intermediate value theorem, the existence of a point b with $c < b < \beta$, such that $f_D(b; a) = f(b)$, completing part i).

The proof of part ii) is more difficult and is left to the reader, although elementary techniques are still adequate. FIGURE 4 provides a "generic" picture of why the theorem is true. The unique value of x between a and b where the approximation errors have the same size is denoted by c; we observe that at $x = c'$ the approxima-

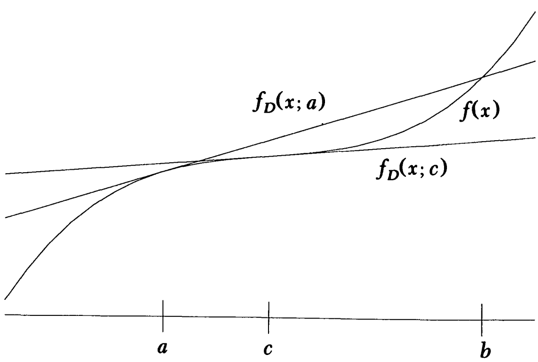


FIGURE 3

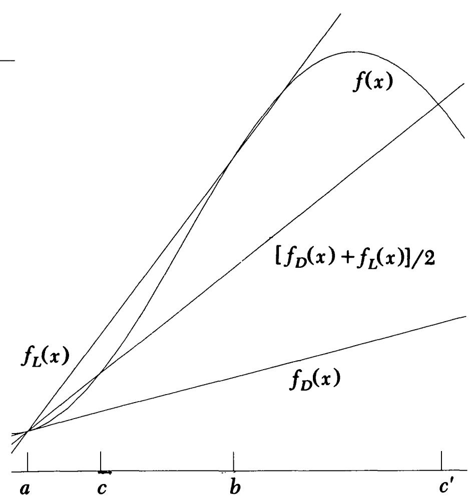


FIGURE 4

tion errors also have the same size, but this value of x lies outside the interval $[a, b]$.

4. An important special case In investigating our problem further, we will completely and explicitly solve the problem in the case when f is quadratic. "Solving the problem" means finding the solution of (3), which we know from Theorem 1 is unique, but in any case will be observed to be unique from the calculations below.

If f is quadratic, say $f(x) = px^2 + qx + r$, then $f(x)$ can be written as $f(x) = f(a) + f'(a)(x - a) + p(x - a)^2$.

This representation can most easily be obtained from Taylor's theorem after noting that $p = f''(a)/2$, since the coefficient of x^2 must be p. It then follows from (1) that

$$f(x) - f_D(x) = p(x - a)^2. \tag{7}$$

On the other hand, with f quadratic and f_L linear, $f(x) - f_L(x)$ is a quadratic function that is zero at $x = a$ and $x = b$ so we must have

$$f(x) - f_L(x) = p(x - a)(x - b). \tag{8}$$

The constant factor can be identified as p since subtracting a linear function from f does not change the coefficient of x^2.

At the x we seek, the solution of (3), the sum of (7) and (8) is zero,

$$p\left[(x - a)^2 + (x - a)(x - b)\right] = 0 \text{ or } p(x - a)(2x - a - b) = 0$$

with solutions $x = a$ and $x = (a + b)/2$. Thus for quadratic functions, differentials are better until halfway between a and b and then the linear interpolant is better. This is a pleasing result and provides a good rule of thumb, as we briefly explain next.

The solution of the quadratic case provides a local analysis of the general problem, applicable when a and b are close together and the hypotheses of Theorem 1 hold for some open interval I containing $[a, b]$. Under these conditions, f is well-approximated by a quadratic function and $x = (a + b)/2$ will be a good approximation to the unique solution of (3). A precise formulation and proof of this fact is somewhat difficult, however, and will not be attempted here.

5. Numerical considerations As discussed in section 3, if $f(x)$ is concave up or down, the problem of determining where each approximation method is better than the other reduces to finding the unique solution of (3) for values of x in the interval (a, b). If we define

$$g(x) \equiv f(x) - f_D(x) + f(x) - f_L(x), \tag{9}$$

then we are seeking the solution of $g(x) = 0$ in the interval (a, b). It turns out that Newton's method is guaranteed to converge to the desired solution if we use $x_0 = b$ as the initial guess for the root of g.

THEOREM 3. *If $f(x)$ satisfies the hypotheses of Theorem 1, then Newton's method applied to the function $g(x)$ in (9), namely, the iteration*

$$x_{n+1} = x_n - g(x_n)/g'(x_n),$$

with starting value $x_0 = b$, will converge to the unique solution $x = c$ of $g(x) = 0$ in the interval (a, b).

Proof. We will not provide a proof here. Results of this sort for Newton's method are well known; see for instance [1, p. 62]. The result basically follows from the observations depicted graphically in FIGURE 5:

i) The sequence $\{x_n\}$ is a decreasing sequence with $x_n > c$. This comes from the concavity of f (and therefore g) in the interval (a, b).

ii) As a decreasing sequence bounded from below, $\{x_n\}$ has a limit, say x^*, and this limit must satisfy $x^* = x^* - g(x^*)/g'(x^*)$ so that $g(x^*) = 0$. It follows that $x^* = c$.

Finally, we briefly consider the possibility of solving (4) by iteration. If we solve (4) for x on the (linear) right side, we can obtain an equation of the form $x = h(x)$. It is tempting to try to solve this equation numerically by the iteration $x_{n+1} = h(x_n)$, or what is the same thing, the iteration

$$[f_D(x_{n+1}) + f_L(x_{n+1})]/2 = f(x_n). \tag{10}$$

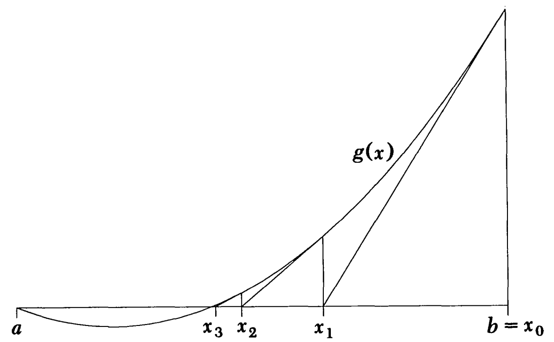


FIGURE 5

This turns out to be rather dangerous, even in the simplest case, where f satisfies the hypotheses of Theorem 1. In the situation depicted in FIGURE 2, for instance, it is easy to see graphically that for the iteration (10), $x = a$ is a stable fixed point, whereas the solution we seek, $x = c$, is an unstable fixed point. Two other possibilities, both with $f'(a) > 0$, $f''(x) < 0$, are depicted in FIGURE 6. If $f'(x) > 0$, $a < x < b$ (see FIGURE 6a), then the solution $x = c$ is a stable fixed point and any x_0 in the interval (a, b) results in a monotone sequence converging to $x = c$. Perhaps most interesting is the case depicted in FIGURE 6b, in which f has a maximum to the left of $x = c$. In this case $x = a$ is an unstable fixed point, but $x = c$ may or may not be stable. In fact this generic situation leads, under certain conditions, to the well-known chaotic maps of an interval onto itself, of which the most famous example is

$$x_{n+1} = 4\lambda x_n(1 - x_n), \tag{11}$$

where chaos occurs for certain values of the parameter λ lying between 0 and 1. (See [2] for a recent discussion at an elementary level.) Indeed, this example arises if we take $f(x) = x - x^2$, $a = 0$ and b a parameter, so that $f_D(x) = x$, $f_L(x) = (1 - b)x$ and (10) becomes (11) with $4\lambda = (1 - b/2)^{-1}$.

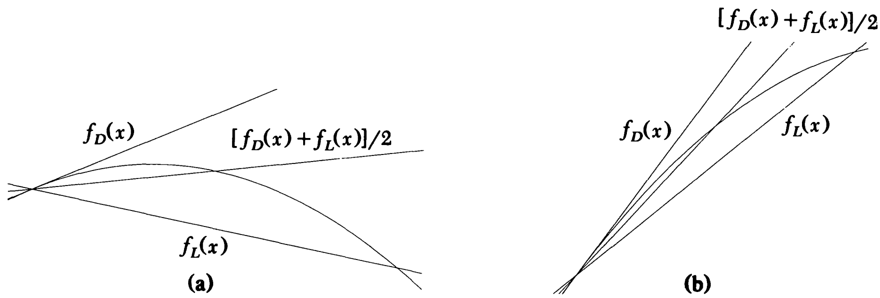


FIGURE 6

6. Concluding remarks We have shown in the preceding discussion, how a simple question concerning the comparison of linear interpolation and the differential approximation leads to some interesting mathematics, using only elementary methods at the calculus level. We discovered classes of functions for which the answer to the question could be well-characterized qualitatively; and obtained a nice exact answer in the case of quadratic functions, which also serves as a local approximation. Finally we showed that Newton's method can be reliably used to find the "cross-over" point in the case of concave up/down functions, while the method of iteration, naively applied, can converge, diverge, or lead to chaotic behavior even in the simplest examples.

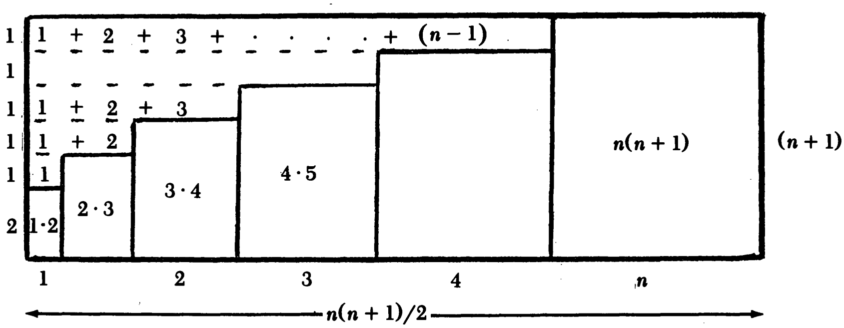
The authors were partially supported by an Air Force Office of Scientific Research contract 91MN062.

REFERENCES

1. Kendall E. Atkinson, *An Introduction to Numerical Analysis*, 2nd edition, John Wiley & Sons, New York, 1989.
2. Robert L. Devaney, The orbit diagram and the Mandelbrot set, *The College Math. J.* 22 (1991), 23–38.

Proof without Words
Sum of Products of Consecutive Integers

$$1\cdot 2+2\cdot 3+3\cdot 4+\cdots+n(n+1)=n(n+1)(n+2)/3.$$



$$1\cdot 2+2\cdot 3+3\cdot 4+\cdots+n(n+1)$$

$$=[n(n+1)/2][n+1]-\left[\begin{array}{l} \quad 1 \\ +1+2 \\ +1+2+3 \\ + \\ \vdots \\ +1+2+3+\cdots+(n-1) \end{array}\right.$$

$$=\frac{n(n+1)^2}{2}-\frac{(n-1)n(n+1)}{6}=\frac{n(n+1)(n+2)}{3}.$$

—James O. Chilaka
Long Island University
Greenvale, NY 11548

6. Concluding remarks We have shown in the preceding discussion, how a simple question concerning the comparison of linear interpolation and the differential approximation leads to some interesting mathematics, using only elementary methods at the calculus level. We discovered classes of functions for which the answer to the question could be well-characterized qualitatively; and obtained a nice exact answer in the case of quadratic functions, which also serves as a local approximation. Finally we showed that Newton's method can be reliably used to find the "cross-over" point in the case of concave up/down functions, while the method of iteration, naively applied, can converge, diverge, or lead to chaotic behavior even in the simplest examples.

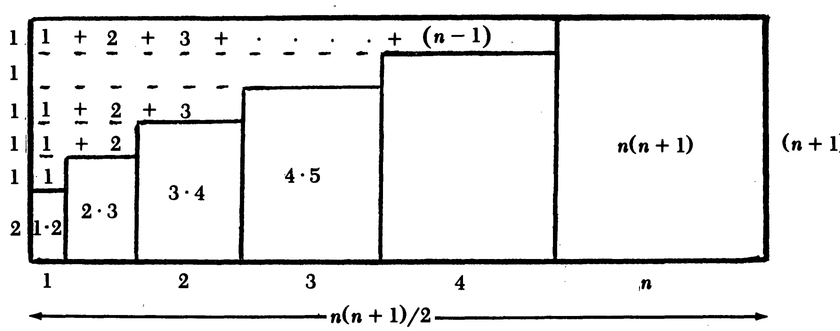
The authors were partially supported by an Air Force Office of Scientific Research contract 91MN062.

REFERENCES

1. Kendall E. Atkinson, *An Introduction to Numerical Analysis*, 2nd edition, John Wiley & Sons, New York, 1989.
2. Robert L. Devaney, The orbit diagram and the Mandelbrot set, *The College Math. J.* 22 (1991), 23–38.

Proof without Words
Sum of Products of Consecutive Integers

$$1\cdot 2+2\cdot 3+3\cdot 4+\cdots+n(n+1)=n(n+1)(n+2)/3.$$



$$1\cdot 2+2\cdot 3+3\cdot 4+\cdots+n(n+1)$$

$$=[n(n+1)/2][n+1]-\left[\begin{array}{l}1\\+1+2\\+1+2+3\\+\\ \vdots\\+1+2+3+\cdots+(n-1)\end{array}\right.$$

$$=\frac{n(n+1)^2}{2}-\frac{(n-1)n(n+1)}{6}=\frac{n(n+1)(n+2)}{3}.$$

—JAMES O. CHILAKA
LONG ISLAND UNIVERSITY
GREENVALE, NY 11548

Counting Centralizers in Finite Groups

SARAH MARIE BELCASTRO*
Haverford College
Haverford, PA 19041

GARY J. SHERMAN*
Rose-Hulman Institute of Technology
Terre Haute, IN 47803

A finite (and all groups mentioned in this paper are finite) abelian group is the direct product of cyclic groups of prime power order. Thus, one expects group structure to become increasingly complex with decreasing 'abelianness.' Indeed, the basic classification scheme for groups reflects the importance of the notion of commutativity in understanding group structure because labeling a group as abelian, nilpotent, supersolvable, solvable, or simple indicates, at least in a qualitative sense, the degree of commutativity the group enjoys.

Beginning abstract algebra students tend to ignore the subtleties of the commutativity issue—$xy = yx$ as far as they are concerned. Thirteen or fourteen years of commutative arithmetic are hard to dismiss. An effective way to deal with this misconception is to address it directly by asking 'How many pairs of elements of a group commute?' or 'What is the probability that two group elements commute?' The formal answers are $\#\mathrm{Com}(G) = |\{(x, y)|xy = yx\}|$ and $\mathrm{PrCom}(G) = \#\mathrm{Com}(G)/|G|^2$, respectively. These questions and their (formal) answers put the notion of commutativity on a numerical basis, which students enjoy, and provide motivation for a delightful excursion through some nice elementary group theory. Indeed, the fact that $\#\mathrm{Com}(G) = k|G|$, where k is the number of conjugacy classes in G, is woven from elementary results on subgroups, centralizers, Lagrange's theorem, and conjugacy classes [3]. The equivalent probabilistic statement, $\mathrm{PrCom}(G) = k/|G|$, leads, unsurprisingly, to a reassuring result,

$$G \text{ is abelian if, and only if, } \mathrm{PrCom}(G) = 1, \tag{1}$$

and, pleasantly, to a surprising result,

$$G \text{ is nonabelian if, and only if, } \mathrm{PrCom}(G) \leq 5/8. \tag{2}$$

Another, less precise, way to say this is that either all of the elements commute or at most $5/8$ of the elements commute. For details of the development of the '$5/8$-bound' for commutativity see [5] or pages 329 and 330 of [4].

Here's another question relating numbers and commutativity: How many distinct centralizers can a group have? Recall that the centralizer of x in G, denoted by $C(x)$, is the subgroup of G consisting of all elements that commute with x; i.e., $C(x) = \{y \in G|xy = yx\}$. If we denote the number of distinct centralizers in G by $\#\mathrm{Cent}(G)$, then $\#\mathrm{Cent}(G) = |\{C(x)|x \in G\}|$ and our question becomes 'What can we say about $\#\mathrm{Cent}(G)$?' This paper is an itinerary for an excursion in elementary group theory motivated by this question. Our goal is to provide some answers and some more questions. We think these are interesting in their own right and useful to those who teach abstract algebra.

*Work supported by National Science Foundation grant DMS-8922674.

Since G is abelian if, and only if, $C(x) = G$ for each element of G, we may begin by stating the analogue of (1).

$$G \text{ is abelian if, and only if, } \#\mathrm{Cent}(G) = 1.$$

Is there a centralizer 'gap' for nonabelian groups like the probability 'gap' between 1 and 5/8?

THEOREM 1. *If G is not abelian, then $\#Cent(G) \geq 4$.*

Proof. G is certainly the union of its centralizers. But the center of G,

$$Z = \{z \mid zx = xz \text{ for each } x \in G\} = \bigcap_{x \in G} C(x),$$

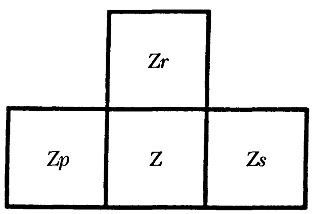
is a subset of each centralizer, so G is the union of its proper centralizers. If $\#\mathrm{Cent}(G) = 2$, then G is one of its proper subgroups, which is impossible. If $\#\mathrm{Cent}(G) = 3$, then G is the union of two proper subgroups, say H and K. This is also impossible because if we choose $x \in H - K$ and $y \in K - H$, we have no place to put xy. For example, $xy \in H$ implies $y \in H$ because $y = x^{-1}xy$.

Can $\#\mathrm{Cent}(G) = 4$? Yes, the dihedral group on four symbols, D_4, and the quaternion group of order eight, Q, are groups with exactly four centralizers. For example, the distinct centralizers of D_4, written as unions of right cosets of the center, are

$$\begin{aligned} D_4 &= Z \cup Z \cdot (1,2,3,4) \cup Z \cdot (1,2)(3,4) \cup Z \cdot (1,3), \\ P &= Z \cup Z \cdot (1,2,3,4), \\ R &= Z \cup Z \cdot (1,2)(3,4), \\ S &= Z \cup Z \cdot (1,3), \end{aligned}$$

where $Z = \langle (1,3)(2,4) \rangle$.

The essence of this example can be captured graphically, and mathematically:



THEOREM 2. *$\#Cent(G) = 4$, if, and only if, $G/Z \cong Z_2 \oplus Z_2$; i.e., G modulo its center is isomorphic to the Klein four group.*

Proof. If $G/Z \cong Z_2 \oplus Z_2$, then there are noncentral elements, p, r, and s, of G such that $G = Z \cup Zp \cup Zr \cup Zs$. It follows that the three proper subgroups of G containing Z are $P = Z \cup Zp$, $R = Z \cup Zr$, and $S = Z \cup Zs$. Let x be one of p, r, or s and let X be the corresponding subgroup. Notice that for $zx \in Zx$, $G \supset C(zx) \supseteq X$. So,

$$[G:X] = [G:C(zx)][C(zx):X] = 2 \text{ and } [G:C(zx)] \neq 1,$$

thus $C(zx) = X$. Therefore the proper centralizers of G are precisely P, R, and S; i.e., $\#\mathrm{Cent}(G) = 4$.

For the converse, it is sufficient to show that $[G:Z] = 4$ because then

$$\text{either } G/Z \cong Z_2 \oplus Z_2 \quad \text{or} \quad G/Z \cong Z_4.$$

As G is nonabelian, G/Z cannot be cyclic—$G = Z \cup Zx \cup Zx^2 \cup Zx^3$ implies that G is abelian—which means the latter case is impossible. So, suppose $\#\mathrm{Cent}(G) = 4$ and let $P = C(p)$, $R = C(r)$, and $S = C(s)$ be the three proper centralizers of G. Since G can not be written as the union of two proper subgroups and since an element must belong to its centralizer, we may choose p, r, and s in $G - (R \cup S)$, $G - (P \cup S)$, and $G - (P \cup R)$, respectively. Moreover, at least one of the proper centralizers, say P, has index two in G. For otherwise,

$$|G| \leq |P| + |R| + |S| - 2|Z| \leq |G|/3 + |G|/3 + |G|/3 - 2 < |G|.$$

Further,

$$P \cap R = P \cap R \cap S = Z$$

because if $x \in (P \cap R) - Z$, then

i) $C(x) \neq G$ because $x \notin Z$,

ii) $C(x) \neq P$ and $C(x) \neq R$ because $p, r \in C(x)$,

iii) $C(x) \neq S$ because $x \notin S$,

which means that $\#\mathrm{Cent}(G)$ must be at least 5.

Now we can compute $|Z|$ using the fact that for subgroups X and Y of G,

$$|X \cap Y| = \frac{|X||Y|}{|XY|} \geq \frac{|X||Y|}{|G|}. \tag{3}$$

Indeed,

$$|Z| = |P \cap R| \geq \frac{|P||R|}{|G|} = \frac{|R|}{2}$$

since $|P| = |G|/2$. But $Z \neq R$, so $|Z| = |R|/2$. Similarly, $|Z| = |S|/2$. Thus

$$|G| = |P| + |R| + |S| - 2|Z| = |G|/2 + 2|Z| + 2|Z| - 2|Z| = |G|/2 + 2|Z|,$$

which implies that $|G|/2 = 2|Z|$; i.e., $[G:Z] = 4$, as desired.

Groups for which $G/Z \cong Z_2 \oplus Z_2$ are as abelian as a nonabelian group can be in the probabilistic sense also. To see this, recall that the order of the conjugacy class of an element is the index of the centralizer of that element. Thus, each conjugacy class of G is of order one or two. Therefore the number of conjugacy classes in G is

$$k = |Z| + (|G| - |Z|)/2 = |G|/4 + 3|G|/8 = 5|G|/8$$

and $\mathrm{PrCom}(G) = 5/8$. This computation also suggests why $\mathrm{PrCom}(G) \leq 5/8$ for nonabelian groups: $\mathrm{PrCom}(G)$ is as large as possible when the center is as large as possible and all of the noncentral elements are in conjugacy classes of size two. This occurs when $[G:Z] = 4$; i.e., when $G/Z \cong Z_2 \oplus Z_2$. The threshold of 'abelianness', as measured by $\#\mathrm{Cent}(G)$ or $\mathrm{PrCom}(G)$, is occupied by the same class of groups:

THEOREM 3. $\mathrm{PrCom}(G) = 5/8$ *and* $\#\mathrm{Cent}(G) = 4$ *are equivalent to* $G/Z \cong Z_2 \oplus Z_2$.

Some readers may recognize the last two statements in our proof of Theorem 1 and D_4 or Q as a solution to an old Putnam problem **[6]**.

Show that a finite group cannot be the union of two of its proper subgroups. Does the statement remain true if "two" is replaced by "three"?

Not long after this problem appeared, Bruckheimer, Bryan, and Muir [1] provided an elementary proof that

> A group can be written as the union of three proper subgroups, H, K, and L, if, and only if, $N = H \cap K \cap L$ is a normal subgroup of G and $G/N \cong Z_2 \oplus Z_2$. (4)

Theorem 2 is seen to be a special case of (4). Indeed, it is an easy exercise to modify our proof of Theorem 2 and thereby prove (4). However, we chose to give a direct proof of Theorem 2 to serve as a warm-up for the more complicated proof of

THEOREM 4. *$\#\mathrm{Cent}(G) = 5$ if, and only if, $G/Z \cong Z_3 \oplus Z_3$ or $G/Z \cong S_3$ where S_3 is the symmetric group on three symbols.*

Proof. If $G/Z \cong Z_3 \oplus Z_3$ or $G/Z \cong S_3$, we may argue as we did in the proof of Theorem 2. Specifically, if $G/Z \cong S_3$, then $G = Z \cup Zx \cup Zx^2 \cup Zy \cup Zyx \cup Zyx^2$ where x^3, y^2, $(xy)^2 \in Z$. There are exactly four proper subgroups of G properly containing Z,

$$Z \cup Zx \cup Zx^2,\ Z \cup Zy,\ Z \cup Zyx,\ Z \cup Zyx^2,$$

each of which is the centralizer of its noncentral elements because it is abelian and of prime index in G.

For the converse, suppose $\#\mathrm{Cent}(G) = 5$ and let P, R, S, and T be the four proper centralizers of G. It is convenient to weave the proof from a sequence of lemmas.

LEMMA 1. *No one of P, R, S, or T is contained in the union of the other three.*

Proof. Suppose to the contrary, and without loss of generality, that T is a subset of $P \cup R \cup S$. Then, since $G = P \cup R \cup S \cup T$, we must have $G = P \cup R \cup S$. It follows from (4) that $G/(P \cap R \cap S) \cong Z_2 \oplus Z_2$. So, if we can show $P \cap R \cap S = Z$, then $\#\mathrm{Cent}(G) = 4$ – a contradiction.

Choose $p \in P - (R \cup S)$, $r \in R - (P \cup S)$, and $s \in S - (P \cup R)$ such that $C(p) = P$, $C(r) = R$ and $C(s) = S$. This is possible because if, for example, no such p existed, we would have $C(p) = T$ for each $p \in P - (R \cup S)$; i.e., $P - (R \cup S) \subseteq T - (R \cup S)$ and we could interchange the roles of P and T. Now, let $x \in (P \cap R \cap S) - Z$ and consider $C(x)$:

$$\begin{aligned} C(x) &\neq G \text{ because } x \notin Z,\\ C(x) &\neq T \text{ because } x \notin T,\\ C(x) &\neq P \text{ because } r, s \in C(x) - P,\\ C(x) &\neq R \text{ because } p, s \in C(x) - R,\\ C(x) &\neq S \text{ because } p, r \in C(x) - S. \end{aligned}$$

Thus $(P \cap R \cap S) - Z$ is empty, and the contradiction, $\#\mathrm{Cent}(G) = 4$, follows.

In view of Lemma 1, a proper centralizer, X, contains an element, x, not belonging to the union of the other three proper centralizers. Thus, $C(x) = X$.

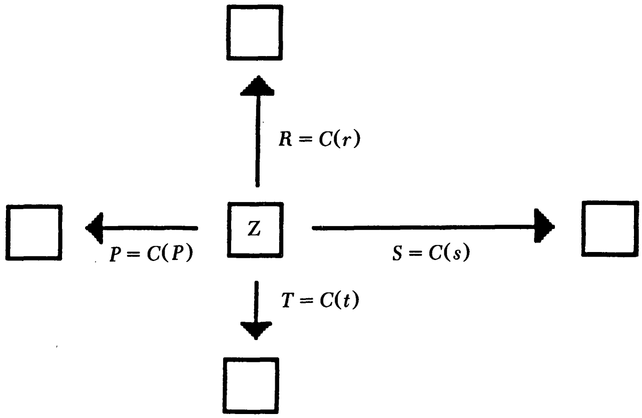
LEMMA 2. *No element of G is in exactly two proper centralizers.*

Proof. Suppose, for example, that $g \in (P \cap R) - (S \cap T)$. The only candidates for $C(g)$ are $P = C(p)$ and $R = C(r)$. However, the fact that both p and r are elements of $C(g)$ means that $C(g)$ can be neither $C(p)$ nor $C(r)$; i.e., $(P \cap R) - (S \cup T)$ is empty.

A similar proof establishes

LEMMA 3. *No element of G is in exactly three proper centralizers.*

The upshot of these lemmas is that G is 'almost' a disjoint union of its four proper centralizers, each of which is a union of cosets of the center.



The arithmetical summary is

LEMMA 4. $|G| = |P| + |R| + |S| + |T| - 3|Z|$.

And so, just as in the proof of Theorem 2, it is essential to constrain $|Z|$.

LEMMA 5. *If X and Y are distinct proper centralizers of G, then*

$$|X|\,|Y|/|G| \le |Z| \le |G|/6.$$

Proof. The lower bound for $|Z|$ follows from (3) because Lemma 2 and Lemma 3 imply $Z = X \cap Y$. The upper bound follows because $[G:Z] > 4$ and $[G:Z] = 5$ imply G is abelian.

To sharpen the lower bound, let's concentrate on the orders of the centralizers. We may assume, without loss of generality, that $|P| \ge |R| \ge |S| \ge |T|$. Thus, Lemma 4 and the fact that P is a subgroup of G imply $|P| = |G|/2$ or $|P| = |G|/3$. Now we can proceed by cases.

Case. $|P| = |G|/2$. In this case Lemma 4 guarantees that $|R| > |G|/6$. On the other hand, Lemma 5 applied to P and R implies $|G|/6 \ge |G|\,|R|/2|G|$; i.e., $|R| \le |G|/3$. Thus $|R|$ is one of $|G|/3$, $|G|/4$ or $|G|/5$. Reapplying Lemma 5 to P and R yields $|G|/10 \le |Z| \le |G|/6$. Thus, $|Z|$ is one of $|G|/6$, $|G|/8$, or $|G|/10$, because Z is a subgroup of P.

i) Let $|Z| = |G|/6$. The only choices for G/Z are S_3 and Z_6, so G/Z is S_3. Recall that if G/Z is cyclic, then G is abelian.

ii) Let $|Z| = |G|/8$. Since Z is a subgroup of R, $|R| = |G|/4$. It follows from Lemma 4 that $|S| + |T| = 5|G|/8$, which is impossible because $|S|, |T| \le |G|/4$.

iii) Let $|Z| = |G|/10$. Since Z is a subgroup of R, $|R| = |G|/5$. It follows from Lemma 4 that $|S| + |T| = 7|G|/10$, which is impossible because $|S|, |T| \le |G|/5$.

Case. $|P| = |G|/3$. Applying Lemmas 4 and 5, as in the previous case, yields

$$|G|/4 \le |R| \le |G|/3 \quad \text{and} \quad |G|/12 \le |Z| \le |G|/6.$$

And again, the fact that Z is a subgroup of P enables us to restrict $|Z|$ to one of $|G|/6$, $|G|/9$, or $|G|/12$.

i) Let $|Z| = |G|/6$. The only choices for G/Z are S_3 and Z_6 so G/Z is S_3.

ii) Let $|Z| = |G|/9$. The only choices for G/Z are Z_9 and $Z_3 \oplus Z_3$ so G/Z is $Z_3 \oplus Z_3$.

iii) Let $|Z| = |G|/12$. If $|R| = |G|/3$, then Lemma 5 applied to P and R yields the contradiction $|Z| \geq |G|/9$. If $|R| = |G|/4$, Lemma 4 implies that $|S| + |T| = 2|G|/3$, which is impossible because $|S|, |T| \leq 1/4$.

A connection with PrCom(G), while looser, still exists when Cent(G) = 5. For example, if $G/Z \cong S_3$ there are $|Z|$ conjugacy classes in Z, $|Z|$ conjugacy classes in $Zx \cup Zx^2$ and $|Z|$ conjugacy classes in $Zy \cup Zyx \cup Zyx^2$ so PrCom(G) = 1/2. Similarly, if $G/Z \cong Z_3 \oplus Z_3$, PrCom(G) = 11/27. Conversely, there is an elementary, but subtle, proof that PrCom(G) = 1/2 implies $G/Z \cong S_3$ [8]. And, if PrCom(G) = 11/27 and the smallest prime divisor of $|G|$ is three, then $G/Z \cong Z_3 \oplus Z_3$.

The proof of the previous statement is lurking in the paragraph preceding Theorem 3. Well, what's really lurking there is a proof of a more general result: If the smallest prime divisor of $|G|$ is p, then PrCom(G) $= (p^2 + p - 1)/p^3$ if, and only if, $G/Z \cong Z_p \oplus Z_p$. How many centralizers does such a group have? For $p = 2$ and for $p = 3$ we have shown that the answer is #Cent(G) = 4 and #Cent(G) = 5, respectively. These results are special cases of

THEOREM 5. *Let p be a prime. If $G/Z \cong Z_p \oplus Z_p$, then* #Cent(G) $= p + 2$. *If p is odd and $G/Z \cong D_p$, then* #Cent(G) $= p + 2$.

Proof. A quick check of our proofs of Theorems 2 and 4 will convince you that the crucial observations are

i) each proper subgroup of G properly containing Z is abelian and of prime index in G (i.e., it is a centralizer) and

ii) both $Z_p \oplus Z_p$ and D_p (p odd) have exactly $p + 1$ proper subgroups.

So, if for each prime there are groups satisfying the hypothesis of Theorem 5, we can make #Cent(G) as large as we like. For $p = 2$, D_4 and Q will do. If p is an odd prime, we can use D_p, since it has a trivial center, or either of the two groups of order p^3,

$$\langle a, b | a^{p^2} = e,\ b^p = e \quad \text{and} \quad b^{-1}ab = a^{1+p} \rangle,$$
$$\langle a, b, c | a^p = b^p = c^p = e,\ ab = bac,\ ca = ac \quad \text{and} \quad cb = bc \rangle,$$

since their centers must be of order p.

Question. Can we make #Cent(G) anything we like; i.e., if n is a positive integer other than two or three, does there exist a group with n = #Cent(G) centralizers? Which values of #Cent(G), other than 4 and 5, characterize G?

It is interesting (or embarrassing) to note that one of our choices for making #Cent(G) arbitrarily large is the class of 'most abelian' nonabelian groups ($G/Z \cong Z_p \oplus Z_p$) and the other is a class of 'fairly' abelian groups (D_p). We say that D_p is 'fairly' abelian because PrCom(D_p) is bounded below by 1/4 (PrCom(D_p) $= k/2p = 1/4 + 3/4p$). Here's an attempt at saving face: Normalize #Cent(G) by defining

$$\text{PrCent}(G) = \#\text{Cent}(G)/|G|.$$

Then, for the 'most abelian' nonabelian groups

$$\text{PrCent}(G) = (p+2)/p^2|Z| = 1/p|Z| + 2/p^2|Z|,$$

which decreases monotonically and reassuringly to 0 (the limiting value of PrCent(G) for abelian G) as p approaches infinity. Could it be that PrCent(G) approaches zero as $|G|$ approaches infinity for arbitrary G? No, not even for our class of 'fairly' abelian groups because PrCent(D_p) is bounded below by $1/2$. But, we can get a '$5/8$-like' bound for PrCent(G).

THEOREM 6. *Let p be the largest prime divisor of* $|G|$.

$$\mathrm{PrCent}(G) \le \begin{cases} 1/2, & \text{if } p = 2. \\ 3/4 + 1/4p, & \text{if } p \text{ is odd}. \end{cases}$$

Proof. Our strategy is to find a tractable partition of G that is compatible with the partition of G that is created when elements are assigned to their centralizers. If $p = 2$, then G is a 2-group and has a nontrivial center. Since elements in the same coset of the center have the same centralizer, the cosets of Z will do:

$$\mathrm{PrCent}(G) \le [G:Z]/|G| = 1/|Z| \le 1/2.$$

If p is odd we aren't guaranteed the existence of a nontrivial center. However, Cauchy's theorem guarantees the existence of an element of order p. Such an element is distinct from its inverse, but, as is the case for all group elements, has the same centralizer as its inverse. The partition we are looking for has sets of the form $\{x, x^{-1}\}$ as its components. Unfortunately, some of these components may consist of only one element. This occurs if, and only if, $x^2 = e$; i.e., $x = e$ or x is an involution. Let's denote the number of such elements in G by #Inv(G) and summarize:

$$\#\mathrm{Inv}(G) < |G|,$$
$$\#\mathrm{Cent}(G) \le \#\mathrm{Inv}(G) + (|G| - \#\mathrm{Inv}(G))/2 = (\#\mathrm{Inv}(G) + |G|)/2.$$

It follows that

$$\mathrm{PrCent}(G) \le 1/2 + \#\mathrm{Inv}(G)/2|G|. \tag{5}$$

And, it is known that if not all elements of G are involutions, then

$$\mathrm{PrInv}(G) = \#\mathrm{Inv}(G)/|G| \le (p+1)/2p. \tag{6}$$

Our result follows immediately from (5) in conjunction with (6).

The result quoted in (6) first appeared in an article published by G. A. Miller in 1905 [7]. We choose not to repeat any of the details here because Miller's paper is perfect reading for a beginning abstract algebra student—elementary and (old enough to be) refreshingly wordy.

While the bounds established in Theorem 6 are sharp for $p = 2$ (PrCent(D_4) = PrCent(Q) = $1/2$) and for $p = 3$ (PrCent(S_3) = $5/6$) that is not the case for $p > 3$. Indeed, if PrCent(G) = $3/4 + 1/4p$, then PrInv(G) must also be a maximum; i.e., PrInv(G) = $(p+1)/2p$. This is precisely the value of PrInv(D_p). A coincidence? No, Miller's paper also establishes that if PrInv(G) = $(p+1)/2p$, then

$$G \cong D_p \oplus Z_2 \oplus Z_2 \oplus \cdots \oplus Z_2. \tag{7}$$

Fortunately, #Cent(G) is a multiplicative function; i.e., the number of distinct centralizers in a direct sum is the product of the numbers of centralizers in each factor. Thus, for G as in (7), #Cent(G) = #Cent(D_p) = $p + 2$ and so PrCent(G) = $(p+2)/2pp^j$ where j is the number of times Z_2 occurs as a direct factor in G. It is easy to check that $(p+2)/2pp^j$ is as large as $3/4 + 1/4p$ if, and only if, $p = 3$ and $j = 0$. Theorem 7 and the succeeding question are natural corollaries to Theorem 6 and this discussion.

THEOREM 7. PrCent(G) $\leq 5/6$ *and* PrCent(G) $= 5/6$ *if, and only if,* $G \cong S_3$.

Question. Is it true that PrCent(G) $< 1/2$ unless G is Q or D_p?

We close by proposing two class discussion topics that are suggested by viewing the centralizer of an element as the stabilizer of that element under the action of G on itself by conjugation:

$$C(x) = \{y \in G | xy = yx\} = \{y \in G | y^{-1}xy = x\}.$$

i) Let G act on its set of subgroups by conjugation. The stabilizer of a subgroup H is referred to as the normalizer of H:

$$N(H) = \{y \in G | y^{-1}Hy = H\}.$$

How many normalizers can a group have? It's a classic result (see pages 138–140 of **[10]**) that G has exactly one normalizer (i.e., each subgroup of G is normal) if, and only if, G is abelian or Dedekind—the direct sum of Q, an abelian group in which each element is of order two and an abelian group in which each element is of odd order. Can a group have exactly two normalizers? Yes, here's an example,

$$Z_3 \oplus \langle a, b | a^9 = e,\ b^3 = e, \text{ and } b^{-1}ab = a^4 \rangle \oplus Q, \tag{8}$$

constructed from groups we have mentioned in this paper. It was shown just recently **[9]** that all groups with exactly two normalizers 'resemble' this example; i.e., for some prime p, G is (loosely speaking) isomorphic to a direct sum of an abelian group of order p^k, a so-called metacyclic group of order p^j, and a Dedekind group whose order is not divisible by p.

A slight abuse of terminology enables us to refer to the normalizer of an element:

$$\text{Norm}(x) = N(\langle x \rangle) = \{y \in G | y^{-1}\langle x \rangle y = \langle x \rangle\} = \{y \in G | y^{-1}xy = x^j \text{ for some } j\}.$$

And then, as with centralizers of elements, we set

$$\#\text{Norm}(G) = |\{N(x) | x \in G\}| \text{ and } \text{PrNorm}(G) = \#\text{Norm}(G)/|G|.$$

What can one say about #Norm(G) and PrNorm(G)? Well, there is not a 'gap' for #Norm(G) because #Norm(G) $= 2$ for the group defined in (8). On the other hand, it isn't clear that #Norm(G) $= 2$ implies that G has exactly two normalizers. Even if it does, the characterization that appears in **[9]** is long and hard, particularly when contrasted with our characterization of those groups with #Cent(G) $= 4$, so moving to #Norm(G) $= 3$ could be tough. But $C(x)$ is a subgroup of $N(x)$ and $N(x) = N(x^{-1})$, so all isn't lost.

ii) Let G be a finite abelian group and let A be its automorphism group **[11]**. The stabilizer of—well, just have your students take it from here.

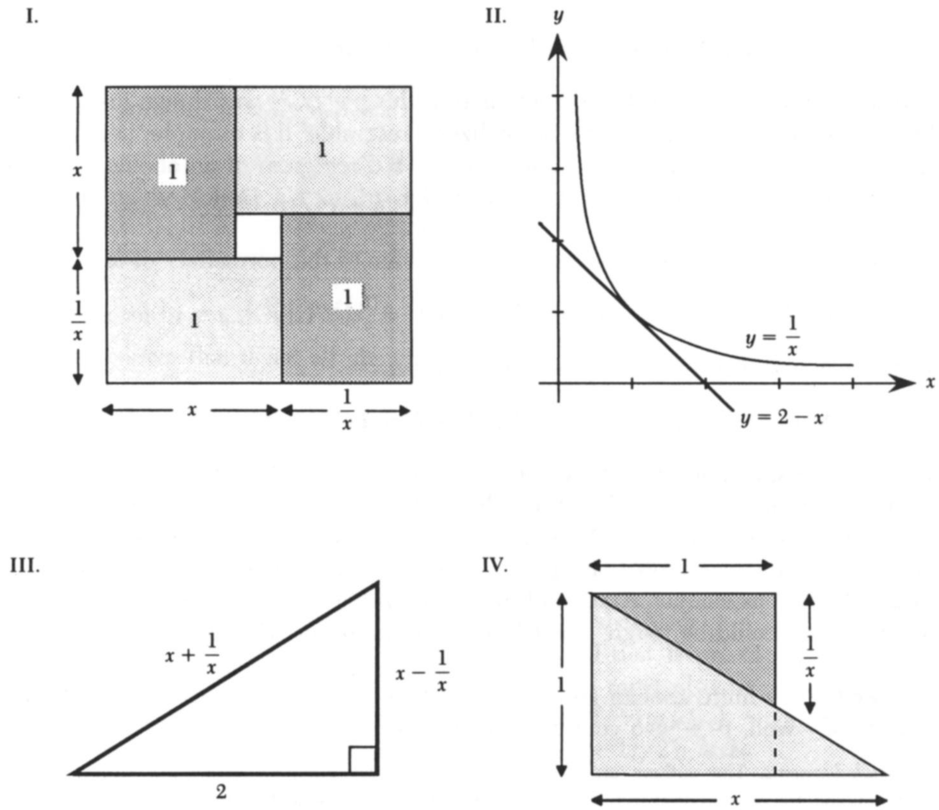
Acknowledgements. The authors are grateful to the referees for their assistance. Group-theoretic experimentation and computation were facilitated with the computer algebra system CAYLEY [2].

REFERENCES

1. M. Bruckheimer, A. C. Bryan, and A. Muir, Groups which are the union of three subgroups, *Amer. Math. Monthly* 77 (1970), 52–57.
2. D. F. Holt, The CAYLEY group theory system, *Notices Amer. Math. Soc.* 35 (1988) No. 8, 1135–1140.
3. P. Erdös and P. Turan, On some problems of a statistical group-theory, IV, *Acta Math. Acad, Sci. Hung.* 19 (1968), 413–415.

4. J. A. Gallian, *Contemporary Abstract Algebra*, 2nd edition, D. C. Heath and Company, Lexington, MA, 1990.
5. W. H. Gustafson, What is the probability that two group elements commute?, *Amer. Math. Monthly* 80 (1973), 1031–1034.
6. *The Thirtieth William Lowell Putnam Mathematical Competition*, MAA, December 6, 1969, Problem B-2.
7. G. A. Miller, Groups containing the largest possible number of operators of order two, *Amer. Math. Monthly* 12 (1905), 149–151.
8. J. Leavitt, G. J. Sherman, and M. J. Walker, Rewriteability in finite groups, *Amer. Math. Monthly* 99 (1992), 446–452.
9. M. D. Perez-Ramos, Groups with two normalizers, *Arch. Math.* 50 (1988), 199–203.
10. D. J. S. Robinson, *A course in the Theory of Groups*, Springer-Verlag New York, 1980.
11. G. J. Sherman, What is the probability an automorphism fixes a group element?, *Amer. Math. Monthly* 82 (1975), 261–264.

Proof without Words: The Sum of a Positive Number and Its Reciprocal Is at Least Two (four proofs)

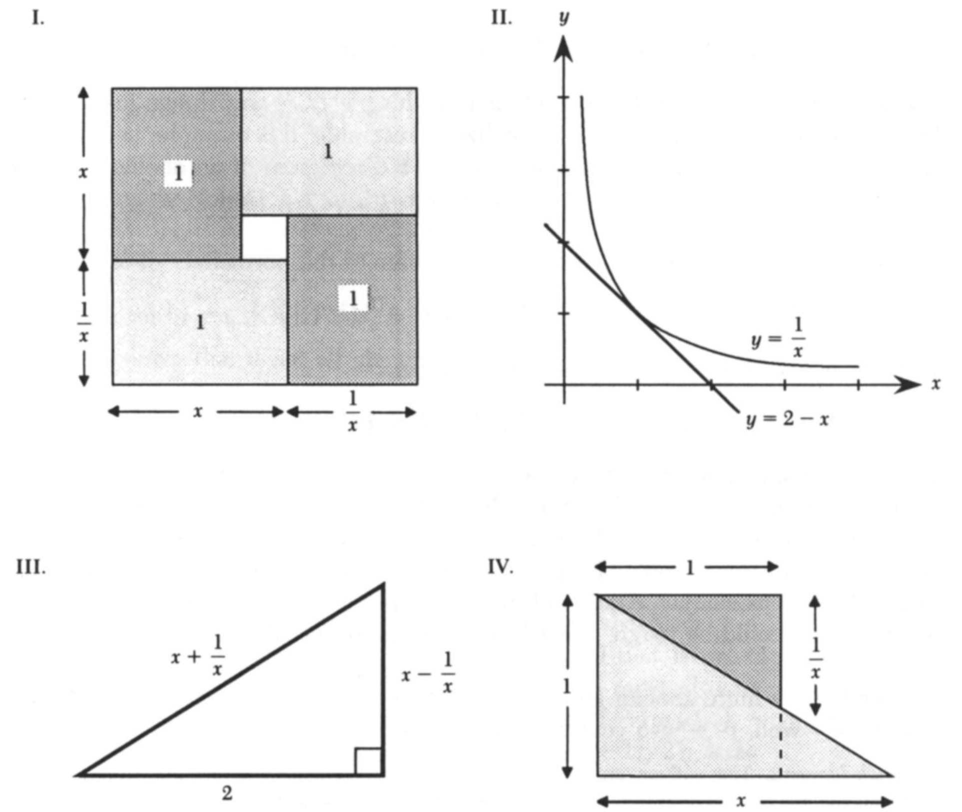


$$x \geq 1 \Rightarrow x + \frac{1}{x} \geq 2.$$

—ROGER B. NELSEN
LEWIS AND CLARK COLLEGE
PORTLAND, OR 97219

4. J. A. Gallian, *Contemporary Abstract Algebra*, 2nd edition, D. C. Heath and Company, Lexington, MA, 1990.
5. W. H. Gustafson, What is the probability that two group elements commute?, *Amer. Math. Monthly* 80 (1973), 1031–1034.
6. *The Thirtieth William Lowell Putnam Mathematical Competition*, MAA, December 6, 1969, Problem B-2.
7. G. A. Miller, Groups containing the largest possible number of operators of order two, *Amer. Math. Monthly* 12 (1905), 149–151.
8. J. Leavitt, G. J. Sherman, and M. J. Walker, Rewriteability in finite groups, *Amer. Math. Monthly* 99 (1992), 446–452.
9. M. D. Perez-Ramos, Groups with two normalizers, *Arch. Math.* 50 (1988), 199–203.
10. D. J. S. Robinson, *A course in the Theory of Groups*, Springer-Verlag New York, 1980.
11. G. J. Sherman, What is the probability an automorphism fixes a group element?, *Amer. Math. Monthly* 82 (1975), 261–264.

Proof without Words: The Sum of a Positive Number and Its Reciprocal Is at Least Two (four proofs)



$$x \geq 1 \Rightarrow x + \frac{1}{x} \geq 2.$$

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

On Length and Curvature

HAO JIANG
JEFFREY NUNEMACHER
Ohio Wesleyan University
Delaware, OH 43015

It is visually obvious that for any positive integer n the graph of the function $f(x) = x^{n+1}$ is longer than the graph of $g(x) = x^n$ over the interval $[0, 1]$. This observation is posed as a problem in the first edition of Marsden and Weinstein's calculus book ([1], p. 468). A nice proof can be given using the basic differential geometry of plane curves, which many students encounter in calculus. In this note, we discuss this proof as an application of curvature to a problem where it may not be thought to be the natural tool.

A first attempt to prove the required inequality is likely to be by a comparison of the two arc-length integrals. But this approach is not fruitful, since the arc-length integrand $\sqrt{1 + (n+1)^2 x^2}$ for f does not always exceed the integrand $\sqrt{1 + n^2 x^{2n-2}}$ for g. Indeed, the graph of f rises more slowly than that of g near 0 but then catches up before reaching 1. A more subtle comparison is necessary, one that relates the two curves along the normal to one of them.

We wish to make precise the following heuristic argument. The functions f and g have the same values at 0 and 1; f always lies below g on $(0, 1)$; so the convexity of g should imply that f is longer. But how does one measure convexity in a useful way? It turns out that the concept of curvature is exactly right for this purpose.

Let $X(s)$ denote the arc-length parametrization of the curve $y = x^n$ for $0 \le s \le s_0$ with $X(0) = (0, 0)$ and $X(s_0) = (1, 1)$. Recall that this represents the curve as a vector-valued function with derivative $X'(s)$ equal to a unit tangent vector $T(s)$ for each s. Let $N(s)$ denote the unit normal vector oriented at a right angle counterclockwise from $T(s)$. Differentiation of the equation $|X'(s)| = 1$ shows that $X''(s)$ is orthogonal to $T(s)$; thus it must be a multiple of $N(s)$. The proportionality factor is the signed curvature $k(s)$, defined as $\phi'(s)$, where $\phi(s)$ is the angle that $T(s)$ makes with the horizontal (see FIGURE 1). In symbols $X''(s) = k(s)N(s)$. A related equation is $N'(s) = -k(s)T(s)$, which follows in the same way from differentiating $|N(s)| = 1$. The convexity of the curve is then expressed by the inequality $k(s) > 0$. For a nice development of the elementary differential geometry of plane curves using these conventions see, for example, [2, p. 531–535].

We now consider a parametrization of $y = x^{n+1}$ derived from $X(s)$ as follows. Let $Y(s)$ denote the point on the graph of f obtained by intersecting this graph with the

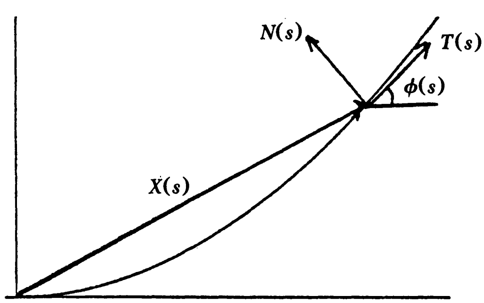


FIGURE 1

normal line to $y = x^n$ at $X(s)$ (see FIGURE 2). Then $Y(s) = X(s) - a(s)N(s)$, where $a(s)$ is a positive function because the graph of f lies below that of g for $0 < s < s_0$. Notice that despite the variable name, $Y(s)$ is not an arc-length parametrization of $y = x^{n+1}$, since the length of the tangent vector $Y'(s)$ will generally not be one. Everything else in the equation depends smoothly on s, so the function $a(s)$ must be continuously differentiable. Thus $Y'(s) = X'(s) - a'(s)N(s) - a(s)N'(s)$. But $N'(s) = -k(s)T(s)$ as noted above, so $Y'(s) = (1 + a(s)k(s))T(s) - a'(s)N(s)$; whence

$$|Y'(s)| = \sqrt{(1 + a(s)k(s))^2 + a'(s)^2}.$$

Thus for all s in $(0, s_0)$ we have that $1 = |X'(s)| < |Y'(s)|$, since both $a(s)$ and $k(s)$ are positive. Recall that the usual arc-length formula holds whether or not the curve is parametrized by arc length. Thus integration gives $\int_0^{s_0} |X'(s)|\, ds < \int_0^{s_0} |Y(s)|\, ds$, which establishes the desired inequality for the lengths of the graphs of f and g.

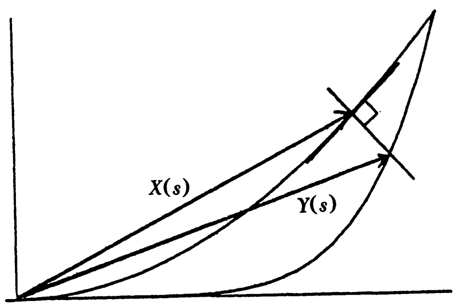


FIGURE 2

This argument generalizes to show that any (rectifiable) curve connecting the endpoints of a smooth convex curve and lying below it will be longer than the original curve. As in the special case above, the convexity implies that $k(s) > 0$ and the relative position of the two curves is expressed by the sign of $a(s)$. Of course, the proof works just as well if both $k(s)$ and $a(s)$ are negative, the case of a curve lying above a concave curve.

A key point in this argument that often is not emphasized in a calculus course is that curvature is the intrinsic measure of convexity for a curve, i.e., the measure of how fast the curve is bending away from its tangent line. For the curve $y = g(x)$ the curvature k at x can be computed in terms of derivatives of g by the formula $k(x) = g''(x)/(1 + g'(x)^2)^{3/2}$. Notice that it is k that gives real geometric information about the graph of g (though of course the sign of g'' does indicate convexity versus concavity).

We do not see an easy way to relate the lengths of two such curves without using differential geometry. It would be interesting to find a conceptually different approach.

REFERENCES

1. Jerrold Marsden and Alan Weinstein, *Calculus*, Benjamin/Cummings Publishing Co., Menlo Park, CA, 1980.
2. George Simmons, *Calculus with Analytic Geometry*, McGraw Hill Book Co., New York, 1985.

How Much Search is Enough?

JET YEE
Economic Research Service
Washington, DC 20005

Introduction Many students may regard the derivation of the distributions of the maximum and minimum values in a random sample to be just a technical exercise (see [2]). There are, however, many interesting economic applications of these distributions. We consider the problem of a consumer searching for the lowest price for a product and a similar example in which an unemployed job seeker searches for the highest wage rate. As will become apparent, these two examples possess a common structure.

Searching for the lowest price Suppose that a consumer faces a distribution of prices for a particular product. Assume that the consumer wants to buy just one unit of the product no matter the price. Different stores charge different prices. It is assumed that the consumer *knows* the distribution of price, but does not know, without searching, which stores are charging which prices. The problem is to determine how many stores the consumer should consult before buying.

Assume that each search (i.e. each price encountered) is an independent random variable from the known distribution of prices. Let $f(p)$ be the known probability density function (p.d.f.) of prices ($0 < p < \infty$) with F its distribution function (d.f.). Let M_n be a random variable defined as $M_n = \min\{p_1, \dots, p_n\}$. That is, M_n is the minimum price encountered after n searches. The d.f. of M_n is $G_M(p) = 1 - [1 - F(p)]^n$. For any random variable X that can take on only nonnegative values, the expected value of X is

$$E(X) = \int_0^\infty x f(x)\, dx = \int_0^\infty [1 - F(x)]\, dx.$$

The expected value of M_n after n searches is then

$$\begin{aligned} E(M_n) &= \int_0^\infty [1 - G_M(p)]\, dp \\ &= \int_0^\infty [1 - F(p)]^n\, dp. \end{aligned}$$

Should the consumer make one additional search? Let

$$M_{n+1} = \min\{p_1, \dots, p_n, p_{n+1}\}.$$

The expected value of M_{n+1} after $(n+1)$ searches is

$$E(M_{n+1}) = \int_0^\infty [1 - F(p)]^{n+1}\, dp.$$

The expected gain, i.e. saving in price, is

$$s_n = E(M_n) - E(M_{n+1}) = \int_0^\infty [1 - F(p)]^n F(p)\, dp. \tag{1}$$

Suppose that the cost of making one additional search (time, transportation costs, etc.) is a constant, c. The consumer should engage in additional search as long as the expected saving in the form of a lower price exceeds c, i.e., $s_n > c$. Since $[1 - F(p)]^n$ and $F(p)$ in (1) are both between 0 and 1, the expected gain is nonnegative, but decreases with n. There is a maximum value of n (call it n^*) such that $s_{n+1} < c < s_n$, at which point it will not pay to engage in additional search. Thus the rule is to search n^* stores and choose the minimum price of the n^* prices encountered. This model is due to Stigler [3]. Note that the optimal number of searches, n^*, can be determined before the search even begins.

This rule however suffers from one important drawback. Since the consumer is assumed to know the distribution of prices, he should recognize a good price when he sees one. The consumer can do better by using the information from the search as it occurs. As we will now show, a better rule is to select a price r and buy at the first store that charges a price no greater than r. We will now proceed to discuss how to determine r.

Let m be the minimum price so far obtained after n searches and p be the price encountered on the $(n+1)$th search. If $p < m$, then the gain from one additional search is $m - p$. If $p > m$, then the gain from one additional search is 0 (the consumer will ignore it). The expected gain s_n from one additional search is

$$\begin{aligned} s_n &= \int_0^m (m-p) f(p)\, dp \\ &= \int_0^m F(p)\, dp \equiv h(m) \end{aligned}$$

on performing integration by parts.

The consumer should engage in additional search as long as $h(m) > c$. By Leibniz's rule

$$\frac{dh(m)}{dm} = \frac{d}{dm}\int_0^m F(p)\, dp = F(m) > 0.$$

So since m is non-increasing in n, $h(m)$ is non-increasing in n. There is a maximum value of n and a corresponding value of m (call it r) such that

$$h(r) = \int_0^r F(p)\, dp \geqslant c, \tag{2}$$

at which point it will not pay to engage in additional search.

That is, the optimal strategy is to search until a price no greater than r is encountered. Note that r is established (from (2)) before the search even begins. The value of r depends on c as well as the shape of the distribution of prices, F. This model of search (in a job search context) is due to McCall [1].

The framework we have discussed can be used to elucidate a variety of observed market phenomena. For example, consumers typically search longer for a new car than for candy (because the expected saving is greater for a car) and full-time workers typically search less than students for the same product (because the cost in terms of foregone income from not working is greater for full-time workers).

Searching for the highest wage rate Now consider the problem of a person who is looking for a new job. Assume that the person *knows* the distribution of wage rate offers, but does not know, without searching, which employers are offering which wage rate offers. The problem is to determine how long (i.e., how many searches) the person should spend in looking.

Assume that each search (i.e. each wage rate offer encountered) is an independent random variable from the known distribution of wage rate offers. Also assume that previous wage rate offers remain open to the person and that the cost of making one additional search is a constant.

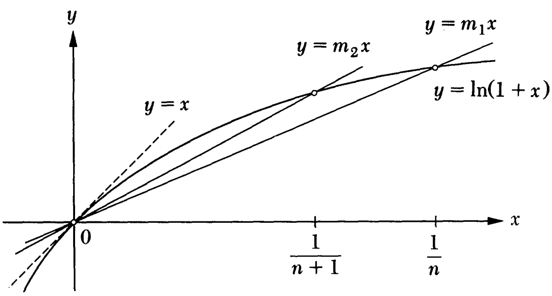
This example is conceptually similar to that of the consumer example. It can be verified that there is again an optimal number of searches to make. It is, however, more realistic to assume that the person will take into consideration the wage rate offers that he has already encountered in deciding whether to make an additional search. It can also be verified that a better rule is to select a wage rate R and accept the first job paying at least R.

REFERENCES

1. McCall, John J., Economics of information and job search, *Quarterly J. of Economics* 84 (1970), pp. 113–26.
2. Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes, *Introduction to the Theory of Statistics*, 3rd edition, McGraw-Hill Book Co., New York, 1974.
3. Stigler, George, The economics of information, *J. of Political Economy* 69 (1961), pp. 213–25.

Proof without Words: A Monotone Sequence Bounded by e

$$\forall n \geq 1, \left(1+\frac{1}{n}\right)^n < \left(1+\frac{1}{n+1}\right)^{n+1} < e.$$



$$n \geq 1 \Rightarrow m_1 < m_2 < 1$$

$$\Rightarrow \frac{\ln\left(1+\frac{1}{n}\right)}{\frac{1}{n}} < \frac{\ln\left(1+\frac{1}{n+1}\right)}{\frac{1}{n+1}} < 1$$

$$\Rightarrow \left(1+\frac{1}{n}\right)^n < \left(1+\frac{1}{n+1}\right)^{n+1} < e$$

—ROGER B. NELSEN
LEWIS AND CLARK COLLEGE
PORTLAND, OR 97219

wage rate offers. The problem is to determine how long (i.e., how many searches) the person should spend in looking.

Assume that each search (i.e. each wage rate offer encountered) is an independent random variable from the known distribution of wage rate offers. Also assume that previous wage rate offers remain open to the person and that the cost of making one additional search is a constant.

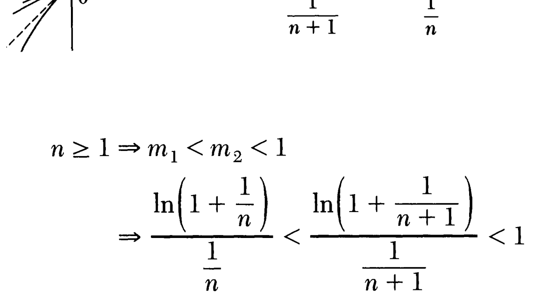
This example is conceptually similar to that of the consumer example. It can be verified that there is again an optimal number of searches to make. It is, however, more realistic to assume that the person will take into consideration the wage rate offers that he has already encountered in deciding whether to make an additional search. It can also be verified that a better rule is to select a wage rate R and accept the first job paying at least R.

REFERENCES

1. McCall, John J., Economics of information and job search, *Quarterly J. of Economics* 84 (1970), pp. 113–26.
2. Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes, *Introduction to the Theory of Statistics*, 3rd edition, McGraw-Hill Book Co., New York, 1974.
3. Stigler, George, The economics of information, *J. of Political Economy* 69 (1961), pp. 213–25.

Proof without Words: A Monotone Sequence Bounded by e

$$\forall n \geq 1, \left(1+\frac{1}{n}\right)^{n} < \left(1+\frac{1}{n+1}\right)^{n+1} < e.$$



$$n \geq 1 \Rightarrow m_1 < m_2 < 1$$

$$\Rightarrow \frac{\ln\left(1+\frac{1}{n}\right)}{\frac{1}{n}} < \frac{\ln\left(1+\frac{1}{n+1}\right)}{\frac{1}{n+1}} < 1$$

$$\Rightarrow \left(1+\frac{1}{n}\right)^{n} < \left(1+\frac{1}{n+1}\right)^{n+1} < e$$

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

Coin Flipping, Dynamical Systems, and the Fibonacci Numbers

STEPHEN F. KENNEDY[1]
Saint Olaf College
Northfield, MN 55057

MATTHEW W. STAFFORD[2]
University of Texas
Austin, TX 78712

In this note we pose a pair of seemingly disparate problems, one from probability and the other from dynamical systems. We will show that these two problems are actually one—i.e., their solutions are identical and each, surprisingly, involves the Fibonacci numbers. The dynamical systems problem is new. Variants of the probability problem have been around for centuries. The oldest reference we have found is Problem LXXIV in Abraham De Moivre's *The Doctrine of Chances* [3], the first edition of which appeared in 1718. (See [6] for a treatment of De Moivre's formulation with a more modern flavor.) After solving both problems and finding the connection between them, we will mention some simple generalizations.

Suppose we are asked to flip a fair coin repeatedly until two consecutive heads appear. What is the probability that this will ever occur? And, more specifically: What is the probability that it will occur in exactly n tosses? The dynamical systems problem arose during the second author's dissertation research [1, 5]. If $f(x)=2x \pmod 1$, for $x \in [0,1)$, where are all the points that visit $[3/4,1)$ for the first time after exactly n iterations of $f(x)$? That is, let $A_n = \{x \mid f^n(x) \in [3/4,1)$, and n is minimal with respect to this property$\}$. What is the geometry of A_n? Is it the case that $\bigcup_{n=0}^{\infty} A_n = [0,1)$? We note that $f^n(x)$ denotes the composition $f \circ f \circ f \circ \cdots \circ f(x)$, of n copies of f.

A moment's thought should convince you that the probability of seeing two heads in a row in the coin flipping problem is very large. Is it one? Represent a sequence of tosses by a sequence of H's and T's in the obvious way. Call a sequence *n-allowable* if it ends the coin tossing after n flips, i.e., has 'HH' as its last two letters and no earlier occurrence of this string. Since all sequences are equally probable, the probability of stopping after n tosses is the number of n-allowable sequences (call this a_n) divided by 2^n. To count the n-allowable sequences notice that every n-allowable sequence is of one of two forms: a 'T' followed by an $(n-1)$-allowable sequence, or, an 'HT' followed by an $(n-2)$-allowable sequence. Moreover, appending 'T' to the beginning of any $(n-1)$-allowable sequence yields an n-allowable one, as does appending 'HT' to any $(n-2)$-allowable sequence. Therefore, $a_n = a_{n-1} + a_{n-2}$ and $a_2 = a_3 = 1$. That is, a_n is the $(n-2)^{\text{nd}}$ Fibonacci number. The probability of stopping after n tosses is then $a_n/2^n$, and the probability of *ever* stopping is $\sum_{n=2}^{\infty} a_n/2^n$. We expect that this sum must converge to one: We'll prove this momentarily.

To solve the dynamical systems problem, it pays to work in dyadic expansions. If $.101101\ldots$ is the dyadic expansion of x, then $.01101\ldots$ is the dyadic expansion of

[1]Much of the work in this paper was performed while the author was a postdoctoral visitor at the University of Delaware.

[2]The author acknowledges the support of The Institute for Mathematics and Its Applications, University of Minnesota.

$f(x)$ ($f(x)$ "forgets" the first digit of x's dyadic expansion). Any number in $[3/4, 1)$ has one as the first two digits of its dyadic expansion, therefore a number that visits this interval for the first time after n iterations of f must have the string '11' in its $(n+1)^{\text{st}}$ and $(n+2)^{\text{nd}}$ places and no earlier occurrences of this string. The number of different $(n+2)$-length strings that could begin the expansion of a point in A_n is equal to the number of $(n+2)$-allowable words in the coin flipping problem. The collection of points that have a specified $(n+2)$-length string at the beginning of their expansion is an interval of length $1/2^{n+2}$. Therefore, A_n is a collection of a_{n+2} disjoint (why?) intervals each of this length. If $n \neq m$, then $A_n \cap A_m = \varnothing$ and the length of $\bigcup_{n=0}^{\infty} A_n = \sum_{n=0}^{\infty} a_{n+2}/2^{n+2}$, a reindexing of the earlier sum.

As we shall see, this converges to one, however, not every point eventually maps into $[3/4, 1)$. For example, $1/3 = .010101\ldots$, $f(1/3) = .101010\ldots = 2/3$, and $f(2/3) = 1/3$. In some sense these points correspond to coin flipping sequences that never terminate (e.g., HTTHTT...), but we have some extra structure in this situation. The set $[0, 1)$ inherits a topology as a subset of the real line and we have a function, $f(x)$, which is *invariant* on the set of points that never visit $[3/4, 1)$. (Invariant means if x is a point in this set, then $f(x)$ is also.) In this topology our set is a Cantor set and together with $f(x)$ forms a chaotic dynamical system [4].

To prove that $\sum_{n=2}^{\infty} a_n/2^n = 1$, first notice that by the ratio test the sum does converge. Let

$$\begin{aligned}
S &= \tfrac{1}{4} + \tfrac{1}{8} + \tfrac{2}{16} + \tfrac{3}{32} + \tfrac{5}{64} + \cdots \\
\tfrac{1}{2}S &= \phantom{\tfrac{1}{4} + {}} \tfrac{1}{8} + \tfrac{1}{16} + \tfrac{2}{32} + \tfrac{3}{64} + \cdots \\
\tfrac{3}{2}S &= \tfrac{1}{4} + \tfrac{2}{8} + \tfrac{3}{16} + \tfrac{5}{32} + \tfrac{8}{64} + \cdots \\
&= 2S - \tfrac{1}{2} \\
\Rightarrow S &= 1.
\end{aligned}$$

This is quite remarkable—the sum whose i^{th} term is the i^{th} Fibonacci divided by 2^{i+2} converges to one! That we can even find the number to which this series converges is unusual; that its sum is the simplest rational seems miraculous. The explanation for this miracle lies in the so-called "generating function" for the Fibonacci numbers. Let $g(x)$ be the function whose power series expansion has n^{th} term the n^{th} Fibonacci number times x^n. It can be shown relatively easily that [2]:

$$g(x) = \sum_{n=0}^{\infty} a_{n+2} x^n = \frac{1}{1 - x - x^2}.$$

Our series is $(1/4)g(1/2) = 1$.

De Moivre [3] asked for the probability of observing three consecutive heads in a sequence of 10 flips. More generally, what is the probability of seeing m consecutive heads in *exactly* n flips? For three consecutive heads the numbers of n-allowable words are $1, 1, 2, 4, 7, 13, \ldots$. The recursion relation is $a_n = a_{n-1} + a_{n-2} + a_{n-3}$. The proof is a slight modification of the original: We can append a 'T' onto any $(n-1)$-allowable word, an 'HT' onto any $(n-2)$-allowable word, or an 'HHT' onto an $(n-3)$-allowable word. If we form the power series whose n^{th} term is a_{n-3} times x^n, we get the generating function $g(x) = 1/(1 - x - x^2 - x^3)$. Notice that $(1/8)g(1/2) = 1$, the probability of ever seeing three heads is one. Table 1 lists the first few terms, the recursion relation, and the resulting generating function for $m = 1, 2, 3, 4, 5$. The patterns persist.

In solving these problems we have also solved the corresponding dynamical systems problems: What is the set of points that visits $[1 - 1/2^m, 1)$ after n iterations of $f(x)$? If we ask what happens if we have a p-sided coin and we want two consecutive heads, the results are summarized in Table 2. Again, it is not difficult to prove that the patterns persist.

TABLE 1

m	Terms	Recursion	Generating Function
1	$1,1,1,1,1,\ldots$	$a_n = a_{n-1}$	$g(x) = \dfrac{1}{1-x}$
2	$1,1,2,3,5,8,\ldots$	$a_n = a_{n-1} + a_{n-2}$	$g(x) = \dfrac{1}{1-x-x^2}$
3	$1,1,2,4,7,13,\ldots$	$a_n = a_{n-1} + a_{n-2} + a_{n-3}$	$g(x) = \dfrac{1}{1-x-x^2-x^3}$
4	$1,1,2,4,8,15,29,\ldots$	$a_n = a_{\ -1} + \cdots + a_{n-4}$	$g(x) = \dfrac{1}{1-x-x^2-x^3-x^4}$
5	$1,1,2,4,8,16,31,\ldots$	$a_n = a_{n-1} + \cdots + a_{n-5}$	$g(x) = \dfrac{1}{1-x-x^2-x^3-x^4-x^5}$

TABLE 2

p	Terms	Recursion	Generating Function
2	$1,1,2,3,5,8,\ldots$	$a_n = a_{n-1} + a_{n-2}$	$g(x) = \dfrac{1}{1-x-x^2}$
3	$1,2,6,16,44,\ldots$	$a_n = 2(a_{n-1} + a_{n-2})$	$g(x) = \dfrac{1}{1-2x-2x^2}$
4	$1,3,12,45,171,\ldots$	$a_n = 3(a_{n-1} + a_{n-2})$	$g(x) = \dfrac{1}{1-3x-3x^2}$
5	$1,4,20,96,464,\ldots$	$a_n = 4(a_{n-1} + a_{n-2})$	$g(x) = \dfrac{1}{1-4x-4x^2}$

The reader is urged to consider the problem of obtaining m consecutive heads from a p-sided coin. Also, the numbers obtained by finding the limits of quotients of consecutive terms for the various values of p have many nice properties, analogous to properties of the golden mean ($p = 2$).

Acknowledgement. The authors would like to thank the many people whose advice and suggestions contributed to this note: R. O. Young, W. T. Butterworth, R. Holmgren, and R. F. Williams who started it all.

REFERENCES

1. J. Ashley, B. Kitchens, and M. Stafford, Boundary sets of Markov partitions, *Trans. AMS*, (333), Sept. 1992, pp. 177–201.
2. G. Bergman and K. D. Fryer, *Introduction to Combinatorics*, Academic Press, London, 1972.
3. A. De Moivre, *The Doctrine of Chances*, 3rd edition, 1756, Chelsea Publishing Co., New York (reprint), 1967.
4. R. Devaney, *An Introduction to Chaotic Dynamical Systems*, 2nd edition, Addison-Wesley Publishing Co., Reading, MA, 1989.
5. M. Stafford, Markov partitions for the doubling map, Ph.D. Thesis, Northwestern University, Evanston, IL, 1989.
6. I. Todhunter, *A History of the Mathematical Theory of Probability*, 1865, Chelsea Publishing Co., New York (reprint) 1965.

A Powerful Procedure for Proving Practical Propositions

SOLOMON W. GOLOMB
University of Southern California
University Park, CA 90089-2565

Dedicated to Martin Gardner

The eminent Oxford don, Charles Lutwidge Dodgson, more widely known as Lewis Carroll, demonstrated the applicability of formal mathematical reasoning to real-life situations with such incontrovertible rigor as evidenced in his syllogism:

"All Scotsmen are canny.
All dragons are uncanny.
Therefore, no Scotsmen are dragons."

While his logic is impeccable, the conclusion (that no Scotsmen are dragons) is not particularly surprising, nor does it shed much light on situations that we are likely to encounter on a daily basis.

The purpose of this note is to exploit this powerful proof methodology, introduced by Professor Dodgson, to a broader range of human experience, with special emphasis on obtaining conclusions having political or moral significance.

THEOREM 1. *Apathetic people are not human beings.*
Proof. All human beings are different.
All apathetic people are indifferent.
Therefore, no apathetic people are human beings.

THEOREM 2. *All rude people are irrelevant.*
Proof. All relevant people are pertinent.
All rude people are impertinent.
Therefore, no rude people are relevant.

THEOREM 3. *All incomplete investigations are biased.*
Proof. Every incomplete investigation is a partial investigation.
Every unbiased investigation is an impartial investigation.
Therefore, no incomplete investigation is unbiased.

Numerous additional examples, closely following Dodgson's original model, could be adduced. However, our next objective is to broaden the approach to encompass other models of mathematical proof. For example, it is a well-established principle that a property P is true for *all* members of a set S if it can be shown to be true for an *arbitrary* member of the set S. We exploit this to obtain the following important result.

THEOREM 4. *All governments are unjust.*
Proof. To prove the assertion for *all* governments, it is sufficient to prove it for an *arbitrary* government. But if a government is arbitrary, it is obviously unjust. And since this is true for an arbitrary government, it is true for all governments.

Finding further theorems of this type, and extending the method to other models of mathematical proof, is left as an exercise for the reader.

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 May 1, 1995.

1459. *Proposed by D. M. Bloom, Brooklyn College of CUNY, Brooklyn, New York.*

The now notorious Newman-Conway sequence $1, 1, 2, 2, 3, 4, 4, 4, 5, 6, 7, 7, \ldots$ is defined by the recurrence $P(1) = P(2) = 1$, $P(n) = P(P(n-1)) + P(n - P(n-1))$, $(n \geq 3)$. Richard Guy, in this MAGAZINE, February 1990, p. 17, wrote: "I have an earlier manuscript of Conway in which he has written '$P(2^k) = 2^{k-1}$ (easy), $P(2n) \leq 2P(n)$ (hard), ...'." Prove Conway's "hard" inequality: $P(2n) \leq 2P(n)$.

1460. *Proposed by Doru Popescu Anastasiu, Slatina, Romania.*

Let ABC be an acute triangle with altitudes AA', BB', CC'. Let A_1, B_1, C_1 be the second intersection points of lines AA', BB', CC' with the circumcircle of triangle ABC. Show that

$$AA_1^2 \sin 2A + BB_1^2 \sin 2B + CC_1^2 \sin 2C > 24\, S_0,$$

where S_0 denotes the area of triangle $A'B'C'$.

1461. *Proposed by Václav Konečný, Ferris State University, Big Rapids, Michigan.*

Given the vertices V_1, V_2 and foci at F_1, F_2 of two parabolas with the same axis, construct a common tangent, if one exists, using only a compass and straightedge. Assume that the unit of length is given.

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.

1462. *Proposed by Arthur L. Holshouser and Benjamin G. Klein, Davidson College, Davidson, North Carolina.*

Let λ be a given positive number. Let $p(x) = ax^2 + bx + c$, where a, b, and c are nonzero real numbers. Assume that no roots of $p(x)$ lie on the line $\Lambda = \{t + i\lambda t: t \in \mathbf{R}\}$. Find a homogeneous fourth degree polynomial $H_\lambda(a, b, c)$ such that the number of roots of $p(x)$ that lie below Λ is given by

$$1 - \frac{\operatorname{sgn}(ab) + \operatorname{sgn}(H_\lambda(a, b, c))}{2}.$$

1463. *Proposed by Christos Athanasiadis, Massachusetts Institute of Technology, Cambridge, Massachusetts.*

Let n be a positive integer. For any partition λ of n and any $1 \le i \le n$, let $m_i = m_i(\lambda)$ be the number of i's in λ. Also let z_λ be the quantity $1^{m_1}m_1!2^{m_2}m_2! \cdots n^{m_n}m_n!$. Show that, for $1 \le k \le n$,

$$\sum_{\lambda \vdash n} \frac{m_k(\lambda)}{z_\lambda} = \frac{1}{k}$$

where, in the summation, λ runs over all partitions of n.

Quickies

Answers to the Quickies are on page 390.

Q826. *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada.*

Which of the two integrals

$$\int_0^1 \left(\sqrt{1 + x^{2r}\tan^2\alpha} - x^r \sec\alpha\right)^{1/s} dx, \qquad \int_0^1 \left(\sqrt{1 + x^{2s}\tan^2\alpha} - x^s \sec\alpha\right)^{1/r} dx$$

is larger, given that $r > s > 0$, and $\pi/2 > \alpha > 0$?

Q827. *Proposed by Herbert Gülicher, Westfälische Wilhelms Universität, Münster, Germany.*

For $a_{ij} > 0$, $i, j = 1, 2, 3$, assume $a_{i1} + a_{i2} + a_{i3} = 1$ and $a_{i1} = a_{i+1,3}$, where addition of subscripts is taken modulo 3. Prove that

$$a_{11} + a_{21} + a_{31} = 1 \quad \text{if, and only if,} \quad \frac{a_{11}a_{21}a_{31}}{a_{12}a_{22}a_{32}} = 1.$$

Q828. *Proposed by Ismor Fischer, University of Wisconsin, Madison, Wisconsin.*

Let $f(x)$ be a continuous, differentiable function, concave down on the interval $[a, b]$. Find the point $c \in (a, b)$ such that the combined area of the two inscribed trapezoids yields the best approximation to $\int_a^b f(x)\,dx$. (A similar question may be asked for $f(x)$ concave up.)

Solutions

Divisibility Condition **December 1993**

1433. *Proposed by Cristian Turcu, London, England.*

For $n \geq 1$, let a_n, b_n, c_n be nonzero integers, and suppose there are indices s and t such that $|a_s c_t - a_t c_s| = 1$. Let P_n denote the set of nonzero rational numbers x such that $(a_n x^2 + c_n)/b_n x$ is an integer. Prove that $\bigcap_{n=1}^{\infty} P_n \neq \varnothing$ if, and only if, b_n divides $a_n + c_n$ for each $n \geq 1$.

Solution by Walter Blumberg, Coral Springs, Florida.

Suppose b_n divides $a_n + c_n$ for each $n \geq 1$. It is then obvious that $x = 1$ belongs to P_n for each $n \geq 1$, so that $\bigcap_{n=1}^{\infty} P_n \neq \varnothing$.

Conversely, suppose that $\bigcap_{n=1}^{\infty} P_n \neq \varnothing$. Then there exists a nonzero rational number x which is a member of P_n for each $n \geq 1$. We may assume that x is positive, since if x is a member of P_n, so is $-x$. Let $x = u/v$, where u and v are relatively prime positive integers. Then for $n \geq 1$,

$$\frac{a_n(u/v)^2 + c_n}{b_n(u/v)} = \frac{a_n u^2 + c_n v^2}{b_n uv}$$

is an integer. Thus, uv divides $a_n u^2 + c_n v^2$. Since $\gcd(u, v) = 1$, it follows that u divides c_n and v divides a_n. Consequently, uv divides $a_s c_t - a_t c_s$, and hence, by our assumption, uv divides 1. This implies that $u = v = 1$, and therefore, $x = 1$. Thus, we have shown that for each $n \geq 1$, $(a_n + c_n)/b_n$ is an integer, or equivalently, that b_n divides $a_n + c_n$. This completes the proof.

Also solved by Aardvark Problem Solving Group (Trenton State College), J. Binz (Switzerland), Con Amore Problem Group (Denmark), Furman University Problem Solving Group, Richard Heeg, Hans Kappus (Switzerland), O. P. Lossers (The Netherlands), David E. Manes, F. C. Rembis, Heinz-Jürgen Seiffert (Germany), John S. Sumner, Trinity Univeristy Problem Group, A. N. 't Woord (The Netherlands), and the proposer.

Ring Homomorphisms and Carmichael Numbers **December 1993**

1434. *Proposed by Syd Bulman-Fleming and Edward T. H. Wang, Wilfrid Laurier University, Waterloo, Ontario, Canada.*

Let $\mathbf{Z}_n$ denote the ring of integers modulo n, $n > 1$. Consider the function $f\colon \mathbf{Z}_n \to \mathbf{Z}_n$ defined by $f(x) = x^n$ for all $x \in \mathbf{Z}_n$. Characterize those values of n for which f is a ring homomorphism.

Solution by David Callan, University of Wisconsin, Madison, Wisconsin.

If f is a homomorphism, then $f(1 + 1) = f(1) + f(1)$, hence $2^n \equiv 2 \pmod{n}$ and, by induction,

$$a^n \equiv a \pmod{n} \quad \text{for all } a. \tag{1}$$

Conversely, if (1) holds, then clearly f is a homomorphism. Now (1) certainly holds for n prime (Fermat's Little Theorem). The composite "pseudoprime" numbers for which (1) holds are called Carmichael numbers. The first few are 561, 1105, 1729, 2465. They have a well-known description as $\{n\colon n$ is squarefree, and $(p - 1)$ divides $n - 1$ for each prime p that divides $n\}$. (See [1] for information about Carmichael numbers.) Recently it has been shown [2] that there are infinitely many Carmichael numbers.

REFERENCES

1. Paulo Ribenboim, *The Little Book of Big Primes*, Springer-Verlag, New York, 1991.
2. Andrew Granville, "Primality testing and Carmichael numbers," *Notices*, Amer. Math. Society, Vol. 39, No. 7, (1992), pp. 696–700.

Also solved by Aardvark Problem Solving Group (Trenton State College), Brian D. Beasley, Ruth I. Berger, Gary F. Birkenmeier, Ron Martin Carroll, Con Amore Problem Group (Denmark), Bill Correll, Jr. (student), Michael W. Ecker, F. J. Flanigan, Furman University Problem Solving Group, Lee O. Hagglund, Richard Heeg, John Koker, H. K. Krishnapriyan, Francis C. Leary, O. P. Lossers (The Netherlands), Mihalis Maliakas, David E. Manes, John D. O'Neill, Richard F. Ryan, Arlo W. Schurle (Guam), R. P. Sealy (Canada), Adam William Sembrato, Lawrence Somer, Trinity University Problem Group, A. N. 't Woord (The Netherlands), and the proposer. There was one unsigned solution.

Inequality between Two Sets of Points on a Sphere **December 1993**

1435. *Proposed by Florin S. Pîrvănescu, Slatina, Romania.*

Let $A_1, \dots, A_n$ be points ($n \geq 3$) on the sphere $S(O, R)$ of radius R and center O, and let G be their centroid. Let M be an arbitrary point in the sphere having OG as a diameter, and let B_k be the other intersection of MA_k with the sphere $S(O, R)$. Show that

$$\sum_{k=1}^{n} MB_k \geq \sum_{k=1}^{n} MA_k.$$

Solution by Jiro Fukuta, Kamimakuwa, Shinsei-cho, Gifu-ken, Japan.

We will show that the result holds more generally for points on a hypersphere in n-dimensional space. For brevity, we assume that PQ, etc., denote lengths of the segments PQ, etc. Also, $u \cdot v$ and $\|u\|$ will denote the inner product and the associated norm, respectively.

By the chord theorem of the hypersphere,

$$A_kM \cdot B_kM = R^2 - OM^2 \geq R^2 - (OG^2 - MG^2), \qquad k = 1, 2, \dots, n, \tag{1}$$

since M is in the hypersphere with diameter OG. Also,

$$\begin{aligned} nR^2 &= \sum_{i=1}^{n} OA_i^2 = \sum_{i=1}^{n} \|OA_i\|^2 = \sum_{i=1}^{n} \|OG + GA_i\|^2 \\ &= nOG^2 + 2\sum_{i=1}^{n} OG \cdot GA_i + \sum_{i=1}^{n} GA_i^2 \\ &= nOG^2 + \sum_{i=1}^{n} GA_i^2 \end{aligned}$$

since $\sum_{i=1}^{n} GA_i = 0$ by definition of G. Whence

$$R^2 - OG^2 = \frac{1}{n}\sum_{i=1}^{n} GA_i^2.$$

Substituting this equality into (1), we have

$$A_kM \cdot B_kM \geq \frac{1}{n}\sum_{i=1}^{n} GA_i^2 + MG^2, \qquad k = 1, 2, \dots, n.$$

Dividing by A_kM and summing over k yields

$$\sum_{k=1}^{n} B_kM \geq \frac{1}{n}\sum_{k=1}^{n} A_kM^{-1}\sum_{i=1}^{n} GA_i^2 + MG^2\sum_{k=1}^{n} A_kM^{-1}. \tag{2}$$

But

$$\begin{aligned}\sum_{i=1}^{n} GA_i^2 &= \sum_{i=1}^{n} \|GA_i\|^2 = \sum_{i=1}^{n} \|GM + MA_i\|^2 \\ &= nGM^2 + 2\sum_{i=1}^{n} GM \cdot MA_i + \sum_{i=1}^{n} MA_i^2 \\ &= nGM^2 + 2\sum_{i=1}^{n} GM \cdot (MG + GA_i) + \sum_{i=1}^{n} MA_i^2 \\ &= nGM^2 + 2\sum_{i=1}^{n} GM \cdot MG + \sum_{i=1}^{n} MA_i^2.\end{aligned}$$

Therefore, $\sum_{i=1}^{n} GA_i^2 = \sum_{i=1}^{n} MA_i^2 - nGM^2$. Substituting this into (2) and changing the index of summation,

$$\sum_{k=1}^{n} B_k M \geq \frac{1}{n}\left(\sum_{k=1}^{n} (A_k M)^{-1}\right)\left(\sum_{k=1}^{n} A_k M^2\right).$$

Applying the arithmetic-harmonic mean inequality,

$$\sum_{k=1}^{n} B_k M \geq n\left(\sum_{k=1}^{n} A_k M\right)^{-1}\left(\sum_{k=1}^{n} A_k M^2\right).$$

Finally, Schwartz's inequality implies that

$$\sum_{k=1}^{n} B_k M \geq n\left(\sum_{k=1}^{n} A_k M\right)^{-1} \frac{1}{n}\left(\sum_{k=1}^{n} A_k M\right)^2 = \sum_{k=1}^{n} A_k M.$$

Also solved by Con Amore Problem Group (The Netherlands), Nick Lord (England), and the proposer.

Murray S. Klamkin notes that a more general version of this problem appears in *Crux Mathematicorum*, 1086, Vol. 13 (1987), 100–102.

Limit of Complex Numbers **December 1993**

1436. *Proposed by Randall K. Campbell-Wright, University of Tampa, Tampa, Florida.*

Suppose that $(z_n)_{n=1}^{\infty}$ is a bounded sequence of complex numbers and that $\lim_{n\to\infty}\alpha_n = 1$. If k is a positive integer such that $\lim_{n\to\infty}(z_n - \alpha_n z_{n+k}) = \beta$, prove that $\beta = 0$.

I. *Solution by Nick Lord, Tonbridge School, Kent, England.*

By the Bolzano-Weierstrass Theorem, there is a subsequence $(z_{n_i})_{i=1}^{\infty}$ that converges; say $z_{n_i} \to z$ as $i \to \infty$. Since $z_{n_i} - \alpha_{n_i} z_{n_i+k} \to \beta$ *as* $i \to \infty$, we deduce that $z_{n_i+k} \to z - \beta$.

Applying the same argument to $z_{n_i+k} - \alpha_{n_i+k} z_{n_i+2k} \to \beta$ we see that $z_{n_i+2k} \to z - 2\beta$ and, inductively, that $z_{n_i+r_k} \to z - r\beta$ *for* $r = 3, 4, 5, \ldots$. But this contradicts the fact that $(z_n)_{n=1}^{\infty}$ is bounded unless $\beta = 0$.

II. *Solution by Michael Golomb, Purdue University, Lafayette, Indiana.*

Since $(z_n)_{n=1}^{\infty}$ is bounded and $\lim_{n\to\infty}(1-\alpha_n) = 0$, we have $\lim_{n\to\infty}(1-\alpha_n)z_{n+k} = 0$. Then, from $\lim_{n\to\infty}(z_n - \alpha_n z_{n+k}) = \beta$ it follows that $\lim_{n\to\infty}(z_{n+k} - z_n) = -\beta$. Hence,

$$z_{n+k} = z_n - \beta + w_n, \qquad \text{where } \lim_{n\to\infty} w_n = 0.$$

Thus, for any positive integer m,

$$z_{n+mk} = z_n - m\beta + \sum_{j=0}^{m-1} w_{m+jk}, \quad |z_{n+mk} - z_n| \geq m|\beta| - \sum_{j=0}^{m-1} |w_{n+jk}|. \quad (*)$$

Now suppose $|\beta| = b > 0$. Then choose N so large that $|w_n| \leq b/2$ for $n \geq N$, thus $\sum_{j=0}^{m-1} |w_{N+jk}| \leq mb/2$. Then by $(*)$, $|z_{N+mk} - z_N| \geq mb/2$. Since this holds for all positive integers m, we have a contradiction to the supposition that $(z_n)_{n=1}^{\infty}$ is bounded. We must conclude that $\beta = 0$.

Addendum. Set $z_n - \alpha_n z_{n+k} = \beta_n$. The assertion $\lim_{n\to\infty}\beta_n = 0$ is valid also if the critical hypothesis of the existence of $\lim_{n\to\infty}\beta_n$ is replaced by the much weaker condition that $\beta_n \in B$ for all sufficiently large n, where B is a wedge in $\mathbb{C}$ with vertex at 0 and argument less than $\pi/2$ in absolute value, provided that the condition $\lim_{n\to\infty}(1 - \alpha_n) = 0$ is replaced by $\sum_{n=1}^{\infty} |1 - \alpha_n| < \infty$.

For the proof of this, we may assume without loss of generality,

$$\beta_n \in B \text{ for } n = 1, 2, \ldots; \qquad B = \{z \in \mathbb{C}\colon |\arg(z)| \leq \pi/2 - \delta, 0 < \delta \leq \pi/2\}.$$

For any positive integer m, we have

$$z_n - z_{n+mk} = \sum_{j=0}^{m-1} \beta_{n+jk} - \sum_{j=0}^{m-1} (1 - \alpha_{n+jk}) z_{n+(j+1)k},$$

$$|\mathrm{Re}(z_n - z_{n+mk})| \geq \sum_{j=0}^{m-1} \mathrm{Re}(\beta_{n+jk}) - \sup_{1 \leq j \leq m} |z_{n+jk}| \sum_{j=0}^{m-1} |1 - \alpha_{n+jk}|.$$

This inequality shows that $(z_n)_{n=1}^{\infty}$ cannot be bounded unless the sum of nonnegative terms $\sum_{j=0}^{\infty} \mathrm{Re}(\beta_{n+jk})$ is convergent. Thus we conclude that $\lim_{m\to\infty} \mathrm{Re}(\beta_{n+mk}) = 0$. Since this is true for $n = 1, 2, \ldots, m-1$, we conclude that $\mathrm{Re}(\beta_n) = 0$. By hypothesis, $z_n \in B$, so $|\mathrm{Im}(\beta_n)| \leq (\cot\delta)\mathrm{Re}(\beta_n)$, and thus $\lim_{n\to\infty}\beta_n = 0$.

Also solved by Michael H. Andreoli, Walter Blumberg, Con Amore Problem Group (Denmark), James P. Crawford, Furman University Problem Solving Group, K. P. Hart (The Netherlands), Richard Heeg, Ben Jacobs, Kee-Wai Lau (Hong Kong), O. P. Lossers (The Netherlands), Andreas Müller (Switzerland), F. C. Rembis, Heinz-Jürgen Seiffert (Germany), John S. Sumner, Nora S. Thornber, A. N. 't Woord (The Netherlands), Wyoming Problem Circle, and the proposer.

Converse of Eisenstein Criterion **December 1993**

1437. *Proposed by Yuanan Diao, Kennesaw State College, Marietta, Georgia.*

Let $\mathbf{Q}[x]$ and $\mathbf{Z}[x]$ be the polynomial rings over the rational numbers and integers respectively. If $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbf{Z}[x]$ satisfies the Eisenstein Criterion, i.e., there is a prime integer p such that $p \nmid a_n, p | a_i, 0 \leq i \leq n-1$, and $p^2 \nmid a_0$, then $f(x)$, hence $af(bx + c)$, is irreducible in $\mathbf{Q}[x]$ for any rational numbers a, b, c, where $ab \neq 0$. Prove or disprove the converse: If $f(x) \in \mathbf{Q}[x]$ is irreducible, then there exist rational numbers $a, b, c, (ab \neq 0)$ such that $af(bx + c) \in \mathbf{Z}[x]$ and satisfies the Eisenstein Criterion.

Solution by Nick Lord, Tonbridge School, Kent, England.

The converse is false, for consider the irreducible polynomial $f(x) = x^2 + 4$. Reducing fractions to lowest form, we see that $g(x) = Af(Bx + C) \in \mathbf{Z}[x]$, $A, B, C \in \mathbf{Q}$, if, and only if, it can be written in the form

$$g(x) = ab^2x^2 + 2abcx + a(c^2 + 4d^2) \qquad a, b, c, d \in \mathbf{Z}.$$

Suppose that Eisenstein's criterion with prime p "works" on $g(x)$. Then $p \nmid ab^2$, so that $p \nmid a$ and $p \nmid b$. Further, $p|2c$ and $p|(c^2 + 4d^2)$, and $p^2 \nmid (c^2 + 4d^2)$.

Now, either $p = 2$, in which case from $p|(c^2 + 4d^2)$ we see that c is even and hence $p^2|(c^2 + 4d^2)$, a contradiction to Eisenstein's hypothesis, or $p > 2$, in which case $p|c$ and, from $p|(c^2 + 4d^2)$, we also have $p|d$, so that $p^2|(c^2 + 4d^2)$, again a contradiction to Eisenstein's hypothesis.

We conclude that $Af(Bx + C)$ does not satisfy the Eisenstein Criterion for any rational numbers A, B, C, where $AB \neq 0$.

Also solved by Con Amore Problem Group (The Netherlands), O. P. Lossers (The Netherlands), David E. Manes, A. N. 't Woord (The Netherlands), and the proposer. Other counterexamples were $x^3 - x - 1$, $x^3 + x + 1$, *and* $x^3 + 9$.

Answers

Solutions to the Quickies on page 385.

A826. The first integral equals the area in the first quadrant bounded by the x and y axes and the curve

$$(y^s + x^r \sec \alpha)^2 = 1 + x^{2r} \tan^2 \alpha,$$

which simplifies to

$$y^{2s} + 2y^s x^r \sec \alpha + x^{2r} = 1.$$

Since the mirror image of the latter curve across the $x = y$ line is

$$y^{2r} + 2y^r x^s \sec \alpha + x^{2s} = 1,$$

the two integrals are equal in value.

A827. Use the second set of identities to rewrite the first set as

$$a_{i2} = 1 - a_{i1} - a_{i-1,1} = 1 - s + a_{i+1,1},$$

where $s = a_{11} + a_{21} + a_{31}$. Then $a_{11}a_{21}a_{31} = a_{12}a_{22}a_{32}$ if, and only if,

$$a_{11}a_{21}a_{31} = (1 - s + a_{11})(1 - s + a_{21})(1 - s + a_{31}).$$

Since all factors are positive, the latter equality holds if, and only if, $s = 1$.

A828. The point is that characterized by the Mean Value Theorem. Let $A = (a, f(a))$, $B = (b, f(b))$, and $C = (c, f(c))$ for any $c \in (a, b)$. We wish to minimize the total area between $f(x)$ and chords $\overline{AC}$ and $\overline{BC}$, and this amounts to maximizing area ΔABC. Since the base $|\overline{AB}|$ is constant, this occurs when the triangle's height is maximized, that is, when $f'(c) =$ slope of $\overline{AB}$.

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.

Cipra, Barry, *What's Happening in the Mathematical Sciences*, vol. 2, American Mathematical Society, 1994; 52 pp, $8 (P) (discounts for quantity orders, down to $4 each for 50 or more). ISBN 0-8218-8998-2.

Beautifully illustrated vignettes of 10 developments in mathematics over the past year or two, treating Fermat's Last Theorem, knots, compact waves, medical imaging, wavelets, random algorithms, soap bubbles, nonlinear codes, Latin rectangles, and the Seifert conjecture on vector fields. The writing is easy to read, and this booklet vividly conveys that mathematics is a living science with exciting developments taking place. (Note: You may already order copies of next year's volume, due out September 1995 at this year's prices.)

Hales, Thomas C., The status of the Kepler conjecture, *Mathematical Intelligencer* 16 (3) (1994) 47–58.

In 1990, Wu-Yi Hsiang (University of California—Berkeley) claimed to have proved that the optimal packing of spheres in 3-space is the face-centered cubic packing. This article reviews Hsiang's preprint and its revisions and notes that many experts have concluded that "his work does not merit serious consideration."

Berlekamp, Elwyn, and David Wolfe, *Mathematical Go: Chilling Gets the Last Point*, A K Peters, 1994; xx + 235 pp, $29.95. ISBN 1-56881-032-6. Also available in paperback as *Mathematical Go Endgames: Nightmares for the Professional Go Player*, Ishi Press International, 0-923891-36-6.

Many mathematicians know how to play Go, whose nature and rules are distinctly geometrical. More so than chess, Go would seem susceptible to fruitful mathematical analysis. This volume investigates positions in the endgame and applies combinatorial game theory to them. It also includes a discussion of the variety of rules under which the game has been played. Although "Go players can find quicker ways to improve their game" and "the most the reader can hope for is to get stronger by a single point," this book may mark the beginning of further mathematical analysis of Go and greater attention to combinatorial game theory.

Peterson, Ivars, Beating a fractal drum: How a drum's shape affects its sound, *Science News* 146 (17 September 1994) cover, 179, 184–185.

How much can you tell about a drum from listening to it? From the spectrum of normal modes of vibration of the drum, you can "hear" the drum's area, perimeter, and number of holes—but not its geometrical shape. That drums of different shapes may have identical spectra was proved by mathematicians in 1991 but has now been tested experimentally by physicists. Experiments showing that convoluted boundaries dampen vibrations have in turn led mathematicians to investigate "fractal drums," ones with fractal boundaries.

Golomb, Solomon W., *Polyominoes: Puzzles, Patterns, Problems, and Packings*, 2nd ed., Princeton University Press, 1994; xii + 184 pp, $24.95. ISBN 0-691-08753-0.

Polyominoes are connected planar arrays of squares, some of which are joined along their edges. This new edition of the 1965 classic book on the subject includes two new chapters, on tiling rectangles and on "truly remarkable" results, and 12 unsolved problems recommended for "readers' research." What's missing in this edition is—too bad!—the set of blue plastic pentominoes that came with the first edition.

Zaslavsky, Claudia, *Fear of Math: How to Get Over It and Get On with Your Life*, Rutgers University Press, 1994; x + 264 pp, $37, $14.95 (P). ISBN 0-8135-2090-8, 0-8135-2099-1.

With a broader approach than most "math anxiety" books, this one identifies sources and causes of negative feelings toward mathematics, suggests how to change attitudes, and provides lists of resources. Importantly, it stresses to the reader why it is important to overcome phobias of mathematics, citing examples of misuses of mathematics and statistics.

Paul, Richard, *Critical Thinking: What Every Person Needs to Survive in a Rapidly Changing World*, 2nd ed., Foundation for Critical Thinking (4655 Sonoma Mountain Road, Santa Rosa, CA 95404), 1992; xvi + 673 pp, $19.95 (P). ISBN 0-944583-07-5.

Many U.S. colleges and universities have some form of "quantitative reasoning" graduation requirement, however eclectically or weakly defined. As accrediting boards of U.S. higher-education institutions force colleges to focus on assessment of student outcomes, the nature of such a requirement—and what results it produces—will come under close scrutiny. Students not too anxious about math usually satisfy such a requirement by taking a minimal mathematics course, thereby bolstering enrollment and low-level employment in the mathematics department. A growing trend, however, is to broaden, supplement, or replace such a requirement with one in *critical thinking*. St. John's University (Minnesota) has even founded an endowed chair in the subject. Less catchy but more descriptive than author Paul's definition is the one used as the basis for the California State University requirement: Critical thinking is "the ability to analyze, criticize, and advocate ideas, to reason inductively and deductively, and to reach factual or judgmental conclusions based on sound inferences drawn from unambiguous statements of knowledge or belief." This book is a compendium of Paul's essays on the nature of critical thinking and how to teach for it; his foundation also publishes handbooks and videotapes on teaching strategies, and sponsors regional workshops and an annual conference. As far as I can tell—there is no index—*mathematics and statistics are not mentioned at all* in this book. What role can/should/will mathematics—and college mathematics faculty and departments—have in moving students to better critical thinking? (Note: The third edition has just appeared.)

Grattan-Guinness, I. (ed.), *Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences*, Routledge, 1994; 2 vols., xxiv + 1806 pp. ISBN 0-415-03785-9.

This two-volume set is an extraordinary endeavor, a joy to dip into and read around for a while. More than 180 essays, by almost as many authors, treat various mathematical and epimathematical topics. By intention, the essays terminate at some time between the two World Wars, and there are no biographical articles. The articles are not overly technical, but faithfully serve the purpose of providing introductions to branches of the history of mathematics. Apart from the areas of pure mathematics and periods of history that you would expect, one of the 13 sections is devoted to each of mechanics, mathematical physics, probability and statistics, and institutions of higher education in various countries.

Adler, Jerry, The numbers game, *Newsweek* (25 July 1994) 56–58.

Do 22% of Americans believe that the Holocaust may not have happened—or is it less than 1%? Are 600,000 Americans homeless, or 2.5 million? Are 50,000 American children abducted by strangers each year—or is it 5,000? Each year, are 4 million American women assaulted by a "domestic partner"—or is it 2 million? or 16 million? Many statistics come from sources "less interested in precisely measuring a given problem than in showing that it's even worse than anyone thought ... Advocacy groups tend to consider questions about their statistics as tantamount to an attack on their goals"—and may deliberately inflate statistics to advance their cause. Meanwhile, what about the folks who bring you the numbers? "[J]ournalists ... consider their job done when they find a number that can be attributed to a credible source." This article totally wimps out at the end, with the weak suggestion that "it would be better if journalists were more skeptical of statistics" and the ridiculous claim that the data are irrelevant anyway.

Maor, Eli, *e: The Story of a Number*, Princeton Univ. Pr., 1994; xiv + 223 pp, $24.95. ISBN 0–691–03390–0. Maor, Eli, The story of *e*, *The Sciences* (July/August 1994) 24–29.

Pi regularly has its day, and its appeal is almost universal; but *e* tends to be encountered only by the more limited audience who study precalculus. Hence what a delight it is at last to see a book (plus a derivative article) devoted to *e*: its rich history; its manifestations in geometry, number theory, trigonometry, and calculus; and its applications in logarithms, exponential growth and decay, and logarithmic scales.

Cannell, D.M., *George Green: Mathematician and Physicist 1793–1841; The Background to His Life and Work*, Athlone Press, 1993; xxxvi + 255 pp, $70. ISBN 0–485–11433–X. Cannell, D.M., and N.J. Lord. George Green, mathematician and physicist 1793–1841. *Mathematical Gazette* 66 (1993) 26–51.

George Green's name is known to students of vector calculus, but they are unlikely to know that he was a miller by trade and a self-taught genius who did his best work before going to university at age 40. The book (from which the article is adapted) gives a pleasant tale of his life and circumstances and includes an account (by M.C. Thornley) of his mathematics.

Soifer, Alexander, *Colorado Mathematical Olympiad: The First 10 Years and Further Explorations*, Center for Excellence in Mathematical Education (858 Red Mesa Dr., Colorado Springs, CO 80906); x + 188 pp, $19.95 (P). ISBN 0–940263–03–3.

This book, by a former judge of the Soviet Union Mathematical Olympiad, offers not only the problems and solutions from the first 10 years of the Colorado Mathematical Olympiad, but also a history of how he founded the contest (with lots of help!), photos of some of the winners, and 40 pages of further mathematical explorations inspired by the problems. In particular, the mathematics may only be beginning when the contest ends.

Passell, Peter, The law as the free market's rogue: Hostage to the Prisoner's Dilemma, *New York Times* (25 March 1994) (National Edition) B12.

Suppose that you're in a dispute over money. Data show that hiring a lawyer increases your expected payoff; similarly, hiring a lawyer increases your opponent's expected payoff. So you both hire lawyers, who then take a big chunk of the disputed amount, leaving less for both winner and loser of the dispute. The incentive for both to engage in noncooperative behavior is typical of the Prisoner's Dilemma, and some law school professors are beginning to interpret behavior of clients and lawyers in terms of that archetype of game theory.

NEWS AND LETTERS

ACKNOWLEDGEMENTS

In addition to our associate editors, the following have assisted the MAGAZINE *by refereeing papers during the past year. We appreciate the time and care they have given.*

Adams, William, *University of Maryland, College Park*
Alexander, James C., *University of Maryland, College Park*
Andrews, George E., *Pennsylvania State University, University Park*
Bailey, Donald F., *Trinity University*
Barbeau, Edward J., *University of Toronto*
Bassein, Susan, *Mills College*
Bauman, Steven, *University of Wisconsin, Madison*
Benjamin, Arthur, *Harvey Mudd College*
Benson, Steve, *Santa Clara University*
Bivens, Irl C., *Davidson College*
Boules, Raouf, *Towson State University*
Burton, David, *University of New Hampshire, Durham*
Casey, Stephen D., *The American University*
Crowe, Donald, *University of Wisconsin, Madison*
Cundy, H. Martyn, *Kendal, Cumbria, England*
Delany, James, *California Polytechnical State University*
Dodge, Clayton, *University of Maine, Orono*
Dunham, William W., *Muhlenberg College*
Faires, Barbara, *Iffley, Oxford, England*
Faires, J. Douglas, *Iffley, Oxford, England*
Fischer, Ismor, *University of Wisconsin, Madison*
Flanders, Harley, *University of Michigan, Ann Arbor*
Fredricks, Gregory, *Lewis and Clark College*
Gerstein, Larry, *University of California, Santa Barbara*
Ghahramani, Saeed, *Towson State University*
Gillman, Leonard, *University of Texas at Austin*
Goodson, Geoffrey, *Towson State University*
Grant, John, *Towson State University*
Green, Judy, *Marymount University*
Grimaldi, Ralph P., *Clemson University*
Growney, Joanne, *Bloomsburg University*
Grünbaum, Branko, *University of Washington*
Guy, Richard, *University of Calgary*
Hager, Anthony, *Wesleyan University, CT*
Hall, Leon, *University of Missouri at Rolla*
Hatcher, Rhonda, *Texas Christian University*
Hilton, Peter, *SUNY at Binghamton*
Hodges, Laurent, *Iowa State University*
Hoehn, Larry, *Austin Peay State University*
Horak, Martin, *Towson State University*
Horn, Roger A., *The University of Utah*
Jones, Dan, *Towson State University*
Jones, Sharon, *Towson State University*
Katz, Victor J., *University of the District of Columbia*
Kimberling, Clark H., *University of Evansville*
Klamkin, Murray S., *University of Alberta*
Kung, Sidney, *Jacksonville University*
LaDuke, Jeanne, *DePaul University*
Larson, Suzanne, *Loyola Marymount University*
Lercher, Bruce, *SUNY at Binghamton*
Lindstrom, Peter, *North Lake College*
Liu, Andrew, *University of Alberta, Edmonton*
Markham, Elizabeth, *St. Joseph's College*
Maurer, Stephen B., *Swarthmore College*
McCleary, John, *Vassar College*
Michaels, John G., *SUNY, College at Brockport*
Millett, Kenneth C., *University of California, Santa Barbara*
Monroe, Laura, *The University of Illinois at Chicago*
Morris, Brent, *Columbia, MD*
Nelsen, Roger B., *Lewis & Clark College*
Nunemacher, Jeffrey L., *Ohio Wesleyan University*
Nutt, Michael D., *Acadia University*
Page, Warren, *New York City Technical College*
Pedersen, Jean, *Santa Clara University*
Pfiefer, Richard, *San Jose State University*
Plank, Donald, *Stockton State College*
Richter, Bruce, *Carleton University*
Rosenstein, George, M., *Franklin and Marshall College*
Ross, Sharon C., *DeKalb Community College, GA*
Rudin, Mary Ellen, *University of Wisconsin, Madison*
Schattschneider, Doris M., *Moravian College*
Schaumberger, Norman, *Hofstra University*
Scheinerman, Ed, *The Johns Hopkins University*
Schulz, William, *Northern Arizona University*
Segal, Sanford, *University of Rochester*
Senechal, Lester, *Mount Holyoke College*
Sherman, Gary, *Rose-Hulman Institute of Technology*
Smith, David A., *Duke University*
Sohrab, Houshang, *Towson State University*
Sokolowsky, Daniel, *Williamsburg, VA*
Stein, Sherman K., *University of California, Davis*
Stern, Leonid, *Towson State University*
Stone, Arthur H., *Northeastern University*
Tucker, Alan, *SUNY at Stony Brook*
Tucker, Thomas W., *Colgate University*
Wagon, Stan, *Macalester College*

Wardlaw, William, *U.S. Naval Academy*
Watkins, William, *California State University, Northridge*
Weiss, Joan P., *Fairfield University*
Wetzel, John E., *University of Illinois, Urbana*
Wilf, Herbert, *University of Pennsylvania*
Zimmerman, Jay, *Towson State University*
Zorn, Paul, *St. Olaf College*

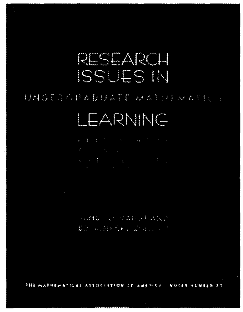
INDEX TO VOLUME 67

AUTHORS

Research Issues in Undergraduate Mathematics Learning

Preliminary Analyses and Reports

James J. Kaput and Ed Dubinsky, Editors



Research in undergraduate mathematics education is important for all college and university mathematicians. If our students are to be more successful in understanding mathematics, then college faculty need to understand how mathematics is learned. This knowledge can guide us in curriculum reform and in improving our own teaching. It can help us make mathematics accessible to all students and it can increase the number of graduate students in mathematics.

This volume of research in undergraduate mathematics education informs us about the nature of student learning in some of the most important topics in the undergraduate curriculum: sets, functions, calculus, statistics, abstract algebra and problem solving. Paying careful attention to the trouble students have in learning mathematics will help us to work with students so they can deal with those difficulties.

A survey of the literature begins the volume. Becker and Pence have brought together an unusually complete list of references on research in collegiate mathematics. Their comments will guide those attempting to begin or to continue a program of research in student learning.

The sad fact that even good calculus students stumble over nonroutine problems is the theme of Selden, Selden, and Mason. Their conclusions point to significant shortcomings in the curriculum. This study of student difficulties is continued by Ferrini-Mundy and Graham who investigate a single student's interactions with the fundamental concepts of the calculus. Baxter studies a group of students to learn how they acquire the concept of set, while Cuoco does the same for the concept of function.

Cooperative learning does help the student. That is the conclusion of Bonsangue, who investigates how two carefully matched classes of students in a statistics course perform on exams. How students learn to write proofs in group theory is the subject considered by Hart. Rosamond breaks new ground by comparing how emotions vary in their effect on the problem solving ability of novices and experts.

All college faculty should read this book to find how they can help their students learn mathematics.

150 pp., Paperbound, 1994
ISBN 0-88385-090-7
List: $24.00
Catalog Number NTE-33

ORDER FROM:

The Mathematical Association of America
1529 Eighteenth Street, NW
Washington, DC 20036
1-(800) 331-1622 Fax (202) 265-2384

Membership Code ______

Qty. Catalog Number Price

Total $________

Name ________________

Address ________________

City ________________

State ___ Zip Code ________

Payment ❑ Check ❑ VISA ❑ MASTERCARD

Credit Card No. ________________

Signature ________________

Exp. Date ________________

CONTENTS

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