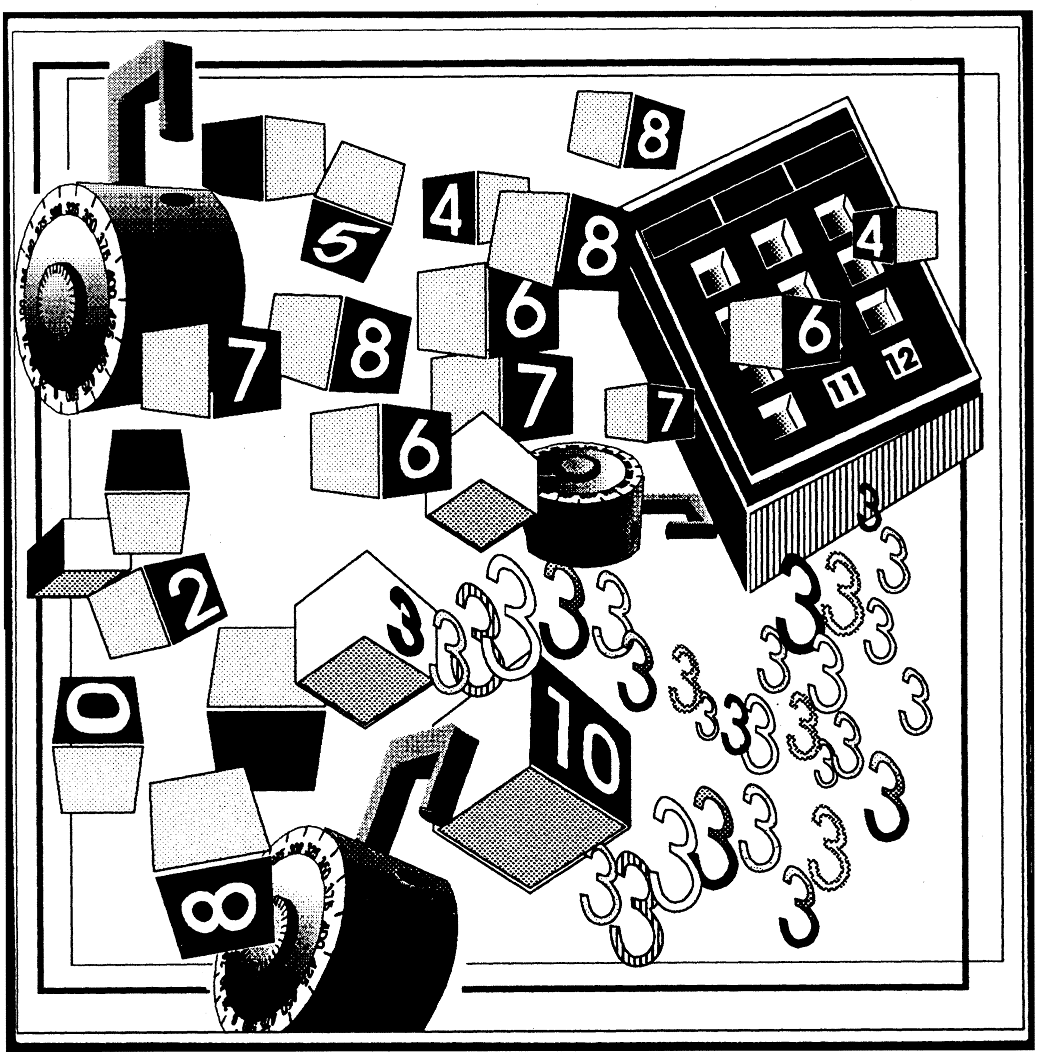
Vol. 68, No. 4 October 1995

MATHEMATICS MAGAZINE



- Permutations and Combination Locks
- Ceva, Menelaus, and the Area Principle
- Lewis Carroll and the Enumeration of Minimal Covers

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.**

AUTHORS

Dan Velleman received his bachelor's degree from Dartmouth College in 1976 and his doctorate from the University of Wisconsin in 1980. He taught at the University of Texas and the University of Toronto before joining the faculty of Amherst College in 1983. He is interested in logic, philosophy of mathematics, and the foundations of quantum mechanics. He is author of the book *How to Prove It,* and a coauthor, with Joe Konhauser and Stan Wagon, of the forthcoming book *Which Way Did the Bicycle Go?*, a collection of problems from the problem-of-the-week series begun by Joe and continued by Stan.

A native of Hanover, NH, **Greg Call** completed his A.B. degree at Dartmouth College in 1980. He did his graduate work at Harvard University under John Tate, receiving his Ph.D. in 1986. He taught for two years at Tufts University before going to Amherst College in 1988. While his primary research interests are Diophantine geometry and algebraic number theory, he is always ready to collaborate with Dan on an interesting problem-of-the-week.

Branko Grünbaum received his Ph.D. from Hebrew University in Jerusalem in 1958. Since 1966, he has taught at the University of Washington. **G. C. Shephard** received his Ph.D. from Cambridge University in 1951, and a D.Sc. from the University of Birmingham in 1967. He is Professor Emeritus at the University of East Anglia.

For more than 30 years, Grünbaum and Shephard have collaborated on various aspects of combinatorial geometry and related topics, and have written more than 50 joint papers. They find elementary and intuitive geometry still a fertile root of interesting and timely mathematics.

Anthony J. Macula received a B.S. degree from SUNY Plattsburgh in 1983 and a Ph.D. from Wesleyan University in 1989. Trained primarily as a point-set topologist, he made the switch to discrete mathematics because it is better suited to undergraduate research —a pedagogical subject in which the author is keenly interested. As this paper represents an initial excursion from topology to combinatorics, it is not surprising that the topic is about "covers." Presently investigating the connections between separation properties and combinatorial search problems, it seems that topological notions are still near and dear to his heart.

Vol. 68, No. 4 October 1995

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

Permutations and Combination Locks

DANIEL J. VELLEMAN
GREGORY S. CALL
Amherst College
Amherst, MA 01002

Consider a combination lock with n buttons, numbered 1 through n. A valid combination consists of a sequence of button-pushes, in which each button is pushed exactly once. If the buttons must be pushed one at a time, then clearly there will be $n!$ possible combinations. But what if we are allowed to push buttons simultaneously?

We can represent a valid combination for such a lock as a sequence of disjoint, nonempty subsets of the set $B = \{1, 2, \dots, n\}$ whose union is B. Each set in the sequence specifies a collection of buttons to be pushed simultaneously. For example, if $n = 3$ then we have the following possible combinations:

$$(\{1\},\{2\},\{3\}),(\{1\},\{3\},\{2\}),(\{2\},\{1\},\{3\}),(\{2\},\{3\},\{1\}),$$
$$(\{3\},\{1\},\{2\}),(\{3\},\{2\},\{1\}),(\{1,2\},\{3\}),(\{1,3\},\{2\}),(\{2,3\},\{1\}),$$
$$(\{1\},\{2,3\}),(\{2\},\{1,3\}),(\{3\},\{1,2\}),(\{1,2,3\}).$$

In this paper we will study formulas for the number of valid combinations for such a lock.

One of the most captivating features of this problem is that it may be solved by a number of elementary methods. In particular, we will use recurrence relations and generating functions to derive a variety of solutions ranging from the values of certain derivatives, to an infinite series, to some double summations with natural combinatorial interpretations. To obtain an asymptotic formula, we will use a basic integral estimate for our infinite series solution. It was more than a coincidence that, while we were developing these results, both of us were teaching second-semester calculus. Indeed we have endeavored to keep our calculus students in mind as we were writing.

A second striking feature of the lock combination problem is that it leads naturally to some well-known integer sequences. The reason we have included "Permutations" in the title is revealed by the appearance in Section 2 of the Eulerian numbers, which count the cardinalities in a natural partitioning of the set of permutations of n objects. We will also observe a connection between our lock problem and the Stirling numbers of the second kind.

We note that a number of authors have studied similar problems (Cayley **[3]**, Good **[7]**, and Gross **[8]**), though it appears that only Borenius, Danielson, and Jansson **[1]** have observed the connection with the Eulerian numbers. While most of our formulas can be found in these sources, we hope that readers will share our delight in collecting them and providing them with elementary derivations. We would like to thank Stan Wagon for sparking our interest in this problem and informing us of a commercially available door lock of this type with five buttons.

1. Two Solutions and an Asymptotic Formula

Let a_n be the number of combinations for a lock with n buttons. If $n = 0$ then the only valid combination is the empty sequence, so $a_0 = 1$. For $n > 0$, a valid combination will consist of a collection of k buttons that are pushed simultaneously, for some $1 \le k \le n$, followed by a combination using the remaining $n - k$ buttons. Thus we are led to the following recurrence relation for a_n:

$$a_0 = 1, \qquad a_n = \binom{n}{1}a_{n-1} + \binom{n}{2}a_{n-2} + \cdots + \binom{n}{n}a_0 \quad \text{for } n > 0.$$

Using this recurrence, we can compute the following values of a_n, for $n \le 10$:

n	0	1	2	3	4	5	6	7	8	9	10
a_n	1	1	3	13	75	541	4683	47293	545835	7087261	102247563

Note that the list of combinations above for the case $n = 3$ confirms the value $a_3 = 13$ in this table.

Filling in the formula for the binomial coefficients in the recurrence above, we find a common factor of $n!$:

$$a_n = n!\left(\frac{a_{n-1}}{1!(n-1)!} + \frac{a_{n-2}}{2!(n-2)!} + \cdots + \frac{a_0}{n!0!}\right).$$

This suggests defining $b_n = a_n/n!$. Dividing the formula above by $n!$ yields the following recurrence relation for b_n:

$$b_0 = 1, \qquad b_n = b_{n-1} + \frac{b_{n-2}}{2!} + \cdots + \frac{b_0}{n!} \quad \text{for } n > 0.$$

This recurrence relation is slightly simpler than the one for a_n, which suggests that we might be able to find a formula for a_n by first solving the recurrence for b_n. We begin with some simple bounds on b_n.

THEOREM 1. *For all* n,

$$\frac{1}{2(\ln 2)^n} \le b_n \le \frac{1}{(\ln 2)^n}.$$

Proof. We proceed by induction on n. The inequalities in the theorem clearly hold when $n = 0$. Now suppose $n > 0$, and assume that the inequalities hold for all $n' < n$. Then

$$\begin{aligned} b_n &= b_{n-1} + \frac{b_{n-2}}{2!} + \cdots + \frac{b_0}{n!} \\ &\le \frac{1}{(\ln 2)^{n-1}} + \frac{1}{2!(\ln 2)^{n-2}} + \cdots + \frac{1}{n!} \\ &= \frac{1}{(\ln 2)^n}\left(\ln 2 + \frac{(\ln 2)^2}{2!} + \cdots + \frac{(\ln 2)^n}{n!}\right) \\ &\le \frac{1}{(\ln 2)^n}(e^{\ln 2} - 1) = \frac{1}{(\ln 2)^n}. \end{aligned}$$

This gives the desired upper bound on b_n. For the lower bound, we begin by applying the induction hypothesis to all terms in the formula for b_n *except* the last:

$$\begin{aligned}b_n &= b_{n-1} + \frac{b_{n-2}}{2!} + \cdots + \frac{b_1}{(n-1)!} + \frac{b_0}{n!}\\ &\geq \frac{1}{2(\ln 2)^{n-1}} + \frac{1}{2!\cdot 2(\ln 2)^{n-2}} + \cdots + \frac{1}{(n-1)!\cdot 2\ln 2} + \frac{1}{n!}\\ &= \frac{1}{2(\ln 2)^n}\left(\ln 2 + \frac{(\ln 2)^2}{2!} + \cdots + \frac{(\ln 2)^{n-1}}{(n-1)!} + \frac{2(\ln 2)^n}{n!}\right).\end{aligned}$$

By Taylor's Theorem, there is a number c such that $0 < c < \ln 2$ and

$$\begin{aligned}e^{\ln 2} &= 1 + \ln 2 + \frac{(\ln 2)^2}{2!} + \cdots + \frac{(\ln 2)^{n-1}}{(n-1)!} + \frac{e^c(\ln 2)^n}{n!}\\ &\leq 1 + \ln 2 + \frac{(\ln 2)^2}{2!} + \cdots + \frac{(\ln 2)^{n-1}}{(n-1)!} + \frac{2(\ln 2)^n}{n!}.\end{aligned}$$

Thus

$$\begin{aligned}b_n &\geq \frac{1}{2(\ln 2)^n}\left(\ln 2 + \frac{(\ln 2)^2}{2!} + \cdots + \frac{(\ln 2)^{n-1}}{(n-1)!} + \frac{2(\ln 2)^n}{n!}\right)\\ &\geq \frac{1}{2(\ln 2)^n}(e^{\ln 2} - 1) = \frac{1}{2(\ln 2)^n},\end{aligned}$$

as required.

A natural way to study the sequence $\{b_n\}$ is to define the generating function

$$f(x) = \sum_{n=0}^{\infty} b_n x^n.$$

Note that by Theorem 1, the sum converges absolutely for $|x| < \ln 2$. Using our recurrence relation for b_n, we can solve for $f(x)$:

$$\begin{aligned}f(x) &= b_0 + \sum_{n=1}^{\infty} b_n x^n = 1 + \sum_{n=1}^{\infty}\sum_{k=1}^{n} \frac{b_{n-k}}{k!}x^n = 1 + \sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\frac{b_{n-k}}{k!}x^n\\ &= 1 + \sum_{k=1}^{\infty}\frac{x^k}{k!}\sum_{n=k}^{\infty} b_{n-k}x^{n-k} = 1 + \sum_{k=1}^{\infty}\frac{x^k}{k!}\sum_{n=0}^{\infty} b_n x^n = 1 + (e^x - 1)f(x),\end{aligned}$$

and therefore

$$f(x) = \frac{1}{2 - e^x}, \qquad |x| < \ln 2.$$

Since b_n is the coefficient of x^n in the Maclaurin series for $f(x)$, we have $b_n = f^{(n)}(0)/n!$. But recall that $a_n = n! b_n$, so we have proven the following theorem:

THEOREM 2. *For all* n,

$$a_n = \frac{d^n}{dx^n}\left(\frac{1}{2 - e^x}\right)\bigg|_{x=0}.$$

To compute values of a_n using Theorem 2, we must compute derivatives of the function $1/(2 - e^x)$. One way to make this computation easier is to rewrite the function as a geometric series:

$$\frac{1}{2 - e^x} = \frac{1/2}{1 - e^x/2} = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{e^x}{2}\right)^k, \qquad x < \ln 2.$$

Differentiating term-by-term we find that

$$\frac{d^n}{dx^n}\left(\frac{1}{2-e^x}\right)=\frac{1}{2}\sum_{k=0}^{\infty}k^n\left(\frac{e^x}{2}\right)^k, \qquad x<\ln 2.$$

(Note that for $n=0$ we use the convention $0^0=1$.) Thus, applying Theorem 2 we get another formula for a_n:

THEOREM 3. *For all* n,

$$a_n=\frac{1}{2}\sum_{k=0}^{\infty}\frac{k^n}{2^k}.$$

Using Theorem 3, we can improve on our estimates in Theorem 1. A natural way to estimate the infinite sum in Theorem 3 is to use the improper integral

$$\int_0^{\infty}\frac{x^n}{2^x}dx.$$

We evaluate this integral by performing the substitution $u=x\ln 2$ and then integrating by parts repeatedly (or recognizing the Γ function):

$$\int_0^{\infty}\frac{x^n}{2^x}dx=\frac{1}{(\ln 2)^{n+1}}\int_0^{\infty}\frac{u^n}{e^u}du=\frac{\Gamma(n+1)}{(\ln 2)^{n+1}}=\frac{n!}{(\ln 2)^{n+1}}.$$

Thus, by Theorem 3 we expect to have

$$a_n\approx\frac{n!}{2(\ln 2)^{n+1}}.$$

For small n, this approximation is remarkably accurate. For example, we have:

$$a_{15}=230283190977853,$$
$$\frac{15!}{2(\ln 2)^{16}}\approx 230283190977853.04.$$

Unfortunately, for larger n the error of the approximation grows, but we can place bounds on this error.

THEOREM 4. *For all* n,

$$\frac{n!}{2(\ln 2)^{n+1}}-\frac{1}{2}\left(\frac{n}{e\ln 2}\right)^n<a_n<\frac{n!}{2(\ln 2)^{n+1}}+\frac{1}{2}\left(\frac{n}{e\ln 2}\right)^n.$$

Proof. It is easy to check that the function $g(x)=x^n/2^x$ is increasing on the interval $[0,n/\ln 2]$ and decreasing on $[n/\ln 2,\infty)$, and thus its maximum value on the interval $[0,\infty)$ is

$$g(n/\ln 2)=\frac{(n/\ln 2)^n}{2^{n/\ln 2}}=\left(\frac{n}{e\ln 2}\right)^n.$$

Let j be the largest integer less than or equal to $n/\ln 2$ and put $r=(n/\ln 2)-j$. Thus $(n/\ln 2)=j+r$, j is a nonnegative integer, and $0\le r<1$.

We now use upper and lower rectangles to overestimate and underestimate the integral of $g(x)$. Upper rectangles of width 1 lead to the following overestimate of the integral (see FIGURE 1):

$$\int_0^\infty \frac{x^n}{2^x}dx < \sum_{k=1}^{j} \frac{k^n}{2^k} + \left(\frac{n}{e\ln 2}\right)^n + \sum_{k=j+1}^{\infty} \frac{k^n}{2^k} \le 2a_n + \left(\frac{n}{e\ln 2}\right)^n.$$

Substituting in the value of the integral gives the required lower bound on a_n.

To underestimate with lower rectangles we again use rectangles of width 1, except that we split the interval from j to $j+1$ into two rectangles, of width r and $1-r$ (see FIGURE 2). This gives us

$$\begin{aligned}\int_0^\infty \frac{x^n}{2^x}dx &> \sum_{k=0}^{j-1} \frac{k^n}{2^k} + r\frac{j^n}{2^j} + (1-r)\frac{(j+1)^n}{2^{j+1}} + \sum_{k=j+2}^{\infty} \frac{k^n}{2^k}\\ &= 2a_n - \left((1-r)\frac{j^n}{2^j} + r\frac{(j+1)^n}{2^{j+1}}\right)\\ &\le 2a_n - \left(\frac{n}{e\ln 2}\right)^n,\end{aligned}$$

which leads to the stated upper bound on a_n.

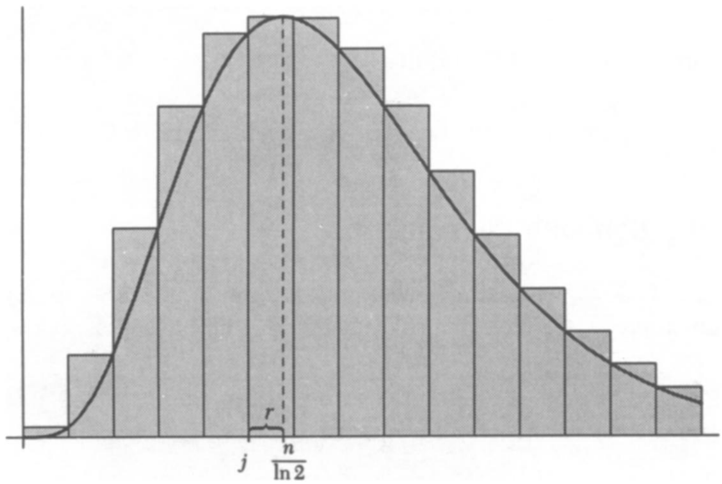


FIGURE 1

Approximating $\int_0^\infty \frac{x^n}{2^x}dx$ with upper rectangles.

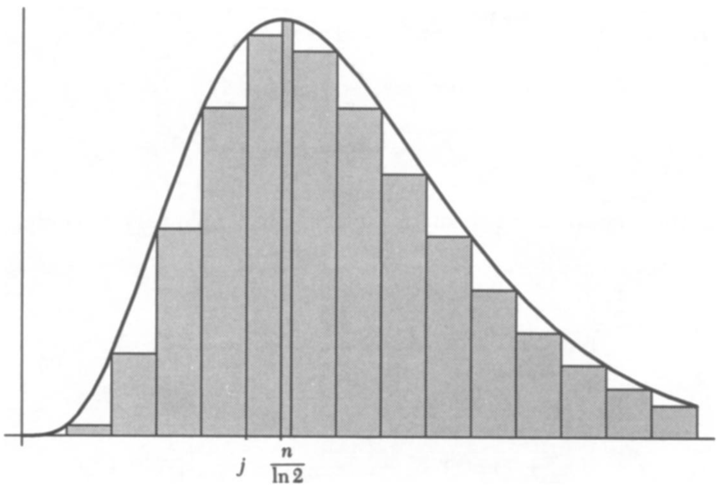


FIGURE 2

Approximating $\int_0^\infty \frac{x^n}{2^x}dx$ with lower rectangles.

Better bounds on a_n can be obtained by using methods from complex analysis (see [1], [7], and [8]). However, our bounds are good enough to prove:

COROLLARY 5.

$$\lim_{n\to\infty} \frac{a_n}{n!/\left[2(\ln 2)^{n+1}\right]} = 1.$$

Proof. By Theorem 4, we have

$$\left| \frac{a_n}{n!/\left[2(\ln 2)^{n+1}\right]} - 1 \right| < \frac{\ln 2 (n/e)^n}{n!}.$$

To complete the proof of the corollary, we show that the quantity on the right side of this inequality approaches 0 as n approaches infinity. Recall Stirling's formula, which says that

$$\lim_{n\to\infty} \frac{\sqrt{2\pi n}\,(n/e)^n}{n!} = 1.$$

Applying Stirling's formula, we find that

$$\lim_{n\to\infty} \frac{\ln 2 (n/e)^n}{n!} = \lim_{n\to\infty} \frac{\ln 2}{\sqrt{2\pi n}} \frac{\sqrt{2\pi n}\,(n/e)^n}{n!} = 0\cdot 1 = 0.$$

2. A Connection with Permutations

So far we have been concentrating on estimates of the sum in Theorem 3, in order to get approximations to a_n. We can also evaluate the sum exactly, as follows. For $n \geq 0$ let

$$h_n(x) = \sum_{k=0}^{\infty} k^n x^k.$$

Then by Theorem 3, $a_n = (1/2)h_n(1/2)$. We now derive a formula for a_n by finding formulas for the $h_n(x)$'s.

Clearly

$$h_0(x) = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \qquad -1 < x < 1.$$

Also, differentiating term-by-term we see that

$$h_n'(x) = \sum_{k=0}^{\infty} k^{n+1} x^{k-1}, \qquad \text{so} \quad x\,h_n'(x) = \sum_{k=0}^{\infty} k^{n+1} x^k = h_{n+1}(x).$$

Applying this recurrence repeatedly, we get the following formulas for $h_n(x)$, for $n \leq 5$:

$$h_0(x) = \frac{1}{1-x}, \qquad h_1(x) = \frac{x}{(1-x)^2},$$

$$h_2(x) = \frac{x+x^2}{(1-x)^3}, \qquad h_3(x) = \frac{x+4x^2+x^3}{(1-x)^4},$$

$$h_4(x) = \frac{x+11x^2+11x^3+x^4}{(1-x)^5}, \qquad h_5(x) = \frac{x+26x^2+66x^3+26x^4+x^5}{(1-x)^6}.$$

It appears that $h_n(x)$ is always a polynomial of degree n divided by $(1-x)^{n+1}$, but it is not immediately clear what the pattern of coefficients in these polynomials is. To study this pattern, we will introduce the following notation. Let $A_{n,k}$ be the coefficient of x^k in the numerator of $h_n(x)$. For example, $A_{5,4}=26$. Then the recurrence for $h_n(x)$ above can be used to derive a recurrence for the $A_{n,k}$'s:

THEOREM 6. *For all* $n \geq 1$,

$$h_n(x) = \frac{\sum_{k=1}^{n} A_{n,k} x^k}{(1-x)^{n+1}},$$

where the numbers $A_{n,k}$ *are given by the following recurrence relation*:

$$A_{n,1} = A_{n,n} = 1, \qquad A_{n+1,k} = kA_{n,k} + (n+2-k)A_{n,k-1} \quad \textit{for } 2 \leq k \leq n.$$

Proof. We proceed by induction. The case $n=1$ is easy to verify. For the induction step, suppose the formula in the statement of the theorem is correct for h_n. Then

$$\begin{aligned} h_{n+1}(x) &= xh_n'(x) = x\frac{d}{dx}\left(\frac{\sum_{k=1}^{n} A_{n,k}x^k}{(1-x)^{n+1}}\right) \\ &= x\frac{(1-x)^{n+1}\sum_{k=1}^{n} kA_{n,k}x^{k-1} + (n+1)(1-x)^n\sum_{k=1}^{n} A_{n,k}x^k}{(1-x)^{2n+2}} \\ &= \frac{\sum_{k=1}^{n} kA_{n,k}(1-x)x^k + \sum_{k=1}^{n}(n+1)A_{n,k}x^{k+1}}{(1-x)^{n+2}} \\ &= \frac{\sum_{k=1}^{n} kA_{n,k}x^k + \sum_{k=1}^{n}(n+1-k)A_{n,k}x^{k+1}}{(1-x)^{n+2}} \\ &= \frac{x + \sum_{k=2}^{n}(kA_{n,k} + (n+2-k)A_{n,k-1})x^k + x^{n+1}}{(1-x)^{n+2}} \\ &= \frac{\sum_{k=1}^{n+1} A_{n+1,k}x^k}{(1-x)^{n+2}}. \end{aligned}$$

The recurrence relation given in Theorem 6 allows us to readily compute the coefficients $A_{n,k}$ for $1 \leq k \leq n$, given the coefficients $A_{n-1,k}$ for $1 \leq k \leq n-1$. Thus it is convenient to display these numbers in the form of a triangular table à la Pascal's triangle. The nth row of the triangle will be

$$A_{n,1} \quad A_{n,2} \quad \cdots \quad A_{n,n}.$$

These rows may be computed by beginning with $A_{1,1}=1$ and generating each row from the preceding row via the recurrence relation of Theorem 6. The first seven rows of the triangle are shown in FIGURE 3.

1						
1	1					
1	4	1				
1	11	11	1			
1	26	66	26	1		
1	57	302	302	57	1	
1	120	1191	2416	1191	120	1

FIGURE 3

Note that the first five rows agree with our computations of $h_n(x)$ above. Like Pascal's triangle, this triangle appears to be symmetric, which suggests that perhaps $A_{n,k} = A_{n,n+1-k}$ for all $1 \leq k \leq n$. An even more striking analogy with Pascal's triangle is revealed by adding the entries across the rows of our triangle. Recall that row n of Pascal's triangle adds up to 2^n, because the entries in this row count the numbers of subsets of $\{1, 2, \dots, n\}$ of different sizes. Adding the entries across the rows of the triangle in FIGURE 3, we find that the sums of the rows are: $1, 2, 6, 24, 120, 720, 5040$; i.e., the factorials! This observation suggests that we look for an interpretation of the numbers $A_{n,k}$, for $1 \leq k \leq n$, in terms of a partition of the set of permutations of $\{1, 2, \dots, n\}$.

For ease of terminology, we will use the term *n-permutation* to denote a permutation of $\{1, 2, \dots, n\}$. We think of an n-permutation as an ordered list of the numbers $1, 2, \dots, n$. Given a permutation in this form, say $s_1 s_2 \dots s_n$, we count the number of increasing runs in the sequence $s_1, s_2, \dots, s_n$ reading from left to right. For example, the permutation 1 4 2 5 3 6 8 7 has four increasing runs; namely, 1 4, 2 5, 3 6 8, and 7. As the next proposition shows, partitioning the set of n-permutations according to their numbers of increasing runs yields the numbers $A_{n,k}$.

PROPOSITION 7. *For all n and k such that $1 \leq k \leq n$, the number of n-permutations with k increasing runs is $A_{n,k}$.*

Proof. See [**4**, Theorem A, pp. 241–2] or [**16**, Problem 12.22(a)].

The numbers $A_{n,k}$, and the interpretation given for them in Proposition 7, are well known in combinatorics (see [**4**], [**13**], [**14**], [**15**], and [**16**]). They are called the *Eulerian numbers* (see [**5**]), and so it is appropriate to refer to the triangular table of numbers we defined above as *Euler's triangle*. MacMahon [**11**] was apparently the first to use Eulerian numbers to classify permutations by increasing runs. The Eulerian numbers have appeared in a variety of contexts in combinatorics and statistics (e.g., see [**2**], [**9**], and [**10**]). A very extensive description of their theory is given by Foata and Schützenberg [**6**].

We can now confirm the conjectures we made based on the first seven rows of Euler's triangle.

COROLLARY 8. *For all $n \geq 1$,*

(a) $\sum_{k=1}^{n} A_{n,k} = n!$ *and*

(b) $A_{n,k} = A_{n,n+1-k}$ *for* $1 \leq k \leq n$.

Proof. (a) is immediate from Proposition 7. For a proof of (b), see [**4**, Theorem B, p. 242] or [**16**, Problem 12.22(b)].

Now we return to our original lock problem. From Theorems 3 and 6, we see that the number of combinations for a lock with $n \geq 1$ buttons, a_n, satisfies:

$$a_n = \frac{1}{2} h_n\left(\frac{1}{2}\right) = \frac{1}{2} \sum_{k=1}^{n} A_{n,k}\left(\frac{1}{2}\right)^{k-n-1} = \sum_{k=1}^{n} A_{n,k} 2^{n-k}. \tag{1}$$

By the symmetry of Euler's triangle stated in Corollary 8(b), we also have:

$$a_n = \sum_{k=1}^{n} A_{n,n+1-k}\, 2^{(n+1-k)-1} = \sum_{k=1}^{n} A_{n,k} 2^{k-1}. \tag{2}$$

Equation (1) has a natural interpretation in terms of lock combinations. To see this, first note that given a lock combination, we can generate a corresponding n-permutation by writing down the numbers of the buttons pressed, in the order in which they are pressed, with the numbers of the buttons pressed simultaneously being written in increasing order. For example, the lock combination $(\{1\},\{3,4\},\{2,5\})$ would correspond to the permutation 1 3 4 2 5. Clearly a lock combination that consists of l steps will lead to a permutation with at most l increasing runs. Furthermore, if we follow this procedure for every lock combination, each permutation with k increasing runs will appear exactly 2^{n-k} times. To demonstrate this, let σ be an n-permutation with k increasing runs, and consider separating this permutation into blocks representing the steps in a corresponding lock combination; for example, we might do this by inserting vertical lines in the permutation to delimit the blocks. The numbers within each block must be in increasing order, so we must at least insert the $k-1$ vertical lines needed to delimit the k increasing runs in σ. There are $n-1$ spaces between the n numbers where vertical lines might go, and $k-1$ of them are now filled, so there are $(n-1)-(k-1)=n-k$ places left unfilled. By choosing a subset of these $n-k$ positions and inserting vertical lines at the chosen positions, we determine a lock combination corresponding to σ. Hence, there are 2^{n-k} lock combinations corresponding to a single n-permutation with k increasing runs.

For example, the permutation 1 3 4 2 5 has two increasing runs: 1 3 4 and 2 5. It corresponds to the $2^{5-2}=8$ lock combinations:

$$(\{1,3,4\},\{2,5\}),(\{1\},\{3,4\},\{2,5\}),(\{1,3\},\{4\},\{2,5\}),(\{1\},\{3\},\{4\},\{2,5\}),$$
$$(\{1,3,4\},\{2\},\{5\}),(\{1\},\{3,4\},\{2\},\{5\}),$$
$$(\{1,3\},\{4\},\{2\},\{5\}),(\{1\},\{3\},\{4\},\{2\},\{5\}).$$

Proceeding by analogy with Pascal's triangle and, in particular, in view of the recurrence relation of Theorem 6, one might hope for an explicit formula giving the numbers $A_{n,k}$ in terms of binomial coefficients. In fact, such a formula is known.

THEOREM 9. *For* $1 \le k \le n$,

$$A_{n,k}=\sum_{i=0}^{k-1}(-1)^i\binom{n+1}{i}(k-i)^n.$$

Proof. See [4, Theorem C, p. 243], [**16**, Problem 12.22(d)], or [**12**, Theorem 2].

Combining equations (1) and (2) with Theorem 9, we obtain two more formulas for a_n.

THEOREM 10. *For all* $n \ge 1$,

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

Our last solution to the combination lock problem involves another well-known family of numbers, the *Stirling numbers of the second kind*. For $1 \le k \le n$, the Stirling number $S(n,k)$ is the number of unordered partitions of the set $\{1,2,\ldots,n\}$ into k nonempty subsets. Since a lock combination with k steps can be thought of as an *ordered* partition of this set into k nonempty subsets, we see that the number of lock combinations with k steps is equal to $k!S(n,k)$. Summing over all k gives us the following formula:

THEOREM 11. *For all* $n \geq 1$,

$$a_n = \sum_{k=1}^{n} k!S(n,k).$$

Like the Eulerian numbers, the Stirling numbers of the second kind can also be expressed in terms of binomial coefficients. In fact, there is a striking resemblance between the formula for $S(n,k)$ given in our next theorem and the formula for $A_{n,k}$ given in Theorem 9.

THEOREM 12. *For all* $1 \leq k \leq n$,

$$S(n,k) = \frac{1}{k!} \sum_{i=0}^{k-1} (-1)^i \binom{k}{i} (k-i)^n.$$

Proof. See [4, Theorem A, pp. 204–205].

Combining Theorems 11 and 12, we obtain our last formula for a_n.

THEOREM 13. *For all* $n \geq 1$,

$$a_n = \sum_{k=1}^{n} \sum_{i=0}^{k-1} (-1)^i \binom{k}{i} (k-i)^n.$$

In conclusion, we note that we can now compute the number of valid combinations for the commercially available five-button combination door lock mentioned in the introduction. This lock has an extra feature; namely, it does not require that all five buttons be used in a valid combination. For example, ({2}), ({1, 5}, {3}), and ({1, 2, 4, 5}) are valid combinations. By viewing the subset of unused buttons as the last block in a combination that uses every button, we see that there is a one-to-one correspondence between the combinations that use all the buttons and those that don't. However, this correspondence includes the empty combination—the combination in which no buttons are pushed—and this combination is not meaningful for a commercial door lock. Thus, the number of meaningful combinations for the lock is $2a_5 - 1 = 1081$.

REFERENCES

1. G. Borenius, S. Danielson and B. Jansson, Number of ways for choosing an alternative. A combinatorial problem related to Eulerian numbers, *FOA Reports* 1, no. 5 (1967), 1–9.
2. L. Carlitz, Eulerian numbers and polynomials, this MAGAZINE, 32 (1959), 247–260.
3. A. Cayley, On the analytical forms called trees. Second Part. *Phil. Magazine*, 18 (1859), 374–378.
4. L. Comtet, *Advanced Combinatorics*, D. Reidel, Dordrecht-Holland, 1974.
5. L. Euler, *Institutiones calculi differentialis*, Vol. 2, St. Petersburg, 1755, pp. 485–487; see also L. Euler, *Opera Omnia*, Series Prima, Vol. 10 (1913), pp. 373–375.
6. D. Foata and M.-P. Schützenberger, *Théorie Géométrique des Polynômes Eulériens*, Springer Lecture Notes 138, Springer-Verlag, Berlin, 1970.
7. I. J. Good, The number of orderings of n candidates when ties are permitted, *Fibonacci Quarterly*, 13 (1975), 11–18.
8. O. A. Gross, Preferential arrangements, *Amer. Math. Monthly*, 69 (1962), 1–4.
9. I. Kaplansky and J. Riordan, The problem of rooks and its applications, *Duke Math. J.*, 13 (1946), 259–268.
10. A. C. Kimber, Eulerian numbers, in *Encyclopedia of Statistical Sciences (Supplement)*, eds. S. Kotz, N. L. Johnson, and C. B. Read, John Wiley & Sons, Inc., New York, 1989, pp. 59–60.
11. P. A. MacMahon, Second memoir on the composition of numbers, *Phil. Trans. Royal Soc. London*, A, 207 (1908), pp. 65–134.
12. D. Neal, The series $\sum_{n=1}^{\infty} n^m x^n$ and a Pascal-like triangle, *College Math J.* 25, no. 2 (1994), pp. 99–101.
13. J. Riordan, *An Introduction to Combinatorial Analysis*, John Wiley & Sons, Inc., New York, 1958.

14. N. J. A. Sloane, *A Handbook of Integer Sequences*, Academic Press, New York, 1973.
15. D. Stanton and D. White, *Constructive Combinatorics*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1986.
16. I. Tomescu, *Problems in Combinatorics and Graph Theory*, translated by R. A. Melter, John Wiley & Sons, Inc., New York, 1985.

A Mnemonic For e

ZEEV BAREL
Hendrix College
Conway, AR 72032-3080

Mnemonic devices for the decimal expansion of π are well known. For example, "May I have a large container of coffee" may be found in **[1]**. The principle is to replace each word by the number of letters in it. Similar mnemonics for π exist in other languages, including Russian and Greek.

On the other hand, the only mnemonics for e known to the author are those given in **[2]** (in English, French, and Spanish) covering 10 places after the decimal point and **[3]** containing a sentence going to 20 digits. Potential authors may have been discouraged by early appearances of zeros in

$$e = \mathbf{2.718281828459045235360287471352662497757}\ldots.$$

The 20-place mnemonic in **[3]** uses the (one-letter) word "O" to represent the zero at place 13.

In the quasi-poem below, zeros are represented by exclamations "...!" whose canonical pronunciation is a short "ah!" (although the readers may substitute their favorite letterless words).

We present a mnemonic
　　To memorize a constant
So exciting that Euler exclaimed: "...!"
　　When first it was found,
Yes, loudly: "...!"
　　My students perhaps
Will compute e,
　　Use power of Taylor series,
An easy summation formula,
　　Obvious, clear, elegant!

The author thanks the referee for making the author aware of **[2]** and **[3]** and for suggesting an improvement.

REFERENCES

1. Howard Eves, *An Introduction to the History of Mathematics*, 5th edition, Saunders College Pub., Philadelphia, 1983, p. 89.
2. Martin Gardner, *The Scientific American Book of Mathematical Puzzles and Diversions*, Simon and Schuster, New York, 1959, p. 103 and p. 109.
3. ______, *The Unexpected Hanging and Other Mathematical Diversions*, Simon and Schuster, New York, 1969, p. 40.

14. N. J. A. Sloane, *A Handbook of Integer Sequences*, Academic Press, New York, 1973.
15. D. Stanton and D. White, *Constructive Combinatorics*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1986.
16. I. Tomescu, *Problems in Combinatorics and Graph Theory*, translated by R. A. Melter, John Wiley & Sons, Inc., New York, 1985.

A Mnemonic For *e*

ZEEV BAREL
Hendrix College
Conway, AR 72032-3080

Mnemonic devices for the decimal expansion of π are well known. For example, "May I have a large container of coffee" may be found in **[1]**. The principle is to replace each word by the number of letters in it. Similar mnemonics for π exist in other languages, including Russian and Greek.

On the other hand, the only mnemonics for e known to the author are those given in **[2]** (in English, French, and Spanish) covering 10 places after the decimal point and **[3]** containing a sentence going to 20 digits. Potential authors may have been discouraged by early appearances of zeros in

$$e = \mathbf{2.71828182845904523536028747135266249775 7}\ldots.$$

The 20-place mnemonic in **[3]** uses the (one-letter) word "O" to represent the zero at place 13.

In the quasi-poem below, zeros are represented by exclamations "...!" whose canonical pronunciation is a short "ah!" (although the readers may substitute their favorite letterless words).

We present a mnemonic
    To memorize a constant
So exciting that Euler exclaimed: "...!"
    When first it was found,
Yes, loudly: "...!"
    My students perhaps
Will compute e,
    Use power of Taylor series,
An easy summation formula,
    Obvious, clear, elegant!

The author thanks the referee for making the author aware of **[2]** and **[3]** and for suggesting an improvement.

REFERENCES

1. Howard Eves, *An Introduction to the History of Mathematics*, 5th edition, Saunders College Pub., Philadelphia, 1983, p. 89.
2. Martin Gardner, *The Scientific American Book of Mathematical Puzzles and Diversions*, Simon and Schuster, New York, 1959, p. 103 and p. 109.
3. ______, *The Unexpected Hanging and Other Mathematical Diversions*, Simon and Schuster, New York, 1969, p. 40.

Ceva, Menelaus, and the Area Principle

BRANKO GRÜNBAUM[1]
University of Washington
Seattle, WA 98195

G. C. SHEPHARD
University of East Anglia
Norwich NR4 7TJ, United Kingdom

1. Introduction

The theorems of Ceva and Menelaus [**3**, p. 220], which will be stated shortly, are among the most attractive and useful results in elementary plane geometry. They are easy to state and are quite general in the sense that, for example, Menelaus' Theorem applies to *any* triangle and to *any* transversal that does not pass through a vertex. These theorems, and their proofs, are classical as is reflected in their names. The Greek Menelaus lived in the first century A.D. and the Italian Giovanni Ceva published his theorem (and rediscovered Menelaus' Theorem) in the 17th century. Recently Hoehn [**8**] obtained a new result of a similar kind, showing that the products of five quotients of certain lengths in a pentagram have the value 1.

The purpose of this note is to show that these and other results, and their extensions to general polygons with arbitrarily many sides, are the consequences of a simple idea which we shall call the *area principle*. This principle is illustrated in FIGURE 1, in which P is the point of intersection of the lines BC and A_1A_2. Denoting the lengths of the segments $[A_1, P]$ and $[A_2, P]$ by $|A_1P|$ and $|A_2P|$, and the areas of the triangles $[A_1, B, C]$ and $[A_2, B, C]$ by $|A_1BC|$ and $|A_2BC|$, respectively, the area principle states

$$\frac{|A_1P|}{|A_2P|} = \frac{|A_1BC|}{|A_2BC|}, \tag{1}$$

and this is true whenever the ratios are well-defined, that is, their denominators do not vanish. The validity of (1) does not depend on whether the points A_1 and A_2 are separated by the line BC (FIGURE 1(a)) or not (FIGURE 9(b)). Later we shall refine (1) by assigning signs to the ratios; just for the present we shall consider all areas and lengths as positive.

The proof of (1) is immediate. Clearly each side of this equation is equal to the ratio of the heights of the two triangles on base $[B, C]$. We note that, although some of the well-known proofs of Ceva's Theorem (see for example [**4**, p. 4]) use area arguments, these are distinct from the "area principle."

As a foretaste of the methods to be used, we shall use the area principle to prove three of the theorems mentioned above. Suppose a transversal cuts the lines BC, CA, AB, determined by the named pairs of vertices of the triangle $[A, B, C]$, in points P, Q, R, respectively, and that these three points are distinct from the vertices of the triangle (see FIGURE 2). Then Menelaus' Theorem states that

$$\frac{|BP|}{|PC|} \cdot \frac{|CQ|}{|QA|} \cdot \frac{|AR|}{|RB|} = 1. \tag{2}$$

[1]Research supported in part by NSF grant DMS-9300657.

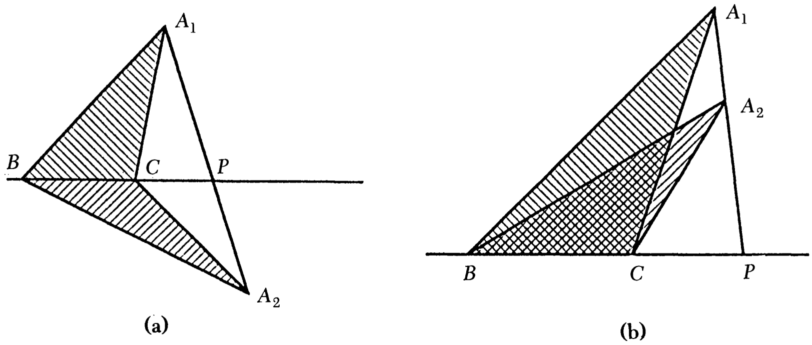


FIGURE 1
The area principle states that $|A_1P|/|A_2P| = |A_1BC|/|A_2BC|$.

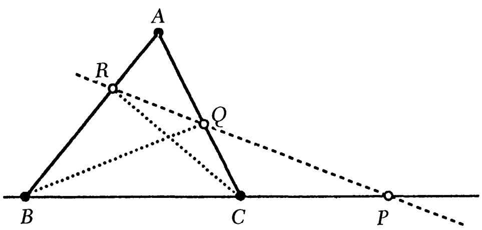


FIGURE 2
Menelaus' Theorem states that $(|BP|/|PC|)\cdot(|CQ|/|QA|)\cdot(|AR|/|RB|) = 1$.

To establish this we apply the area principle to the triangles with base $[R,Q]$, and we see that

$$\frac{|BP|}{|PC|} = \frac{|BRQ|}{|CRQ|}, \qquad \frac{|CQ|}{|QA|} = \frac{|CRQ|}{|ARQ|}, \qquad \frac{|AR|}{|RB|} = \frac{|ARQ|}{|BRQ|}.$$

Substituting these values in (2) the areas of the triangles all cancel, showing that (2) is valid. This proves the theorem.

Let O be a point such that the lines AO, BO, CO meet the opposite sides of the triangle $[A, B, C]$ in P, Q, R, respectively. We shall suppose that these three points are distinct from the vertices A, B, C (see FIGURE 3). Then Ceva's Theorem states that the relationship (2) holds in this case also. To establish this we consider triangles with base $[A, O]$ and then the area principle yields

$$\frac{|BP|}{|PC|} = \frac{|AOB|}{|COA|}.$$

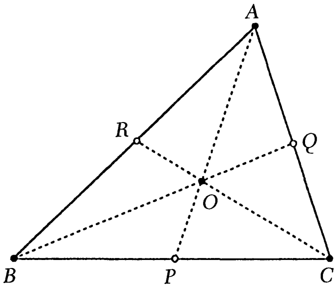


FIGURE 3
Ceva's Theorem states that $(|BP|/|PC|)\cdot(|CQ|/|QA|)\cdot(|AR|/|RB|) = 1$.

Similarly, using triangles with bases $[B, O]$ and $[C, O]$, we have

$$\frac{|CQ|}{|QA|} = \frac{|BOC|}{|AOB|}, \qquad \frac{|AR|}{|RB|} = \frac{|COA|}{|BOC|}.$$

Substituting these values in (2), the areas of the triangles all cancel to yield the value 1, which proves the theorem.

Finally, we consider Hoehn's Theorem, which states that for a pentagon, using the notation indicated in FIGURE 4,

$$\frac{|V_1W_1|}{|W_2V_3|}\cdot\frac{|V_2W_2|}{|W_3V_4|}\cdot\frac{|V_3W_3|}{|W_4V_5|}\cdot\frac{|V_4W_4|}{|W_5V_1|}\cdot\frac{|V_5W_5|}{|W_1V_2|} = 1 \tag{3}$$

and

$$\frac{|V_1W_2|}{|W_1V_3|}\cdot\frac{|V_2W_3|}{|W_2V_4|}\cdot\frac{|V_3W_4|}{|W_3V_5|}\cdot\frac{|V_4W_5|}{|W_4V_1|}\cdot\frac{|V_5W_1|}{|W_5V_2|} = 1. \tag{4}$$

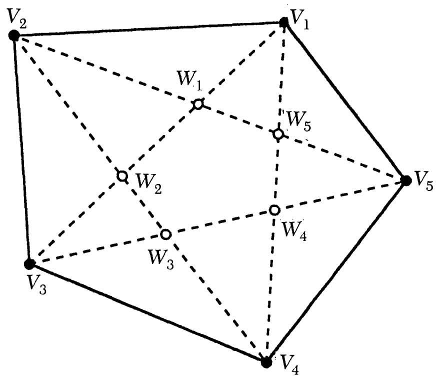


FIGURE 4
Hoehn's Theorem states that

$$\frac{|V_1W_1|}{|W_2V_3|}\cdot\frac{|V_2W_2|}{|W_3V_4|}\cdot\frac{|V_3W_3|}{|W_4V_5|}\cdot\frac{|V_4W_4|}{|W_5V_1|}\cdot\frac{|V_5W_5|}{|W_1V_2|} = 1$$

and

$$\frac{|V_1W_2|}{|W_1V_3|}\cdot\frac{|V_2W_3|}{|W_2V_4|}\cdot\frac{|V_3W_4|}{|W_3V_5|}\cdot\frac{|V_4W_5|}{|W_4V_1|}\cdot\frac{|V_5W_1|}{|W_5V_2|} = 1.$$

In Hoehn's original proof, Menelaus' Theorem was applied to various triangles and transversals in the diagram; here we give a more direct proof using the area principle. Consider the pentagon shown in FIGURE 4. (In this initial treatment we restrict attention to the case in which $[V_1, V_2, V_3, V_4, V_5]$ is a convex pentagon. The more general case, without any assumption of convexity, will be dealt with in Section 4.) Using triangles with base $[V_2, V_4]$, we have

$$\frac{|V_1W_2|}{|W_2V_3|} = \frac{|V_1V_2V_4|}{|V_3V_4V_2|}.$$

Hence $|V_1V_3|/|W_2V_3| = (|V_1W_2| + |W_2V_3|)/|W_2V_3| = (|V_1V_2V_4| + |V_3V_4V_2|)/|V_3V_4V_2| = |V_1V_2V_3V_4|/|V_3V_4V_2|$ where $|V_1V_2V_3V_4|$ is the area of the quadrilateral $[V_1, V_2, V_3, V_4]$. Similarly,

$$\frac{|V_1V_3|}{|V_1W_4|} = \frac{|V_1V_2V_3V_5|}{|V_1V_2V_5|}$$

and so

$$\frac{|V_1W_1|}{|W_2V_3|} = \frac{|V_1V_2V_5|}{|V_1V_2V_3V_5|}\cdot\frac{|V_1V_2V_3V_4|}{|V_3V_4V_2|}.$$

In general, for $i = 1, 2, \ldots, 5$,

$$\frac{|V_iW_i|}{|W_{i+1}V_{i+2}|} = \frac{|V_iV_{i+1}V_{i+4}|}{|V_iV_{i+1}V_{i+2}V_{i+4}|} \cdot \frac{|V_iV_{i+1}V_{i+2}V_{i+3}|}{|V_{i+2}V_{i+3}V_{i+1}|}. \tag{5}$$

Substituting these values in the left side of (3) we see that the areas of the triangles and quadrilaterals all cancel, yielding the value 1 as required. The second assertion (4) of Hoehn's Theorem can be proved in a similar manner.

The pattern of the above proofs will be applied repeatedly: We shall express ratios of lengths as ratios of areas, and then show that, in a product of such ratios, cancellation takes place to yield a constant value.

Before describing the extensions of these results, we put them in a more general context.

2. Affine Geometry and Polygons

Recall that affine geometry [**3**, Chapter 13] is concerned with geometric properties that are affine invariant, which means that they are invariant under *affinities* (that is, non-singular linear transformations combined with translations). Geometrically, such transformations can be thought of as rotations, reflections, translations and shears, or any combination of these. Incidences, ratios of lengths on parallel lines and ratios of areas are preserved under affinities and hence belong to affine geometry; in contrast, lengths, angles and areas do not. All our results belong to affine geometry, but clearly remain valid in the more restrictive Euclidean geometry.

By a *polygon* $P = [V_1, V_2, \ldots, V_n]$ we mean a cyclic sequence of $n \geq 3$ points V_i in the affine plane, together with n closed line segments $S_i = [V_i, V_{i+1}]$. The points V_i are the *vertices* of P and the segments S_i are the *edges* of P. Here and throughout, all subscripts i are reduced modulo n so that they satisfy $1 \leq i \leq n$; moreover, to avoid special cases and degeneracies, we shall always assume that adjacent vertices V_i, V_{i+1} are distinct. A polygon is regarded as oriented and is unchanged by any cyclic permutation of the vertices; however, reversal of the order of the vertices produces a new polygon P', which we shall say is obtained from P by *reversing the orientation*.

A polygon with n vertices, and therefore n edges, is called an *n-gon*. Notice that these polygons are very general; non-adjacent vertices may coincide, and edges may cross or partially or wholly overlap. It is to such polygons that our theorems will apply. Sometimes it is necessary to place further restrictions on the positions of the vertices. These will be mentioned when appropriate.

It is convenient to introduce the notion of a *side* V_iV_{i+1} of a polygon; this is the line containing the edge $[V_i, V_{i+1}]$. A *diagonal* V_iV_j of a polygon is the line defined by two non-adjacent vertices V_i, V_j of P. It is defined if, and only if, V_i and V_j are distinct points[2].

We use uppercase letters $V, X, W, \ldots$, with or without subscripts, for points, and the corresponding lowercase letters $v, x, y, \ldots,$ for their position vectors relative to an arbitrarily chosen origin O. Position vectors can be used, in a perfectly rigorous manner, in affine geometry.

Suppose A, B, C, D are four points such that AB is parallel to CD. (In other words, the directions of the vectors $b - a$ and $d - c$ either coincide or are directly

[2]Notice that the symbols $[A, B]$ and $[A, B, C]$ differ in meaning from the frequently used $[AB] = \text{conv}(A, B)$ and $\text{conv}(A, B, C)$ in that the latter refer to *sets*, whereas the symbols used here also imply an *orientation* of each set.

opposite.) Let λ be the real number defined by $\lambda(d-c)=(b-a)$. Then λ is an affine invariant. It is simply the ratio of the length of $[A, B]$ to that of $[C, D]$, with a plus or a minus sign according to whether these line segments have the same, or opposite directions. We shall denote this ratio by $[A\ B/C\ D]=\lambda$. Clearly, relations such as

$$\left[\frac{A\ B}{C\ D}\right]=-\left[\frac{B\ A}{C\ D}\right]=-\left[\frac{A\ B}{D\ C}\right]=\left[\frac{B\ A}{D\ C}\right]$$

hold, and cancellation is permitted so that, for example,

$$\left[\frac{A\ B}{C\ D}\right]\cdot\left[\frac{C\ D}{E\ F}\right]=\left[\frac{A\ B}{E\ F}\right].$$

These properties follow directly from the definitions.

In particular, if a point W is defined in some prescribed manner on the side V_iV_{i+1} of a polygon P, and is distinct from both V_i and V_{i+1}, then $[V_iW/WV_{i+1}]$ will be referred to as an *edge-ratio* corresponding to that side. In a similar manner, if W lies on a diagonal V_iV_j, then $[V_iW/WV_j]$ is called a *diagonal-ratio*.

Similar considerations apply to areas. In the Euclidean plane we define the signed area $a(A, B, C)$ of a triangle $[A, B, C]$ to be its area (in the usual sense) prefixed by a plus or a minus sign. The $+$ sign is used if the triangle is positively oriented, that is, the vertices are named in a counterclockwise direction; and the $-$ sign is used if the triangle is negatively oriented, that is, the vertices are named in a clockwise direction. For two triangles $[A, B, C]$ and $[D, E, F]$ in the plane, the ratio $a(A, B, C)/a(D, E, F)$ of their signed areas is an affine invariant that will be denoted by $\lambda=[A\ B\ C/D\ E\ F]$. In terms of position vectors, λ is defined by $(a\times b+b\times c+c\times a)=\lambda(d\times e+e\times f+f\times d)$, where $\times$ signifies a vector (cross) product. Since all the points lie in a plane, these product vectors are all parallel. Equivalently, λ is defined by

$$\det\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ 1 & 1 & 1 \end{bmatrix}=\lambda\det\begin{bmatrix} d_1 & e_1 & f_1 \\ d_2 & e_2 & f_2 \\ 1 & 1 & 1 \end{bmatrix},$$

where (a_1, a_2) are the coordinates of the point A, that is, the components of the position vector a, etc.

Since cyclic permutation of vertices does not change area, but reversing the order of the vertices changes the sign of the area, relations such as

$$\left[\frac{A\ B\ C}{D\ E\ F}\right]=\left[\frac{B\ C\ A}{D\ E\ F}\right]=\left[\frac{C\ A\ B}{D\ E\ F}\right]=-\left[\frac{B\ A\ C}{D\ E\ F}\right]=\left[\frac{B\ C\ A}{E\ F\ D}\right]=\left[\frac{B\ A\ C}{E\ D\ F}\right]$$

hold. Moreover, as in the case of line segments, cancellation is permitted; for example,

$$\left[\frac{A\ B\ C}{D\ E\ F}\right]\cdot\left[\frac{D\ E\ F}{G\ H\ J}\right]=\left[\frac{A\ B\ C}{G\ H\ J}\right].$$

In this notation the area principle (1) may be written in the slightly more powerful form

$$\left[\frac{A_1P}{A_2P}\right]=\left[\frac{A_1BC}{A_2BC}\right],$$

which takes account of the signed lengths and signed areas of the triangles.

The *signed area* of an n-gon ($n > 3$) can also be defined; we triangulate the n-gon in any way, and orient the triangles *coherently* (see FIGURE 5). This means that whenever two triangles of the triangulation have an edge $[B, C]$ in common, then their orientations are such that they induce opposite directions on this common edge. The area of the polygon is then defined as the sum of the signed areas of the component triangles. For the validity of this definition it is, of course, essential to show that the signed area is independent of the triangulation used. This is a routine calculation, and we do not give details here. In fact, if one interprets the polygon as an oriented curve, its signed area is exactly what one obtains by applying to it the familiar integrals of calculus.

We invite the reader to adapt the proofs given above using signed areas and the new notation, and so provide general proofs for the theorems we discussed. In particular, if

$$\left[\frac{V_iW_i}{W_{i+1}V_{i+2}}\right] = \left[\frac{V_iV_{i+1}V_{i+4}}{V_iV_{i+1}V_{i+2}V_{i+4}}\right] \cdot \left[\frac{V_iV_{i+1}V_{i+2}V_{i+3}}{V_{i+2}V_{i+3}V_{i+1}}\right], \tag{5a}$$

is used instead of (5), a proof of Hoehn's Theorem is obtained that does not depend on the convexity, simplicity or orientation of the pentagon $[V_1, V_2, V_3, V_4, V_5]$.

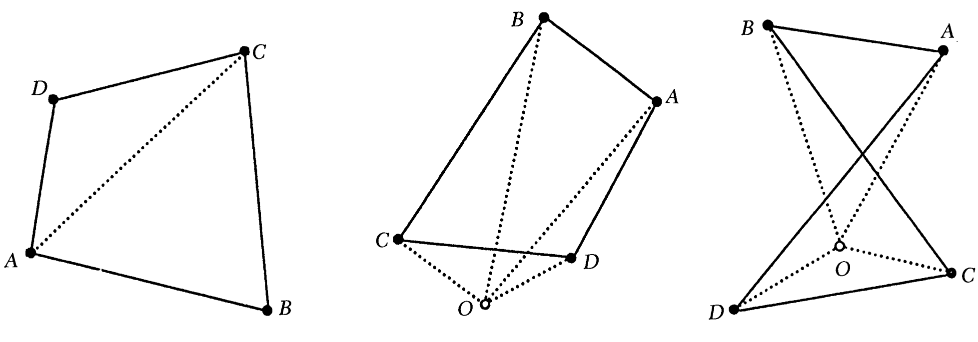


FIGURE 5

Three examples of coherent triangulations of a quadrangle $[A, B, C, D]$. In (a), the quadrangle is triangulated into two triangles $[A, B, C]$ and $[A, C, D]$. In (b) and (c) the triangles are $[O, A, B]$, $[O, B, C]$, $[O, C, D]$, $[O, D, A]$.

3. Generalizations of the Theorems of Ceva and Menelaus

In this and the following section we shall define points W_i on sides or diagonals of an n-gon as points of intersection with other lines. Throughout we shall assume, without further remark, that these points exist (the lines concerned are not parallel) and are distinct from the vertices (so that the relevant edge- and diagonal-ratios are well-defined). This section is concerned with problems of the following type. On each side V_iV_{i+1} of an n-gon, a point W_i is defined in some geometrically meaningful way. Under what circumstances is it possible to make an assertion about the value of the product of the edge-ratios $[V_iW_i/W_iV_{i+1}]$? We shall also consider similar problems for diagonal-ratios.

Our first result of this kind is the generalization of Menelaus' Theorem to n-gons. This is not new—see Section 5. The case $n = 5$ is shown in FIGURE 6.

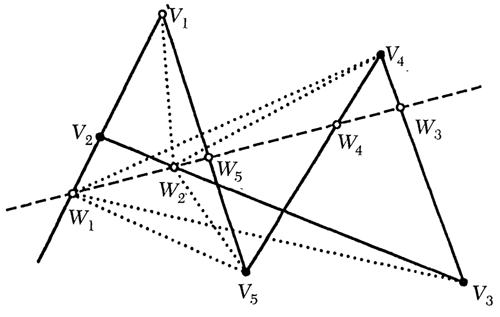


FIGURE 6
Menelaus' Theorem for a pentagon states that

$$\left[\frac{V_1W_1}{W_1V_2}\right]\cdot\left[\frac{V_2W_2}{W_2V_3}\right]\cdot\left[\frac{V_3W_3}{W_3V_4}\right]\cdot\left[\frac{V_4W_4}{W_4V_5}\right]\cdot\left[\frac{V_5W_5}{W_5V_1}\right]=-1.$$

THEOREM 1 (Menelaus' Theorem for n-gons). *Let $P=[V_1,\dots,V_n]$ be an n-gon and suppose that, for $i=1,\dots,n$, a transversal cuts the side V_iV_{i+1} in W_i. Then the product of the edge-ratios is constant, in fact*

$$\prod_{i=1}^{n}\left[\frac{V_iW_i}{W_iV_{i+1}}\right]=(-1)^n \tag{6}$$

for all n-gons P.

In the case where n is odd, the product is negative, which shows that an odd number of intersections of the transversal with the sides V_iV_{i+1} of the n-gon must be *external*, that is, do not lie on the edge $[V_i, V_{i+1}]$. This is, of course, a familiar fact in the case $n=3$; it is known as *Pasch's axiom* [**3**, §12.2]. If n is even, an even number (possibly zero) of intersections of the transversal with the sides of the n-gon must be external.

Proof. Select any two of the points W_i, say W_1, W_2 (see FIGURE 6). Then the area principle for triangles with base $[W_1, W_2]$ yields

$$\left[\frac{V_iW_i}{W_iV_{i+1}}\right]=-\left[\frac{V_iW_1W_2}{V_{i+1}W_1W_2}\right].$$

Substituting for each of the n factors on the left side of (5) we obtain a product of terms each of which is the quotient of the areas of triangles. These cancel to yield the value $(-1)^n$ as required.

A trivial variation of this theorem applies to diagonals instead of edges. If $1<j<n$, and W_i is defined as the intersection of a transversal with the diagonal V_iV_{i+1} for $i=1,\dots,n$, then

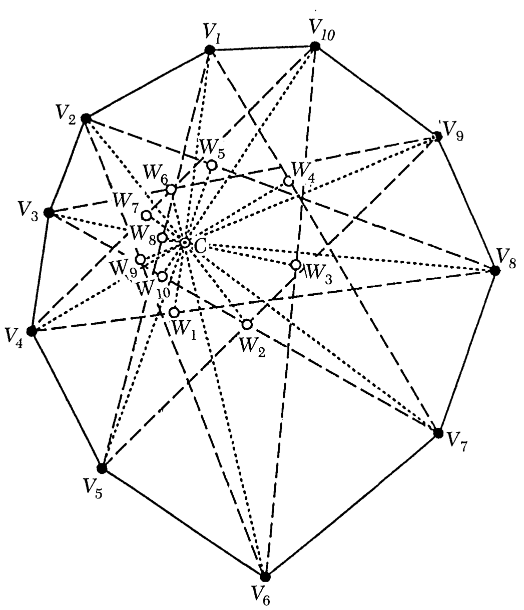
$$\prod_{i=1}^{n}\left[\frac{V_iW_i}{W_iV_{i+j}}\right]=(-1)^n.$$

Despite the apparent generality, this can easily be deduced from Theorem 1 by renaming the vertices $V_1, V_2, V_3,\dots$ as $V_i, V_{i+j}, V_{i+2j},\dots$. This renumbering may (if j is not prime to n) split the original polygon into two or more polygons, but this does not affect the result or its proof.

THEOREM 2 (Ceva's Theorem for n-gons). *Let $P=[V_1,\dots,V_n]$ be an arbitrary n-gon, C a given point, and k a positive integer such that $1\le k<n/2$. For $i=1,\dots,n$ let W_i be the intersection of the lines CV_i and $V_{i-k}V_{i+k}$. Then,*

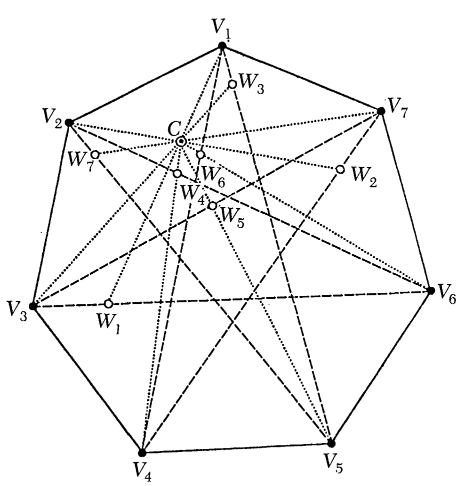
$$\prod_{i=1}^{n}\left[\frac{V_{i-k}W_i}{W_iV_{i+k}}\right]=1. \tag{7}$$

Note that this result concerns diagonal-ratios except in the case n odd and $k=(n-1)/2$. The latter is the only generalization of Ceva's Theorem that we have been able to find in the literature (see, for example, [**6**, p. 86]). However, the result of Theorem 2 is not as general as it appears at first sight. If n and k have highest common factor $d>1$ then the product on the left side of (7) can be split into d products, each with n/d factors, and each of these factors has the value 1. If n and $2k$ are relatively prime then it is possible (by renumbering the vertices as in the case of Menelaus' Theorem) to transform the left side of (7) into a product of n edge-ratios for an appropriate star n-gon. Hence the only genuinely new assertion in Theorem 2 is in the case where n is even and k is prime to n. An example of this case ($n=10, k=3$) is shown in FIGURE 7(a). Other examples appear in FIGURE 7(b) with diagonal-ratios ($n=7, k=2$) and in FIGURE 7(c) with edge-ratios ($n=5, k=2$).

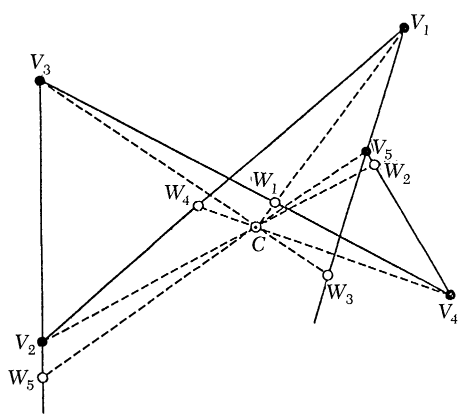


(a)

FIGURE 7
Examples of Ceva's Theorem for n-gons in the cases (a) $n=10, k=3$; (b) $n=7, k=2$; and (c) $n=5, k=2$. In cases (a) and (b) the theorem makes an assertion about diagonal-ratios $[V_iW_{i+3}/W_{i+3}V_{i+6}]$ and $[V_iW_{i+2}/W_{i+2}V_{i+4}]$, respectively, and in (c) an assertion about edge-ratios $[V_iW_{i+3}/W_{i+3}V_{i+1}]$.



(b)



(c)

Proof. The proof follows similar lines to that of Theorem 1. We observe that, applying the area principle to triangles with base $[C, V_i]$ we obtain, for $i = 1, \dots, n$,

$$\left[\frac{V_{i-k}W_i}{W_iV_{i+k}}\right] = \left[\frac{CV_iV_{i-k}}{CV_{i+k}V_i}\right].$$

Substituting these terms in the left side of (6), we obtain a product of n terms each of which is a quotient of the areas of the triangles. These cancel to yield the value 1 as required.

In each of the above theorems we have defined W_i as the intersection of a side or diagonal (a *basis*) with a line CD (the corresponding *transversal*). Menelaus' Theorem corresponds to the case where C and D are fixed points that are not vertices, so the transversal CD is a fixed line for all bases. In Ceva's Theorem C is a fixed point and D is a vertex V_i of the polygon P. Then W_i is the point of intersection of the basis $V_{i-k}V_{i+k}$ with the transversal CV_i. These remarks suggest that there may exist corresponding results for products of edge- or diagonal-ratios when the transversal is determined by two vertices of P. Indeed, we have the following theorem, which appears to be new.

THEOREM 3 (The Selftransversality Theorem). *Let j, r, s be integers distinct* (mod n) *and let W_i be the point of intersection of the basis (side or diagonal) V_iV_{i+j} of the n-gon $P = [V_1, \dots, V_n]$ with the transversal $V_{i+r}V_{i+s}$. Then a necessary and sufficient condition for*

$$\prod_{i=1}^{n}\left[\frac{V_iW_i}{W_iV_{i+j}}\right] = (-1)^n \tag{8}$$

is that either (i) $n = 2m$ *is even,* $j \equiv m$ *and* $s \equiv r + m$; *or that n is arbitrary and either* (ii) $s \equiv 2r$ *and* $j \equiv 3r$; *or* (iii) $r \equiv 2s$ *and* $j \equiv 3s$. *All congruences are* mod n.

In case (i) the terms in (8) cancel in pairs so the statement of the theorem becomes trivial. Cases (ii) and (iii) are essentially the same since r and s play a symmetrical role in determining the transversal V_{i+r}, V_{i+s}, and so may be interchanged.

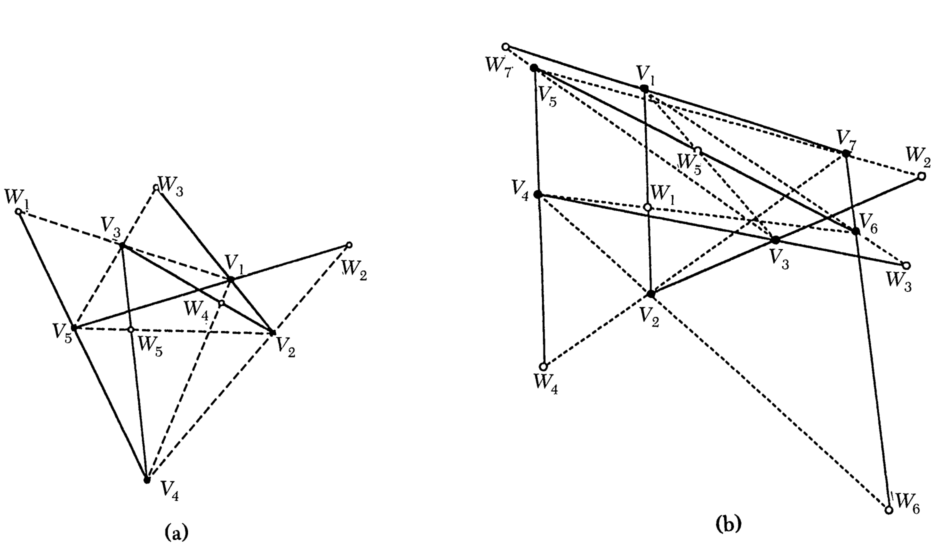


FIGURE 8

Examples of the Selftransversality Theorem for (a) $n = 5, j = 2, r = 3, s = 4$; and (b) $n = 7$, $j = 1, r = 3, s = 5$. In each case the theorem shows that the product of the n ratios $[V_iW_i/W_iV_{i+j}]$ (where $i = 1, \dots, n$) is -1.

For example, in FIGURE 8(a), $n=5, j=2, r=3, s=4$, and in FIGURE 8(b), $n=7$, $j=1, r=3, s=5$ so each is an instance of case (iii); since n is odd, the sign on the right side of (8) is negative.

Proof. Using the area principle for triangles with base $[V_{i+r}, V_{i+s}]$ and apexes V_i and V_{i+j}, we obtain

$$\left[\frac{V_iW_i}{W_iV_{i+j}}\right] = -\left[\frac{V_iV_{i+r}V_{i+s}}{V_{i+j}V_{i+r}V_{i+s}}\right]. \tag{8a}$$

We substitute these expressions for each of the n factors on the left side of (8) and determine when exactly the same triangles occur in both the numerator and denominator and so their areas (as expressed in terms of determinants) cancel to yield the value ± 1 as required. The term $V_iV_{i+r}V_{i+s}$ in the numerator will cancel with the term $V_{(i+h)+j}V_{(i+h)+r}V_{(i+h)+s}$ in the denominator if, and only if, either

(i) $h \equiv -j,\ r \equiv s-j$ and $s \equiv r-j$; or
(ii) $h \equiv -r,\ r \equiv s-r$ and $s \equiv j-r$; or
(iii) $h \equiv -s,\ s \equiv r-s$ and $r \equiv j-s$.

These alternatives correspond to the three cases given in the statement of the theorem. Notice that each cancellation produces the factor -1 in case (i) and $+1$ in the other two cases, leading to the term $(-1)^n$ on the right side of (8). Thus the theorem is proved.

It is of interest to determine, for a given basis, the corresponding transversals for which the theorem holds. Let us consider case (ii). Writing $j+t$ instead of r and $-u$ instead of s, it is easily verified that the given condition is equivalent to $u \equiv t$ and $2j+3t \equiv 0 \pmod n$. This can be solved explicitly for t (and so r and s determined) in terms of j as follows:

(ii_a) If $n \equiv 3k+1$, then $t \equiv 2kj$ for $j = 1, 2, \ldots, [(n-1)/2]$.
(ii_b) If $n \equiv 3k-1$, then $t \equiv (k-1)j$ for $j = 1, 2, \ldots, [(n-1)/2]$.
(ii_c) If $n \equiv 3k$, then $j \equiv 3i$ and $t \equiv -2i,\ k-2i$ or $2k-2i$, for $i = 1, 2, \ldots, [(k-1)/2]$.

For example, in FIGURE 9 we show the case $n=9$, $k=3$ and $i=1$. The transversals V_2V_3, V_6V_8 and V_5V_9 are indicated by dotted lines and the basis V_1V_4 by a dashed line. All these satisfy the conditions of the theorem. (The other eight bases that occur in (8) are not shown since to indicate these would make the figure so complicated as to be unintelligible.)

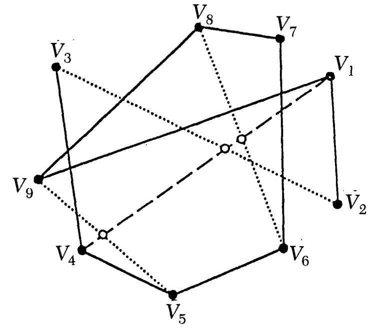


FIGURE 9
The three transversals V_2V_3, V_6V_8, V_5V_9 of the basis V_1V_4 that satisfy the conditions of the Selftransversality Theorem for $n=9$.

4. Generalizations of Hoehn's Theorem

In this section, we consider theorems similar to those in the previous section in that, for every n-gon, the product of n ratios of lengths of line segments has some constant value, namely $+1$ or -1. However, we now consider those cases where, as in Hoehn's Theorem, the line-segments lie on the same line but are not contiguous. Thus we define two points W_i, Z_i on certain edges or diagonals $[V_i, V_j]$ of an n-gon, and consider products of the form

$$\prod_{i=1}^{n}\left[\frac{V_iW_i}{Z_iV_j}\right].$$

THEOREM 4 (Hoehn's First Theorem for n-gons). *Let $P=[V_1,\dots,V_n]$ be a given n-gon and j an integer such that, for each i, the integers $i, i+j, i+2j, i+3j, i+4j$ are distinct* (mod n). *Define W_i (for $i=1,\dots,n$) to be the point of intersection of V_iV_{i+2j} and $V_{i+j}V_{i+3j}$. Then the points W_i and W_{i+j} lie on the line $V_{i+j}V_{i+3j}$ and*

$$\prod_{i=1}^{n}\left[\frac{V_{i+j}W_i}{W_{i+j}V_{i+3j}}\right]=1. \tag{9}$$

It will be observed that if $V_i, V_{i+j}, V_{i+2j}, V_{i+3j}, V_{i+4j}$ are not distinct points, then the identity becomes meaningless or trivial. Further, it is only necessary to consider the cases $j=1,2,\dots,[(n-1)/2]$ since other values of j lead to repetitions of the same result.

FIGURE 10 illustrates the theorem for $n=7, j=1,2$ and 3. The original statement (3) of Hoehn's Theorem (see FIGURE 4) corresponds to the case $n=5, j=1$.

Proof. For the first assertion we note that W_i is the intersection of V_iV_{i+2j} and $V_{i+j}V_{i+3j}$, and W_{i+j} is the intersection of $V_{i+j}V_{i+3j}$ and $V_{i+2j}V_{i+4j}$. Hence both of these points lie on $V_{i+j}V_{i+3j}$ as stated.

For the second assertion we note that, as in (5a), using triangles with bases $[V_i, V_{i+2j}]$ and $[V_{i+2j}, V_{i+4j}]$ we obtain

$$\left[\frac{V_{i+j}W_i}{W_{i+j}V_{i+3j}}\right]=\left[\frac{V_iV_{i+j}V_{i+2j}}{V_{i+2j}V_{i+3j}V_{i+4j}}\right]\cdot\left[\frac{V_{i+2j}V_{i+3j}V_{i+4j}V_{i+j}}{V_iV_{i+j}V_{i+2j}V_{i+3j}}\right].$$

Inserting these expressions in the left side of (9) we obtain a product in which the areas of the triangles, as well as the areas of the quadrilaterals, all cancel. Hence the product has the value 1, and the theorem is proved.

THEOREM 5 (Hoehn's Second Theorem for n-gons). *Let $P=[V_1,\dots,V_n]$ be an n-gon, and j,k positive integers such that $j+2k=n$, and for each $i=1,\dots,n$, the integers $i, i+k, i+j, i+j+k$ are distinct* (mod n) *and the integers $i, i+k, i+2k, i+3k$ are also distinct* (mod n). *Define W_i as the intersection of V_iV_{i+k} and $V_{i+j}V_{i+j+k}$. Then the points W_i and W_{i+2k} lie on the line V_iV_{i+k} and*

$$\prod_{i=1}^{n}\left[\frac{V_iW_i}{W_{i+2k}V_{i+k}}\right]=1. \tag{10}$$

Proof. For the first assertion, we note that W_i is the intersection of V_iV_{i+k} and $V_{i+j}V_{i+j+k}$, and W_{i+2k} is the intersection of $V_{i+2k}V_{i+3k}$ and $V_{i+j+2k}V_{i+j+3k}$, which is the same line as V_iV_{i+k} since $j+2k=n$.

For the second assertion, we again use the area principle, as in (5a), for triangles with bases $[V_{i+j}, V_{i+j+k}]$ and $[V_{i+2k}, V_{i+3k}]$ to yield

$$\left[\frac{V_i W_i}{W_{i+2k} V_{i+k}}\right] = \left[\frac{V_i V_{i+j} V_{i+j+k}}{V_{i+k} V_{i+2k} V_{i+3k}}\right] \cdot \left[\frac{V_{i+k} V_{i+2k} V_i V_{i+3k}}{V_i V_{i+j} V_{i+k} V_{i+j+k}}\right].$$

Substituting in the left side of (10) and using the relation $j + 2k = n$ we see that the areas of the triangles, and the areas of the quadrilaterals in the resulting product, all cancel to yield the value 1. This proves the theorem.

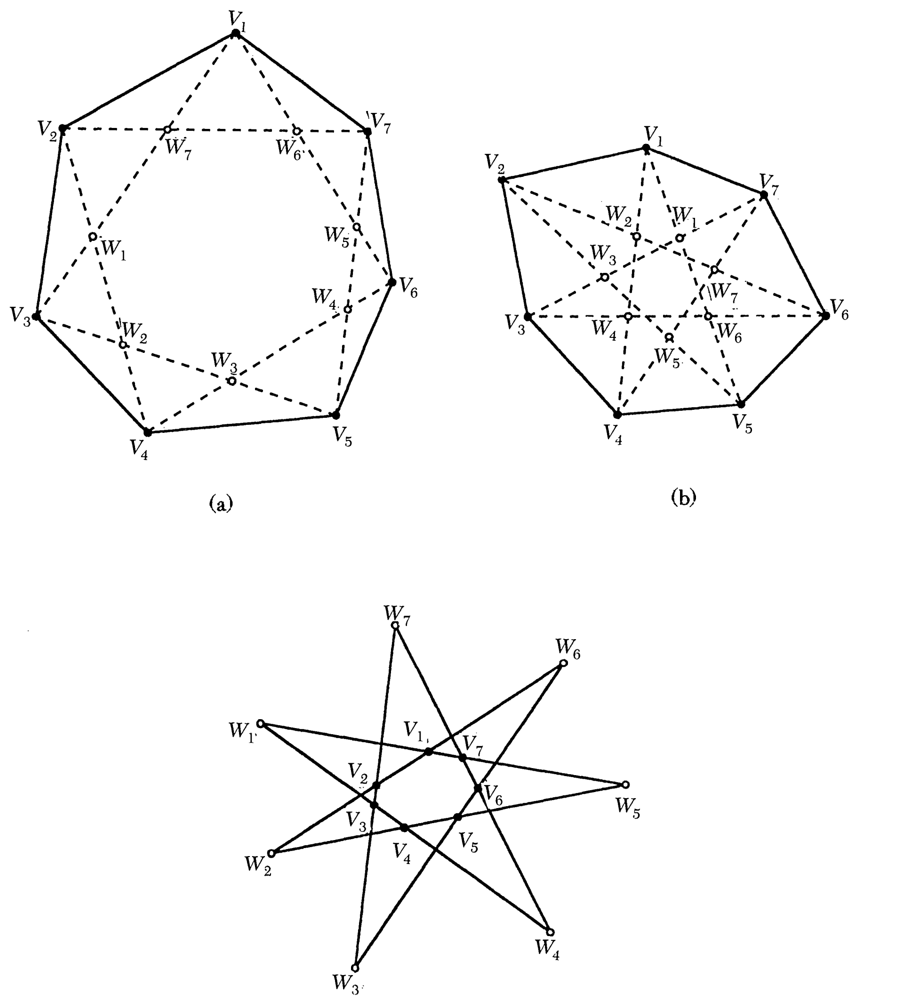


FIGURE 10

Examples of Hoehn's First Theorem for n-gons, for $n = 7$ and (a) $j = 1$; (b) $j = 2$; and (c) $j = 3$. In each case the product of the seven ratios $[V_{i+j}W_i/W_{i+j}V_{i+3j}]$ (for $i = 1, \ldots, 7$) takes the value 1.

The theorem is illustrated in FIGURE 11 for $n = 7$ and $(j, k) = (1, 3)$, $(3, 2)$ and $(5, 1)$. The original statement (4) of Hoehn's Theorem corresponds to the case $n = 5, j = 1, k = 2$.

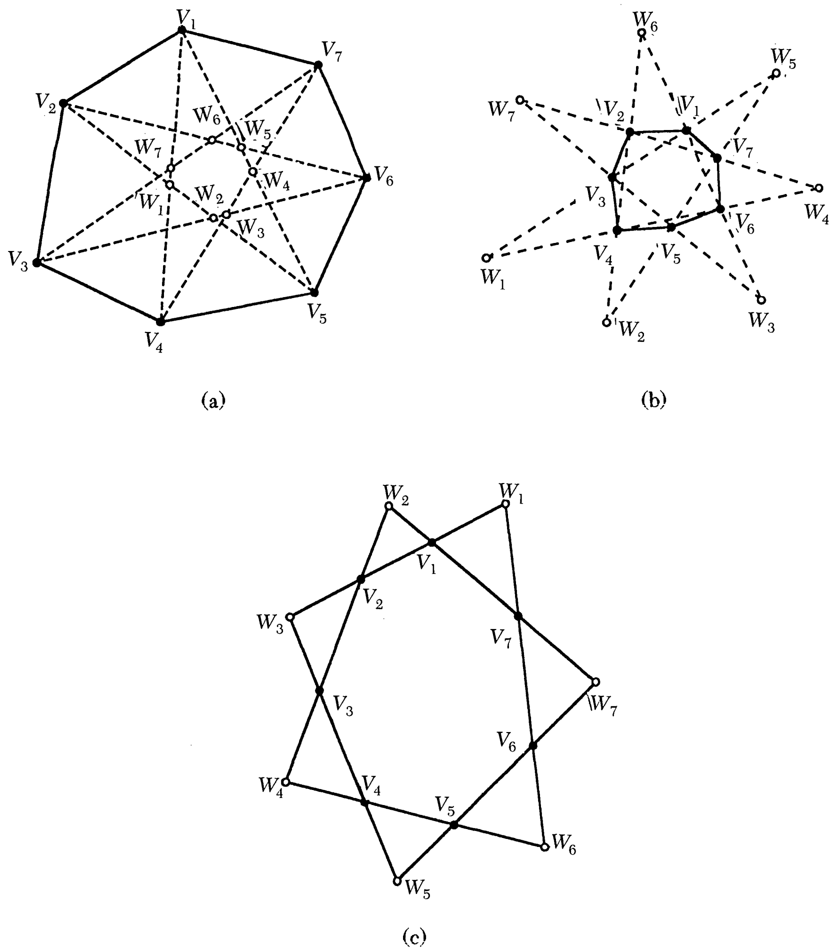


FIGURE 11

Hoehn's Second Theorem for n-gons, for $n = 7$ and (a) $j = 1, k = 3$; (b) $j = 3, k = 2$; and (c) $j = 5$, $k = 1$. In each case the product of the seven ratios $[V_iW_i/W_{i+2k}V_{i+k}]$ (for $i = 1, \dots, 7$) takes the value 1.

5. Comments

Menelaus' Theorem for n-gons is not new—it has been known since the early nineteenth century, see [**2**, p. 295], [**6**, p. 75], [**7**, p. 63], [**9**, p. 75]. As mentioned above, special cases of Ceva's Theorem for n-gons have also been established; see, for example, [**6**, p. 86], [**7**, p. 64].

The most interesting and unexpected feature of the results of this paper is *not* that the various products of edge- and diagonal-ratios are equal to $+1$ or -1, but that the values of these products are independent of the polygon P from which we start the construction (subject to the restrictions stated at the beginning of Section 3). In view of this, one might suppose that all products of n ratios in an n-gon have the same property; however, this is not the case. For example, in FIGURE 4, the value of

$$\prod_{i=1}^{5} \left[\frac{W_iW_{i+1}}{V_iV_{i+2}} \right] \tag{11}$$

depends on the pentagon P which is chosen—in spite of the fact that such a product appears, at first sight, to be very similar to those that occur in the original statement of Hoehn's Theorem. That (11) is not constant can be shown by a little experimentation with numerical examples. In fact, we suspect that the product (11) attains its maximum value when $P = [V_1, \ldots, V_5]$ is an affine image of a regular pentagon. We have no proof of this statement, and a discussion of problems of this nature does not seem appropriate here.

It is not difficult to see that Theorems 4 and 5 cover all the cases mentioned in the introduction to Section 4, namely those in which a product of n factors takes a constant value that does not depend on the choice of the original polygon P. If we extend our investigation to consider products of $2n, 3n, 4n, \ldots$ factors (each of which is the quotient of lengths of line-segments in an n-gon) then many more possibilities arise. The reader may wish to investigate these. However, the results we have stated here will illustrate the versatility and power of the area principle in proving results of this nature.

The case $n = 4, j = 2, k = 1$ of Theorem 5 deserves special mention. It is illustrated in FIGURE 12, and it is clear that it can be interpreted as stating that the product of certain ratios of lengths in a complete quadrilateral equals 1. Although complete quadrilaterals have been thoroughly investigated for two centuries, we were unable to find any mention of this particular result.

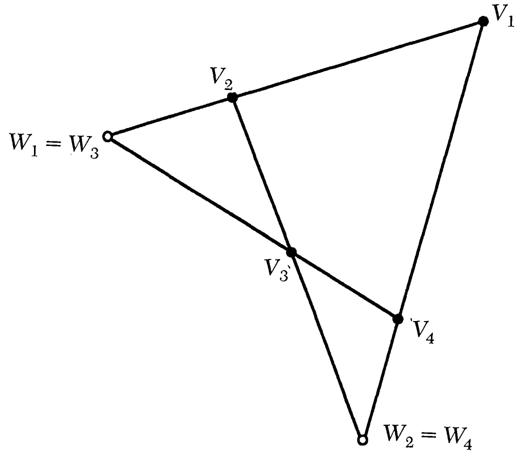


FIGURE 12

The case $n = 4, j = 2, k = 1$ of Hoehn's Second Theorem. The product of the four ratios $[V_iW_i/W_{i+2}V_{i+1}]$ (for $i = 1, 2, 3, 4$) equals 1.

Since all the theorems in this paper are so straightforward and their proofs so elementary, it is surprising that they were not discovered two or three centuries ago. This may be partly explained by the fact that, unlike us, earlier authors did not have the advantages of modern technology. When we started this investigation we were, of course, familiar with Ceva's and Menelaus' Theorems for triangles; it was Hoehn's Theorem that suggested to us that it might be worthwhile to investigate products of other ratios in n-gons with $n > 3$. Using a simple program in Mathematica® we were able to accurately calculate the values of circular products of various ratios for large numbers of n-gons, and the results suggested our Theorems 1 to 5. (Later on we discovered that the n-gonal forms of Ceva's and Menelaus' Theorems are known and can be found in the literature. However, so far we were unable to find any result that can be interpreted as even a particular case of our Selftransversality Theorem.) We observed that the "area principle" was an easy method of proving (in the traditional sense) all the results that had been discovered empirically. Further, it suggested other possibilities as well as enabling us to state the theorems for n-gons with arbitrary n. It also led to higher-dimensional generalizations, which will be presented elsewhere.

Note. Since the completion of the manuscript we have learned from Baptist [**1**, p. 61] that the "area principle" was used—without any special name—in Crelle's proof of Ceva's Theorem [**5**]. But its general utility was not noticed and so it was, for all practical purposes, completely forgotten.

REFERENCES

1. P. Baptist, *Die Entwicklung der neueren Dreiecksgeometrie*, BI Wissenschaftsverlag, Mannheim, 1992.
2. L. N. M. Carnot, *Géometrie de position*, Duprat, Paris, 1803.
3. H. S. M. Coxeter, *Introduction to Geometry*, 2nd edition, John Wiley & Sons, Inc., New York, 1969.
4. H. S. M. Coxeter and S. L. Greitzer, *Geometry Revisited*, Random House, New York, 1967.
5. A. L. Crelle, *Über einige Eigenschaften des ebenen geradlinigen Dreiecks rücksichtlich dreier durch die Winkel-Spitzen gezogenen geraden Linien*, Berlin, 1816.
6. H. Eves, *A Survey of Geometry*, Allyn and Bacon, Boston, 1963.
7. H. Eves, *College Geometry*, Jones and Bartlett, Boston, 1995.
8. L. Hoehn, A Menelaus-type theorem for the pentagram, this MAGAZINE 66 (1993), 121–123.
9. J.-V. Poncelet, *Traité des propriétés projectives des figures*, Mallet-Bachelier, Paris 1821; 2nd edition, Gauthier-Villars, Paris, 1865.

Carl B. Allendoerfer Awards 1995

The Carl B. Allendoerfer Awards, established in 1976, are made to authors of expository articles published in *Mathematics Magazine*. The awards are named for Carl B. Allendoerfer, a distinguished mathematician at the University of Washington and President of the Mathematical Association of America, 1959–1960.

Lee Badger
"Lazzarini's Lucky Approximation of π"
***Mathematics Magazine* 67 (1994), 83–91.**

This interesting article combines a famous problem in probability with statistical methods of data analysis that can demonstrate fraud—or what the author more charitably calls "hoaxes." The example is concrete, the mathematics is rich and detailed, and the material is accessible to students with no more than a calculus background. There are broader lessons for the reader too: rigged data are a fact of life, but so are the statistical tools that can detect them.

Tristam Needham
"The Geometry of Harmonic Functions"
***Mathematics Magazine* 67 (1994), 92–108.**

The geometric approach of this carefully crafted and well-written article is enlightening; there is much here for the reader to learn about the interplay between geometry and analysis. In particular, we see how material almost invariably treated by analytic methods can be rethought—perhaps in ways closer to those used when the ideas were being developed. The author also helps us to understand what something means and not just why it might be possible to prove it true. In this article, geometry becomes a powerful tool for conveying meaning.

sense) all the results that had been discovered empirically. Further, it suggested other possibilities as well as enabling us to state the theorems for n-gons with arbitrary n. It also led to higher-dimensional generalizations, which will be presented elsewhere.

Note. Since the completion of the manuscript we have learned from Baptist [**1**, p. 61] that the "area principle" was used—without any special name—in Crelle's proof of Ceva's Theorem [**5**]. But its general utility was not noticed and so it was, for all practical purposes, completely forgotten.

REFERENCES

1. P. Baptist, *Die Entwicklung der neueren Dreiecksgeometrie*, BI Wissenschaftsverlag, Mannheim, 1992.
2. L. N. M. Carnot, *Géometrie de position*, Duprat, Paris, 1803.
3. H. S. M. Coxeter, *Introduction to Geometry*, 2nd edition, John Wiley & Sons, Inc., New York, 1969.
4. H. S. M. Coxeter and S. L. Greitzer, *Geometry Revisited*, Random House, New York, 1967.
5. A. L. Crelle, *Über einige Eigenschaften des ebenen geradlinigen Dreiecks rücksichtlich dreier durch die Winkel-Spitzen gezogenen geraden Linien*, Berlin, 1816.
6. H. Eves, *A Survey of Geometry*, Allyn and Bacon, Boston, 1963.
7. H. Eves, *College Geometry*, Jones and Bartlett, Boston, 1995.
8. L. Hoehn, A Menelaus-type theorem for the pentagram, this MAGAZINE 66 (1993), 121–123.
9. J.-V. Poncelet, *Traité des propriétés projectives des figures*, Mallet-Bachelier, Paris 1821; 2nd edition, Gauthier-Villars, Paris, 1865.

Carl B. Allendoerfer Awards 1995

The Carl B. Allendoerfer Awards, established in 1976, are made to authors of expository articles published in *Mathematics Magazine*. The awards are named for Carl B. Allendoerfer, a distinguished mathematician at the University of Washington and President of the Mathematical Association of America, 1959–1960.

Lee Badger
"Lazzarini's Lucky Approximation of π"
***Mathematics Magazine* 67 (1994), 83–91.**

This interesting article combines a famous problem in probability with statistical methods of data analysis that can demonstrate fraud—or what the author more charitably calls "hoaxes." The example is concrete, the mathematics is rich and detailed, and the material is accessible to students with no more than a calculus background. There are broader lessons for the reader too: rigged data are a fact of life, but so are the statistical tools that can detect them.

Tristam Needham
"The Geometry of Harmonic Functions"
***Mathematics Magazine* 67 (1994), 92–108.**

The geometric approach of this carefully crafted and well-written article is enlightening; there is much here for the reader to learn about the interplay between geometry and analysis. In particular, we see how material almost invariably treated by analytic methods can be rethought—perhaps in ways closer to those used when the ideas were being developed. The author also helps us to understand what something means and not just why it might be possible to prove it true. In this article, geometry becomes a powerful tool for conveying meaning.

Lewis Carroll and the Enumeration of Minimal Covers

ANTHONY J. MACULA
SUNY at Geneseo
Geneseo, NY 14454

1. Introduction

Let $[n]=\{1,2,\dots,n\}$. A *cover* of $[n]$ is a family γ of non-empty subsets of $[n]$ such that $[n]=\bigcup\{S\colon S\in\gamma\}$. An *ordered k-cover* of $[n]$ is a k-tuple $\tau=(S_1,\dots,S_k)$ such that the family $\{S_i\}$ is a cover of $[n]$. A cover γ is called *minimal* if the removal of one member destroys the covering property, i.e., for each $S\in\gamma$ there is an $i\in S$ such that $i\notin T$ for all $T\in\gamma$ with $S\neq T$. Any non-trivial partition of $[n]$ is a minimal cover, but $\{\{2\},\{1,3\},\{3,4\}\}$ is a minimal cover of [4] that is not a partition. In this way the notion of a minimal cover can be thought of as extending the idea of a partition. The number of partitions of $[n]$ into k nonempty subsets, denoted by $S(n,k)$, is given by

$$S(n,k)=\frac{1}{k!}\sum_{i=1}^{k}(-1)^{k-i}\binom{k}{i}i^n.$$

These numbers are the well-known *Stirling numbers* (of the second kind) and can also be recursively defined by the relation $S(n,k)=S(n-1,k-1)+kS(n-1,k)$ where $S(n,1)=S(n,n)=1$ and $2\le k\le n-1$.

Analogously, let $\mu(n,k)$ be the number of minimal covers of $[n]$ with k members. In **[4]**, $\mu(n,k)$ was computed, and values for a few n and k appear in Table 2. In this note, we compute $\mu(n,k)$ by a different method involving a straightforward application of the seldom seen set diagrams of Lewis Carroll and the basic combinatorial concept of placing objects in boxes. We say $\sigma=(S_1,\dots,S_k)$ is a *minimal ordered k-cover* of $[n]$ if the family $\{S_i\}$ is a minimal cover of $[n]$. $OM_k(n)$ denotes the family of minimal ordered k-covers and $|OM_k(n)|$ denotes its cardinality. Our plan is to compute $|OM_k(n)|$; then

$$\mu(n,k)=\frac{1}{k!}\,|OM_k(n)|.$$

2. Lewis Carroll's Set Diagrams

Have you ever tried to draw a Venn diagram, depicting all possible intersections, using four, five, or six sets? It can be quite cumbersome. In **[2]**, Lewis Carroll provides an easy method for drawing set diagrams depicting all possible intersections. Using his method, one can actually draw these diagrams for ten or more sets. The only limitations are the patience of the drafter and the size of the paper on which the diagram is drawn.

We affectionately refer to the Lewis Carroll set diagram on the k sets $S_1,S_2,\dots,S_k$ as the *Lew k-gram.* S_i' denotes the complement of S_i. The Lew 1-gram and Lew 2-gram are depicted in FIGURES 1 and 2, respectively.

In FIGURE 1, the left half is S_1 and the right half is S_1'. In FIGURE 2, starting in the

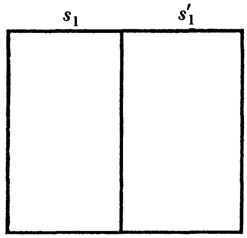


FIGURE 1
Lew 1-gram

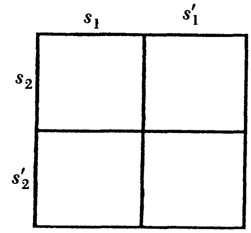


FIGURE 2
Lew 2-gram

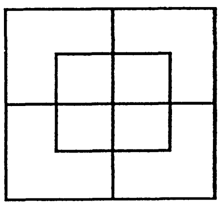


FIGURE 3
Lew 3-gram

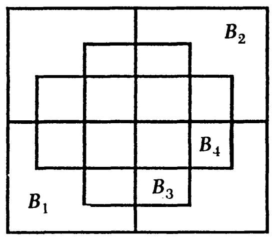


FIGURE 4
Lew 4-gram

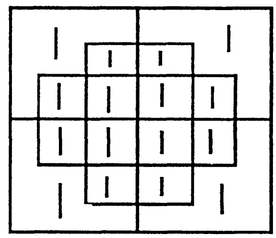


FIGURE 5
Lew 5-gram

upper left corner and proceeding clockwise, these boxes denote $S_1 \cap S_2, S_1' \cap S_2, S_1' \cap S_2', S_1 \cap S_2'$; thus the left half is all of S_1 and the right half is all of S_1' (just as in FIGURE 1). The top half is all of S_2 and the bottom half is all of S_2'.

The Lew k-grams for $k = 3$, 4, and 5 are depicted in FIGURES 3, 4, and 5, respectively.

In FIGURE 3, the left half is all of S_1 and the right half is all of S_1'. The top half is all of S_2 and the bottom half is all of S_2' (just as it is in the Lew 2-gram). Everything inside the interior square is S_3 and everything outside the interior square is S_3'. In FIGURE 4, the left half is all of S_1, the right half is all of S_1', the top half is all of S_2, and the bottom half is all of S_2'. Everything inside the vertical rectangle is S_3 and everything outside the vertical rectangle is S_3' (just as it is in the Lew 3-gram). Additionally, everything inside the horizontal rectangle is S_4 and everything outside the horizontal rectangle is S_4'. To get the Lew 5-gram in FIGURE 5, one just places a vertical line segment in each of the 16 skewed boxes of the Lew 4-gram. Then $S_1, \dots, S_4$ are as they were in the Lew 4-gram, and anything to the left of one of these vertical line segments is in S_5; anything to the right is in S_5'. Notice to get the Lew 5-gram, one simply places a Lew 1-gram in each of the 16 skewed boxes of the Lew 4-gram. FIGURES 6, 7, and 8 depict the Lew k-grams for $k = 6$, 7, and 8, respectively. Each of them is built up from the Lew 4-gram in a manner analogous to the construction of the Lew 5-gram. To get the Lew k-gram for $k = 6$, 7, and 8, simply place a Lew k-gram with $k = 2$, 3, and 4, respectively, in each of the 16 skewed boxes of the Lew 4-gram. We discuss the Lew 8-gram. From this, the interpretations of the Lew 6-gram and Lew 7-gram follow analogously. To interpret the Lew 8-gram, we start with a Lew 4-gram drawn with bold lines. This represents S_1, S_2, S_3, S_4 (and their complements). In each of the 16 skewed boxes, we place a smaller Lew 4-gram drawn with thinner lines. S_5 is everything on the left side of any one of these smaller Lew 4-grams and S_5' is everything on the right side of any one of these smaller Lew 4-grams. S_6 is the top half of all these smaller Lew 4-grams and S_6' is the bottom half of all the smaller Lew 4-grams. Everything inside the vertical rectangle of any one of these smaller Lew 4-grams is S_7; everything outside any one of these vertical rectangles is S'_7. Finally, everything inside the horizontal rectangle of any one of these smaller Lew 4-grams is S_8; everything outside any one of these horizontal rectangles is S'_8. Notice that any intersection of the form $T_1 \cap T_2 \cap \dots \cap T_8$, where $T_i = S_i$ or S'_i is represented by a unique skewed box of the Lew 8-gram and vise versa. For example, $S_1 \cap S'_2 \cap S_3 \cap S'_4 \cap S_5 \cap S'_6 \cap S_7 \cap S_8$ is represented by box E in the lower left of FIGURE 8.

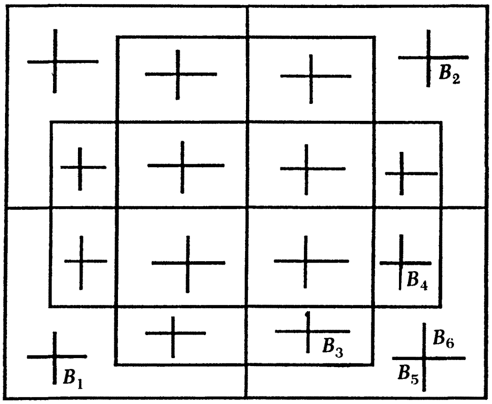


FIGURE 6
Lew 6-gram

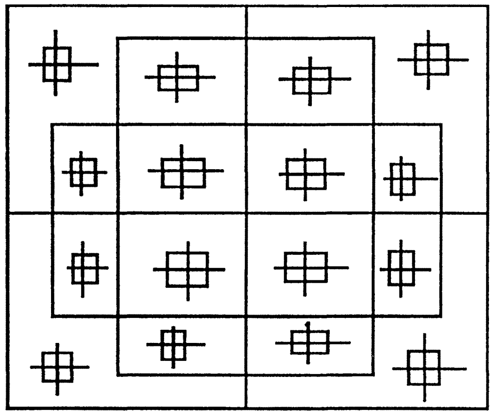


FIGURE 7
Lew 7-gram

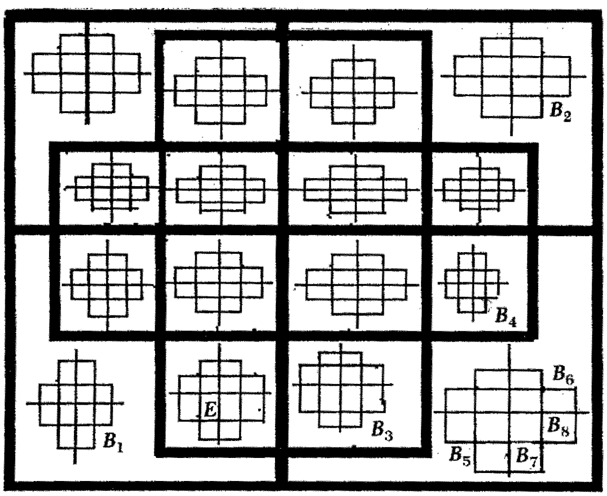


FIGURE 8
Lew 8-gram

By iterating, one can construct a Lew k-gram for any k: If k is a multiple of four, then one constructs the Lew-gram on $k+1$, $k+2$, $k+3$, and $k+4$ many sets by placing a Lew k-gram in the skewed boxes of the Lew 1-gram, Lew 2-gram, Lew 3-gram, and Lew 4-gram, respectively. The interpretations are analogous to those above.

Notice that the Lew k-gram contains 2^k skewed boxes. This observation is clear once one realizes that the skewed boxes of the Lew k-gram are in one-to-one correspondence with all the k-tuples of 0s and 1s. Each skewed box can be labeled by intersections of the form $T_1 \cap T_2 \cap \dots \cap T_k$ where $T_i = S_i$ or S'_i. We let the k-tuple

$(t_1, t_2, \ldots, t_k)$, defined by $t_i = 1$ if $T_i = S_i$ and $t_i = 0$ if $T_i = S_i'$, be a way of coding for the intersection $T_1 \cap T_2 \cap \ldots \cap T_k$. For example, $(1,0,1,0,1,0,1,1)$ is the code for skewed box E in FIGURE 8.

3. Lew k-grams and Ordered k-covers

Let $\tau = (S_1, \ldots, S_k)$ be an ordered k-cover of $[n]$ and consider the Lew k-gram. For each $j \in [n]$, we define $\tau^*(j) = (s_1(j), s_2(j), \ldots, s_k(j))$ where $s_i(j) = 1$ if $j \in S_i$ and $s_i(j) = 0$ if $j \notin S_i$. Thus, $\tau^*(j)$ tells us exactly to which skewed box of the Lew k-gram j belongs. More formally, for each τ, we have a function τ^* between $[n]$ and the skewed boxes of the Lew k-gram. This function can be thought of as the actual placing of the elements of $[n]$ into the Lew k-gram. It is easy to see that if $\tau_1 \neq \tau_2$, then $\tau_1^* \neq \tau_2^*$. From this it follows that different ordered k-covers on $[n]$ correspond to different ways of filling in the skewed boxes of the Lew k-gram with elements of $[n]$. Note, if $\tau = (S_1, \ldots, S_k)$ is an ordered k-cover, then τ^* sends nothing to the skewed box labeled $(0, 0, \ldots, 0)$.

Is there any special way a minimal ordered k-cover of $[n]$ tells us how to fill in the Lew k-gram? Notice the boxes $B_1, B_2, \ldots, B_k$ denoted in FIGURES 8, 9, and 10. For each i, $1 \leq i \leq k$, the k-tuple that has 1 in the ith entry and 0s elsewhere codes for the box B_i. *These boxes are significant because B_i is in S_i, but not in any other S_j with $j \neq i$.* If $\sigma = (S_1, \ldots, S_k)$ is a minimal ordered k-cover on $[n]$, then each of the k distinguished skewed boxes $B_1, B_2, \ldots, B_k$ in the Lew k-gram must receive at least one element of $[n]$. This is because if $B_i = \emptyset$, then S_i can be removed from σ, and what remains is still an ordered $(k-1)$-cover of $[n]$. This observation ultimately allows us to compute $\mu(n, k)$.

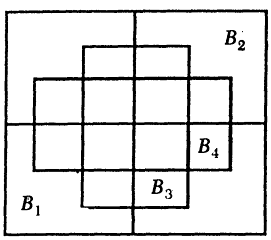


FIGURE 9

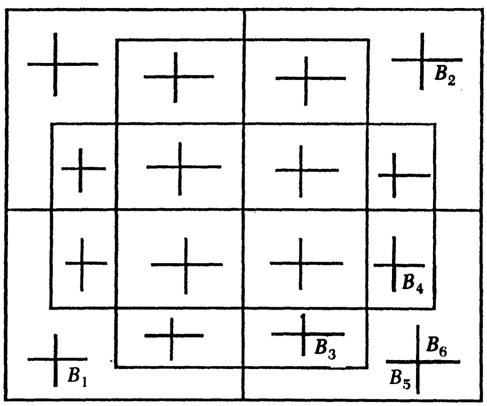


FIGURE 10

Example 1. Let $\sigma = (S_1, \ldots, S_6)$ be given by $S_1 = \{1, 2, 9, \ldots, 12, 15, 16, 17, 20\}$, $S_2 = \{3, 11, 12, 14, 18\}$, $S_3 = \{4, 5, 9, \ldots, 13, 15, \ldots, 18\}$, $S_4 = \{6, 11, \ldots, 17, 20\}$, $S_5 = \{7, \ldots, 11, 13, 15, 16, 17, 20\}$, and $S_6 = \{11, \ldots, 17, 19, 20\}$. Then σ is a minimal 6-cover of $[20]$. FIGURE 12 is the Lew 6-gram which has been filled by σ^*. Notice that each of the distinguished skewed boxes $B_1, \ldots, B_6$ in FIGURE 11 is filled.

Let $\sigma = (S_1, \ldots, S_k)$ be a minimal ordered k-cover of $[n]$. As with any function on $[n]$, the inverse image of the points in the range of the function σ^* defines a partition p_σ on $[n]$. We now view σ^* as a one-to-one function from p_σ to the set of skewed boxes of the Lewis k-gram, that must, by virtue of the minimality of σ, place one member of p_σ in each of the distinguished skewed boxes $B_1, B_2, \dots, B_k$ of the Lew k-gram. Let $P(n)$ denote the partitions of $[n]$ and $SB_k(n)$ denote the skewed boxes of the Lew k-gram. Then there is a bijective correspondence between $OM_k(n)$ and the set of injective functions f with domain$(f) \in P(n)$, codomain$(f) = SB_k(n)$, and the distinguished set of skewed boxes $\{B_1, B_2, \dots, B_k\} \subset$ range(f). *Thus to compute* $|OM_k(n)|$, *we enumerate this set of functions.*

Example 2. Let σ be the minimal 6-cover of **[20]** defined in Example 1. Then $p_\sigma = \{\{1,2\},\{3\},\{4,5\},\{6\},\{7,8\},\{9,10\},\{11\},\{12\},\{13\},\{14\},\{15,16,17\},\{18\},\{19\},\{20\}\}$.

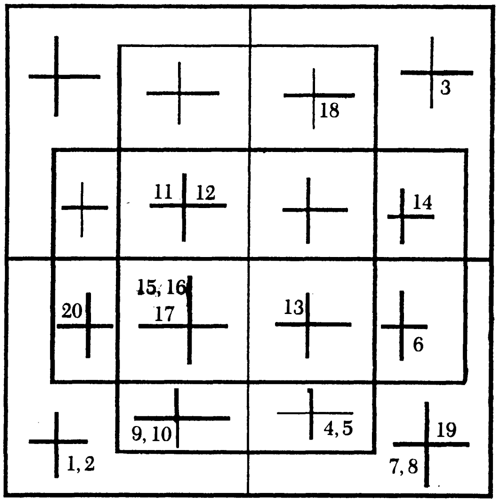


FIGURE 11

Example 3. We count $|OM_4(n)|$. We have $2^4 - 1 = 15$ possible skewed boxes of the Lew 4-gram to fill. (Recall we never fill the skewed box $(0, 0, \dots, 0)$.) We want to count the different ways of filling these skewed boxes with members of some partition p of $[n]$ so that each skewed box gets no more than one member of p, and the distinguished skewed boxes B_1, B_2, B_3, and B_4 get exactly one member. We use the partition length as a parameter. (If a partition has m members, we say it has *length* m.) Since we need to fill the skewed boxes B_1, B_2, B_3, and B_4, we only consider partitions with length at least 4. Letting $\alpha_4 = \min(n, 2^4 - 1)$, it is clear that we can fill at most α_4 of the skewed boxes of the Lew 4-gram. Hence, we only need to consider partitions of length m where $4 \le m \le \alpha_4$. First consider the partitions of length four. Since we must fill the four distinguished boxes, there are $4!S(n,4)$ ways to do this because for each of the $S(n,4)$ partitions of length four, there are 4! ways of making the assignment. Next we consider the partitions of length five. We now fill the four distinguished skewed boxes and an additional one. Since there are $15 - 4 = 11$ skewed boxes from which we may choose the additional one, there are $\binom{11}{1}5!S(n,5)$ ways to do this. Now we consider the partitions of length six. As before, we fill the four distinguished skewed boxes and any two additional skewed boxes. Since there are still $15 - 4 = 11$ skewed boxes from which we may choose the two additional ones, there are $\binom{11}{2}6!S(n,6)$ ways to do this. In general, if we take partitions of length m, where $4 \le m \le \alpha_4$, then there are $\binom{11}{m-4}m!S(n,m)$ ways to fill the skewed boxes of the Lew 4-gram so that a minimal ordered 4-cover results. Summing over our parameter m, we have

TABLE 1. $S(n,k)$

n/k	1	2	3	4	5	6	7
1	1						
2	1	1					
3	1	3	1				
4	1	7	6	1			
5	1	15	25	10	1		
6	1	31	90	65	15	1	
7	1	63	301	350	140	21	1

$$|OM_4(n)| = \sum_{m=4}^{\alpha_4} \binom{11}{m-4} m!S(n,m),$$

and

$$\mu(n,4) = \frac{1}{4!}\sum_{m=4}^{\alpha_4} \binom{11}{m-4} m!S(n,m).$$

In general, to enumerate $OM_k(n)$, one begins with the Lew k-gram. Letting $\alpha_k = \min(n, 2^k - 1)$, one then considers the partitions of length m, with the parameter m ranging over all values $k \le m \le \alpha_k$. With m fixed, one then fills the distinguished $B_1, B_2, \dots, B_k$ skewed boxes and $m-k$ additional skewed boxes in the Lew k-gram with members of a partition of length m. Since there are $2^k - k - 1$ skewed boxes in addition to the boxes $B_1, B_2, \dots, B_k$, there are $\binom{2^k-k-1}{m-k} m!S(n,m)$ ways of doing this. Therefore

$$|OM_k(n)| = \sum_{m=k}^{\alpha_k} \binom{2^k-k-1}{m-k} m!S(n,m),$$

and

$$\mu(n,k) = \frac{1}{k!}\sum_{m=k}^{\alpha_4} \binom{2^k-k-1}{m-k} m!S(n,m). \tag{1}$$

From Tables 1 and 2 notice that $S(n+1,3) = \mu(n,2)$. This can be verified by invoking the recursive definition for $S(n,k)$ and using (1) to get that $\mu(n,2) = S(n,2) + 3S(n,3)$. One question that we cannot answer is: *Can $\mu(n,k)$ be defined recursively in terms of $\mu(n-1,k-1)$ and $\mu(n-1,k)$?*

TABLE 2. $\mu(n,k)$

n/k	1	2	3	4	5	6	7
1	1						
2	1	1					
3	1	6	1				
4	1	25	22	1			
5	1	90	305	65	1		
6	1	301	3410	2540	171	1	
7	1	966	33621	77350	17066	420	1

REFERENCES

1. Ian Anderson, *A First Course in Combinatorial Mathematics*, 2nd edition, Oxford University Press, Oxford, 1990.
2. Lewis Carroll, *The Mathematical Recreations of Lewis Carroll*, Dover Publications, Mineola, NY, 1958.
3. Louis Comtet, *Advanced Combinatorics*, D. Reidel, Dordrecht, 1974.
4. T. Hearne and C. Wagner, Minimal covers of finite sets, *Disc. Math.* 5 (1973) 247–251.

NOTES

The Golden Section and the Piano Sonatas of Mozart

JOHN F. PUTZ
Alma College
Alma, MI 48801

Ubiquitous in nature, the golden section embodies its elegant proportion in the starfish and the chambered nautilus, in the pine cone and the sunflower, and in leaf patterns along the stems of plants [**11**, pp. 98, 113–114; **12**, pp. 3–14; **31**, pp. 150–169]. Perhaps it is because the golden section is in some sense natural, that artists, architects, and composers have often been influenced by it [**14**, **15**, **21**, **22**]. And perhaps this is to be expected, even when not deliberate, insofar as art imitates nature.

The *golden section* is defined to be that division of a line segment into two unequal segments such that the length a of the shorter segment is to the length b of the longer, as the length of the longer is to the whole. See FIGURE 1. That is,

$$\frac{a}{b} = \frac{b}{a+b}.$$

For convenience, we let the segment to be divided have length 1 and the shorter segment have length x as in FIGURE 2. Then the golden section is that division of the segment for which

$$\frac{x}{1-x} = \frac{1-x}{1}.$$

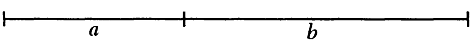


FIGURE 1
The golden section.

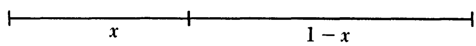


FIGURE 2

Solving this equation and setting φ equal to the common ratio, we have

$$\varphi = \frac{x}{1-x} = \frac{1-x}{1} = \frac{\sqrt{5}-1}{2} \approx 0.6180.$$

This ratio (or its reciprocal) is variously called the "golden ratio," "golden number," or "divine proportion," and is sometimes said to offer the most aesthetically pleasing proportion [**4**, p. 94; **16**, pp. 62–65; **36**, p. 74]. Whether or not this is accurate, the effect of identical ratios between the parts and between the parts and the whole is to unify the structure in a fundamental way.

At an age in excess of 24 centuries [**32**, p. 291], the golden section has become the subject of modern debate. In [**23**], for instance, G. Markowsky debunks some popularly held beliefs, such as, that the Parthenon, United Nations Building, and Great Pyramid conform to the golden section, and that rectangles with golden ratio proportions are the most pleasing. On the other hand, J. Benjafield and J. Adams-Webber [**2**] have advanced the "golden section hypothesis" that whenever people must divide a whole into two unequal parts, they tend to make the division near the golden section. In this paper, we take up the particular case of the piano sonatas of W. A. Mozart (1756–1791). Some have said that these works do reflect the golden section [**7**; **35**, p. 242]. Here, we analyze some collected data to judge whether there is convincing evidence to support this claim.

Even a listener who is only casually acquainted with the music of Mozart will hear something familiar in it; from Mozart's mind came melody not only delightful but memorable. On another level, the genius of the composer is manifested in form and balance. His music has been revered, among other things, for its "beautiful and symmetrical proportions" [**34**, p. 217]. In 1853, Henri Amiel opined that "the balance of the whole is perfect" [**1**, p. 54]. Hanns Dennerlein described Mozart's music as reflecting the "most exalted proportions," and the composer himself as having "an inborn sense for proportions" [quoted in **7**, p. 1], a thought echoed by H. C. Robbins Landon [**20**, p. 268]. Eric Blom wrote that Mozart had "an infallible taste for saying exactly the right thing at the right time and at the right length" [**5**, p. 265].

Music and mathematics having been happily entwined from antiquity, it is not surprising when talent in one accompanies enthusiasm for the other. Mozart's sister, Nannerl, recalled that when her brother was learning arithmetic, he gave himself entirely to it and that "he talked of nothing, thought of nothing but figures" [**19**, p. 124]. She recalled that he once covered the walls of the staircase and of all the rooms in their house with figures, then moved on to do the neighbors' houses as well. When he was 14, Mozart wrote to her asking that she send him arithmetical tables and more exercises in arithmetic [**25**, letter of April 21, 1770, p. 130, and letter of May 19, 1770, p. 137]. The margins of the manuscript of the *Fantasia and Fugue in C major* contain Mozart's calculations of the probability of winning a lottery [**18**, p. 178]. Alfred Einstein, one of Mozart's biographers wrote: "The pleasure of playing with figures remained with Mozart all his life long. Thus he once took up the problem, very popular at the time, of composing minuets 'mechanically,' by putting two-measure melodic fragments together in any order. And we possess a page of musical sketches on which he had begun to figure out the sum which the chess player would have received from the King in the famous Oriental story" [**9**, p. 25].

By the age of 18, Mozart had composed his first sonata for piano [**17**, p. 42; **29**, p. 45]. He wrote 19 altogether [**6**, pp. vi–vii], most of them during the next four years of his life, and almost all of them comprising three movements. In Mozart's time, the sonata-form movement was conceived in two parts [**26**, pp. 30–35; **28**, pp. 160, 163; **30**, pp. 1–2; **33**, pp. 14–15]: the Exposition in which the musical theme is introduced, and the Development and Recapitulation in which the theme is developed and revisited. See FIGURE 3. As a rule, each section was to be repeated (as indicated by the symbol :‖) in performance [**24**, p. xxxiii]. It is this separation into two distinct sections, together with the foregoing, which gives cause to wonder how Mozart apportioned these works.

FIGURE 3
Sonata-form movement.

Table 1 is a collection of data for all Mozart's sonata movements that are divided into two distinct sections, both of which are to be repeated in performance. Of the 56 movements, 29 are constructed in this way. The data are measure counts (taken from

[6]) of the lengths of the two sections: a represents the length of the Exposition, and b the length of the Development and Recapitulation. (When codas were present, they were not included as part of the second section.) The first column identifies the piece and movement by the Köchel cataloging system. K. 498a was not included because its authenticity is in doubt [**6**, p. vi; **29**, p. 56]. The first movement of the first sonata, K. 279, is 100 measures in length and is divided so that the Development and Recapitulation section has length 62. Note that 100φ, rounded to the nearest natural number, is 62. (These lengths are necessarily natural numbers because they are measure counts.) This is a perfect division according to the golden section in the following sense: A 100-measure movement could not be divided any closer (in natural numbers) to the golden section than 38 and 62. This is true of the second movement of this sonata as well. That is, a 74-measure movement cannot be divided any closer to the golden section than 28 and 46. Mozart did not, however, divide the third movement of K. 279

TABLE 1

Köchel	a	b	$a+b$
279, I	38	62	100
279, II	28	46	74
279, III	56	102	158
280, I	56	88	144
280, II	24	36	60
280, III	77	113	190
281, I	40	69	109
281, II	46	60	106
282, I	15	18	33
282, III	39	63	102
283, I	53	67	120
283, II	14	23	37
283, III	102	171	273
284, I	51	76	127
309, I	58	97	155
311, I	39	73	112
310, I	49	84	133
330, I	58	92	150
330, III	68	103	171
332, I	93	136	229
332, III	90	155	245
333, I	63	102	165
333, II	31	50	81
457, I	74	93	167
533, I	102	137	239
533, II	46	76	122
545, I	28	45	73
547a, I	78	118	196
570, I	79	130	209

exactly in golden section. An exact division would require b to be 98, not 102.

To evaluate the degree of consistency in these proportions, we use a scatter plot of b against $a+b$. If Mozart was consistent, there should be a linearity to the data, and if he divided movements near to the golden section, then the data points should fall near the line $y=\varphi x$. FIGURE 4 shows this scatter plot. Certainly, the linearity in the data is striking. The r^2 value is 0.990, confirming an extremely high degree of linearity. To this plot, we add two lines (FIGURE 5): the line $y=\varphi x$, and the regression line whose equation is $y=-0.003241+0.6091x$. The line $y=\varphi x$ scarcely differs from the line of best fit and, at the scale of FIGURE 5, it is difficult to see any difference at all between them. Of course, the line $y = \varphi x$ is a little above the regresssion line because of its slightly larger slope. Finally, a histogram (FIGURE 6) of the ratio $b/(a+b)$ reflects the centrality of φ.

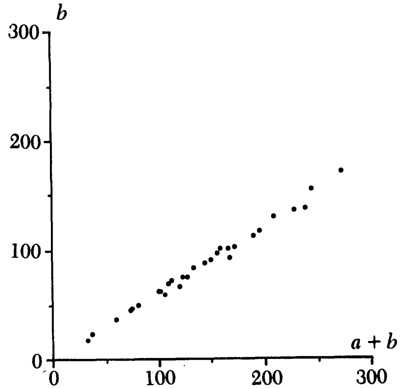


FIGURE 4

Scatter plot of b against $a+b$.

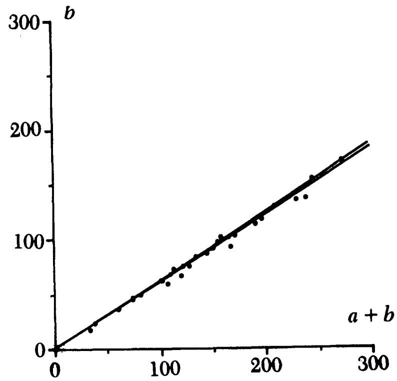


FIGURE 5

Scatter plot of b and $a+b$ with the line $y = \varphi x$ (top) and the regression line (bottom).

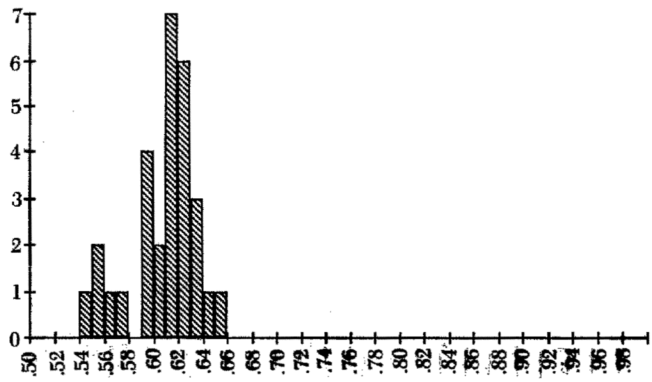


FIGURE 6

Frequency distribution of $\frac{b}{a+b}$.

This is impressive evidence that Mozart did, with considerable consistency, partition sonata movements near the golden section. Before we become convinced, however, let us analyze these data in another way. If a movement is divided in golden section, then both a/b and $b/(a+b)$ should be near φ. We have focused on $b/(a+b)$; let us concentrate now on a/b. FIGURE 7 is a scatter plot of a against b. Again, the data look very linear, though not so much so as b and $a+b$. To this plot, we again add the line $y = \varphi x$ and the regression line whose equation, this time, is $y = 1.360 + 0.6260x$. See FIGURE 8. The line of best fit, which is the one on the top in FIGURE 8, and the line $y = \varphi x$ again differ very little. The r^2 value of 0.938 verifies somewhat less goodness of fit. A histogram (FIGURE 9), however, is more revealing, showing much more variance than FIGURE 6 and, therefore, less evidence for the centrality of φ.

Now, why is this? It is possible, of course, to put a slant on the interpretation of data by the choice of presentation, but there seems to be something other than that going on here. And there is: It is a theorem **[10]** that what we have observed in these data is true for all data; $b/(a+b)$ is always nearer to φ than is a/b.

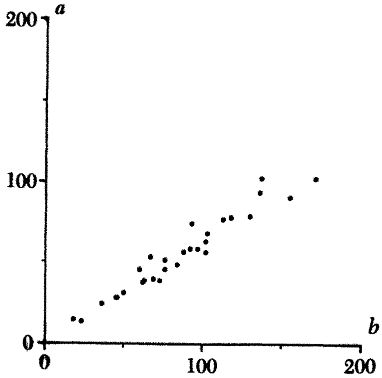


FIGURE 7
Scatter plot of a against b.

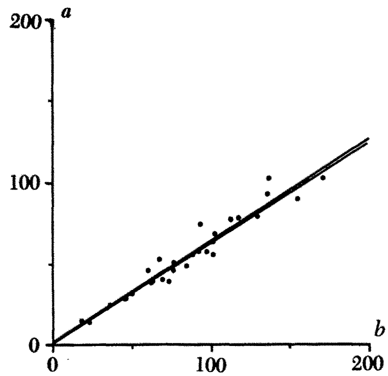


FIGURE 8
Scatter plot of a against b with the line $y = \varphi x$ (bottom) and the regression line (top).

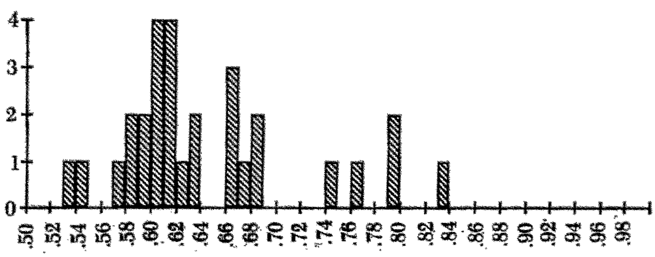


FIGURE 9
Frequency distribution of $\dfrac{a}{b}$.

THEOREM. $\left|\dfrac{b}{a+b} - \varphi\right| \le \left|\dfrac{a}{b} - \varphi\right|$ *where* $0 \le a \le b$.

Proof. Let $x = a/b$. Then we must show that

$$\left|\frac{1}{x+1} - \varphi\right| \le |x - \varphi|$$

for all $x \in [0, 1]$. Let $f(x) = 1/(x+1)$. By the Mean Value Theorem, for all $x \in [0, 1]$ there is a $\xi \in (0, 1)$ such that

$$|f(x) - f(\varphi)| = |f'(\xi)|\,|x - \varphi|.$$

Now $f'(x) = -1/(x+1)^2$ satisfies

$$\tfrac{1}{4} < |f'(x)| < 1$$

for $x \in (0, 1)$. A simple calculation will show that φ is a fixed point of f, that is, that $f(\varphi) = \varphi$. So, for all $x \in [0, 1]$,

$$\left|\frac{1}{x+1} - \varphi\right| \ge |x - \varphi|$$

with equality when $x = \varphi$.

We note, in passing, that the fixed-point algorithm of numerical analysis works on this principle, and that this theorem says that the ratio of consecutive terms of any Fibonacci-like sequence ($f_1 = a, f_2 = b, f_{n+2} = f_n + f_{n+1}$ with a and b not both zero) converges to φ.

Thus we know that, given *any* pair a and b, $0 \leqslant a \leqslant b$, the ratio $b/(a+b)$ will be closer to φ than a/b will. An enthusiast wishing to demonstrate a golden ratio relationship between, say, shoe size s and ACT score t, should present data in the form $t/(s+t)$ instead of s/t, because $t/(s+t)$ will be biased toward φ. Table 2 shows some data to demonstrate this. As the theorem predicts, in every case $t/(s+t)$ is nearer to 0.6180 than s/t is. One lesson to be learned from this is clear: If we are to analyze data in this way, then we must confine our investigations to the ratio a/b. Otherwise, we are likely to find φ even when it is not there. Unfortunately, some investigations have focused on the ratios $b/(a+b)$ [**2**, **3**, **8**].

TABLE 2

Shoe size s	ACT score t	$\frac{t}{s+t}$	$\frac{s}{t}$
8	26	0.7647	0.3077
10	22	0.6875	0.4545
9	28	0.7568	0.3214
12	25	0.6757	0.4800

Drawing our attention, then, to a/b and following the ideas in [**10**], let us ask what values we might reasonably expect the ratio to have. It would be absurd to think that any composer, at least in the classical period, would write, for example, a 200-measure sonata movement and divide it into two parts so lopsidedly as 1 and 199, or 2 and 198, or even 10 and 190. There simply would not be enough room in that to accomplish the purpose of the first section: the exposition of the theme. Quantz suggests some balance in the relationship in saying that "the first part must be somewhat shorter than the second" [**27**, p. 591]. So if we let the length of the movement $m = a + b$ be fixed, then a must be bounded below at some practical distance away from 0, and bounded above by $m/2$. For the moment, let us suppose that $m/4 \le a \le m/2$. This interval satisfies the conditions at least and has the appeal of simplicity. If we assume that a is randomly distributed, then an estimate of the expected value of a/b is

$$\begin{aligned} E(a/b) &\approx \frac{1}{m/4}\int_{m/4}^{m/2} \frac{x}{m-x}\,dx \\ &= \frac{4}{m}(x + m\ln|x-m|)\Big]_{m/2}^{m/4} \\ &= 4\ln\frac{3}{2} - 1 \\ &\approx 0.6219. \end{aligned}$$

This estimate differs from φ by about 0.6%. Of course, infinitely many other intervals also conform to the assumptions, and the expected values vary widely. For example, $0.3m \leqslant a \le 0.4m$ gives $E(a/b) \approx 0.5415$. The data in Table 1 satisfy $0.348m \le a \le 0.455m$. Using this interval, $E(a/b) \approx 0.6753$. On the other hand, intervals satisfying the conditions can be chosen so that the expected value is exactly φ. One such interval is $[rm, (r + 1/5)m]$ where

$$r = \frac{1 - (4/5)e^{(\varphi+1)/5}}{1 - e^{(\varphi+1)/5}}.$$

The point is this: The sonata form itself imposes restrictions. Depending on the assumptions we make about a to conform to them, these restrictions can induce on

a/b central tendency in the vicinity of φ and, in some cases, very near φ. And this is true even when a is determined, not by thoughtful design, but by uninspired randomness.

Still, we must remember that these sonatas *are* the work of a genius, and one who loved to play with numbers. Mozart may have known of the golden section and used it. That there is considerable deviation from it (FIGURE 9) suggests otherwise, however. Perhaps the golden section does, indeed, represent the most pleasing proportion, and perhaps Mozart, through his consummate sense of form, gravitated to it as the perfect balance between extremes. It is a romantic thought.

> *The "music" of nature and the music of man belong to two distinct categories. The transition from the former to the latter passes through the science of mathematics. An important and pregnant proposition. Still, we should be wrong were we to construe it in the sense that man framed his musical system according to calculations purposely made, the system having arisen through the unconscious application of pre-existent conceptions of quantity and proportion, through subtle processes of measuring and counting; but the laws by which the latter are governed were demonstrated only subsequently by science.*
>
> —Eduard Hanslick, 1854 [**13**, p. 110]

REFERENCES

1. H. F. Amiel, *Amiel's Journal*, 2nd edition, Bretano's, New York, 1928.
2. J. Benjafield and J. Adams-Webber, The golden section hypothesis, *British Journal of Psychology* 67 (1976) 11–15.
3. J. Benjafield and C. Davis, The golden number and the structure of connotation, *Journal of Aesthetics and Art Criticism* 36 (1978) 423–427.
4. David Bergamini, *Mathematics*, Time-Life Books, New York, 1972.
5. Eric Blom, *Mozart*, J. M. Dent and Sons Ltd., London, 1962.
6. Nathan Broder, *Mozart: Sonatas and Fantasies for the Piano*, revised ed., Theodore Presser, Bryn Mawr, PA, 1960.
7. E. J. P. Camp, Temporal proportion: A study of sonata forms in the piano sonatas of Mozart, Ph.D. dissertation, Florida State University, 1968.
8. G. Duckworth, *Structural Patterns and Proportions in Vergil's Aeneid*, University of Michigan Press, Ann Arbor, 1962.
9. Alfred Einstein, *Mozart: His Character, His Work*, translated by Arthur Mendel and Nathan Broder, Oxford University Press, London, 1945.
10. Roger Fischler, How to find the "golden number" without really trying, *Fibonacci Quarterly* 19 (1981) 406–410.
11. Matila Ghyka, The Pythagorean and Platonic scientific criterion of the beautiful in classical western art, in *Ideological Differences and World Order: Studies in the Philosophy and Science of the World's Cultures*, edited by F. S. C. Northrop, Yale University Press, New Haven, CT, 1949.
12. Jay Hambidge, *The Elements of Dynamic Symmetry*, Yale University Press, New Haven, CT, 1959.
13. Eduard Hanslick, *The Beautiful in Music*, translated by Gustav Cohen, Liberal Arts Press, Indianapolis, 1957.
14. H. Hedian, The golden section and the artist, *Fibonacci Quarterly* 14 (1976) 406–418.
15. Roy Howat, *Debussy in Proportion: A Musical Analysis*, Cambridge University Press, Cambridge, 1983.
16. H. E. Huntley, *The Divine Proportion*, Dover Publications, Mineola, NY, 1970.
17. Arthur Hutchings, The keyboard music, in *The Mozart Companion*, edited by H. C. Robbins Landon and Donald Mitchell, W. W. Norton, New York, 1969.
18. A. Hyatt King, Mozart's piano music, *The Music Review* 5 (1944) 163–191.
19. A. Hyatt King, *Mozart in Retrospect*, Greenwood Press, Westport, CT, 1976.
20. H. C. Robbins Landon, The concertos: (2) Their musical origin and development, in *The Mozart Companion*, edited by H. C. Robbins Landon and Donald Mitchell, W. W. Norton, New York, 1969.

21. Le Corbusier, Le modulor, in *Le Corbusier, Œuvre Complète 1946–1952*, 3rd edition, Editions Girsberger, Zurich, 1961.
22. Ernö Lendvai, *Béla Bartók: An Analysis of His Music*, Kahn & Averill, London, 1971.
23. George Markowsky, Misconceptions about the golden ratio, *Col. Math. J.* 23 (1992) 2–19.
24. F. Helena Marks, *The Sonata, Its Form and Meaning as Exemplified in the Piano Sonatas by Mozart*, William Reeves, London, 1921.
25. Wolfgang A. Mozart *et al.*, *The Letters of Mozart and His Family, Vol. I*, 2nd edition, translated by Emily Anderson, Macmillan, London, 1966.
26. William S. Newman, *The Sonata in the Classic Era*, University of North Carolina Press, Chapel Hill, 1963.
27. Johann Joachim Quantz, From the *Versuch einer Anweisung die Flöte traversiere zu spielen*, in *Source Readings in Music History*, edited by Oliver Strunk, W. W. Norton, New York, 1950.
28. Leonard Ratner, Harmonic aspects of classic form, *J of the Amer Musicological Soc* 2 (1949) 159–168.
29. Thomas Richner, *Interpreting Mozart's Piano Sonatas*, Paterson's Publications, London, 1978.
30. Charles Rosen, *Sonata Forms*, W. W. Norton, New York, 1980.
31. Edmund W. Sinnott, *Plant Morphogenesis*, McGraw-Hill, New York, 1960.
32. David Eugene Smith, *History of Mathematics, Vol. II Special Topics of Elementary Mathematics*, Ginn and Co., Boston, 1925.
33. J. Raymond Tobin, *Mozart and the Sonata Form*, Da Capo Press, New York, 1971.
34. Donald Francis Tovey, *The Forms of Music*, Meridian Books, New York, 1957.
35. J. H. Douglas Webster, Golden-mean form in music, *Music & Letters* 31 (1950) 238–248.
36. Margaret F. Willerding, *Mathematical Concepts: A Historical Approach, Vol. 5*, Prindle, Weber & Schmidt, Boston, 1967.

Can a Mathematician See Red?

Consider the sphere—
a hollow rounded surface
with no thickness.
Each point that we see
from the outside
is also a point we can see
from the inside.

If I paint red
all over the outside,
is the inside red?

The mathematician says NO,
for the layer of paint
forms a new sphere
that is outside the outside
and not a bit inside.

A mathematician
takes safe pleasure
in surface mysteries.

A poet
will see red
inside.

—JoANNE GROWNEY
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
BLOOMSBURG UNIVERSITY, BLOOMSBURG, PA 17815

21. Le Corbusier, Le modulor, in *Le Corbusier, Œuvre Complète 1946–1952*, 3rd edition, Editions Girsberger, Zurich, 1961.
22. Ernö Lendvai, *Béla Bartók: An Analysis of His Music*, Kahn & Averill, London, 1971.
23. George Markowsky, Misconceptions about the golden ratio, *Col. Math. J.* 23 (1992) 2–19.
24. F. Helena Marks, *The Sonata, Its Form and Meaning as Exemplified in the Piano Sonatas by Mozart*, William Reeves, London, 1921.
25. Wolfgang A. Mozart *et al.*, *The Letters of Mozart and His Family, Vol. I*, 2nd edition, translated by Emily Anderson, Macmillan, London, 1966.
26. William S. Newman, *The Sonata in the Classic Era*, University of North Carolina Press, Chapel Hill, 1963.
27. Johann Joachim Quantz, From the *Versuch einer Anweisung die Flöte traversiere zu spielen*, in *Source Readings in Music History*, edited by Oliver Strunk, W. W. Norton, New York, 1950.
28. Leonard Ratner, Harmonic aspects of classic form, *J of the Amer Musicological Soc* 2 (1949) 159–168.
29. Thomas Richner, *Interpreting Mozart's Piano Sonatas*, Paterson's Publications, London, 1978.
30. Charles Rosen, *Sonata Forms*, W. W. Norton, New York, 1980.
31. Edmund W. Sinnott, *Plant Morphogenesis*, McGraw-Hill, New York, 1960.
32. David Eugene Smith, *History of Mathematics, Vol. II Special Topics of Elementary Mathematics*, Ginn and Co., Boston, 1925.
33. J. Raymond Tobin, *Mozart and the Sonata Form*, Da Capo Press, New York, 1971.
34. Donald Francis Tovey, *The Forms of Music*, Meridian Books, New York, 1957.
35. J. H. Douglas Webster, Golden-mean form in music, *Music & Letters* 31 (1950) 238–248.
36. Margaret F. Willerding, *Mathematical Concepts: A Historical Approach, Vol. 5*, Prindle, Weber & Schmidt, Boston, 1967.

Can a Mathematician See Red?

Consider the sphere—
a hollow rounded surface
with no thickness.
Each point that we see
from the outside
is also a point we can see
from the inside.

If I paint red
all over the outside,
is the inside red?

The mathematician says NO,
for the layer of paint
forms a new sphere
that is outside the outside
and not a bit inside.

A mathematician
takes safe pleasure
in surface mysteries.

A poet
will see red
inside.

—JoAnne Growney
Department of Mathematics and Computer Science
Bloomsburg University, Bloomsburg, PA 17815

The Phi Number System Revisited

CECIL ROUSSEAU
University of Memphis
Memphis, TN 38152

1. Introduction In the November-December 1957 issue of this MAGAZINE, George Bergman, then 12 years old, published an article on a number system in which the base is the golden ratio ($\phi = (1+\sqrt{5})/2$) [**1**]. Since then, the *phi number system* has become part of the folklore of elementary mathematics and has, for example, appeared in an exercise in Knuth's *The Art of Computer Programming* [**2**, problem 35, p. 85]. One of the basic results involving the phi number system is that every positive integer has a finite expansion.

THEOREM. *For every positive integer n, there is a corresponding finite sequence of distinct integers* $k_1, k_2, \ldots, k_m$ *such that* $n = \phi^{k_1} + \phi^{k_2} + \cdots + \phi^{k_m}$.

In his article, Bergman showed that every positive real number has an expansion of the form $\phi^{k_1} + \phi^{k_2} + \phi^{k_3} + \cdots$, where $k_1 > k_2 > k_3 > \cdots$ and thus every such number has a "base ϕ" representation using digits 0 and 1. Moreover, this representation is unique if no two consecutive digits are 1's and the expansion does not become "01010101$\cdots$" after some finite number of digits. In this note, we prove the theorem about finite expansions for positive integers (and a little bit more) using some common notions from algebraic number theory.

2. Proof Let $\mathbb{Z}[\phi] = \{a + b\phi | a, b \in \mathbb{Z}\}$. The *conjugate* of $\alpha = a + b\phi \in \mathbb{Z}[\phi]$ is $\bar{\alpha} = a - b/\phi$ and the *norm* is $N(\alpha) = \alpha\bar{\alpha} = a^2 + ab - b^2$. Note that $\bar{\alpha} = \alpha - b\sqrt{5}$, so

$$N(\alpha) = \alpha(a - b/\phi) = \alpha(\alpha - b\sqrt{5}). \tag{1}$$

We shall use the fact that $N(\alpha\beta) = N(\alpha)N(\beta)$. Clearly $N(\phi) = N(\phi^{-1}) = -1$, and hence $|N(\phi^k)| = 1$ for every $k \in \mathbb{Z}$. Also $N(\alpha) = 0$ if, and only if, $\alpha = 0$. For an introduction to algebraic number theory and uses of the norm, see [**3**, chapter 5].

In what follows, we shall regard $\mathbb{Z}[\phi]$ as a subset of $\mathbb{R}$, ordered in the usual way. We shall prove that every positive number in $\mathbb{Z}[\phi]$ can be expressed as the sum of a finite number of distinct integral powers of ϕ. This can be proved in other, more elementary, ways, but it is of some interest to note that the norm given by (1) provides a natural vehicle for obtaining the expansion and establishing that it is finite. One key fact is needed.

LEMMA. *If* $\alpha \in \mathbb{Z}[\phi]$ *satisfies* $1 \le \alpha < \sqrt{5}$, *then* $|N(\alpha)| > |N(\alpha - 1)|$.

Proof. Note that for any element $\gamma = c + d\phi \in \mathbb{Z}[\phi]$ satisfying $0 \le \gamma < \sqrt{5}$, (1) yields

$$|N(\gamma)| = \gamma|c - d/\phi| = \gamma|\gamma - d\sqrt{5}| = \begin{cases} \gamma(d\sqrt{5} - \gamma) & \text{if } d \ge 1, \\ \gamma(\gamma - d\sqrt{5}) & \text{if } d < 1. \end{cases} \tag{2}$$

Given that $\alpha = a + b\phi$ satisfies $1 \le \alpha < \sqrt{5}$, we can apply (2) in computing both $|N(\alpha)|$ and $|N(\alpha - 1)|$. If $b \ge 1$, then $b\sqrt{5} > 2\alpha - 1$ and

$$|N(\alpha)| = \alpha(b\sqrt{5} - \alpha) > (\alpha - 1)(b\sqrt{5} - (\alpha - 1)) = |N(\alpha - 1)|.$$

Similarly, $b < 1$ yields $b\sqrt{5} < 2\alpha - 1$ and

$$|N(\alpha)| = \alpha(\alpha - b\sqrt{5}) > (\alpha - 1)((\alpha - 1) - b\sqrt{5}) = |N(\alpha - 1)|.$$

Hence $|N(\alpha)| > |N(\alpha - 1)|$ holds in all cases.

The proof that every positive number in $\mathbb{Z}[\phi]$ can be expressed as a finite sum of integral powers of ϕ now proceeds by the greedy method. Let σ be a positive number in $\mathbb{Z}[\phi]$. Choose $k \in \mathbb{Z}$ so that $\phi^k \le \sigma < \phi^{k+1}$ and set $\alpha = \phi^{-k}\sigma$. Then $1 \le \alpha < \phi < \sqrt{5}$ and by the lemma,

$$|N(\sigma)| = |N(\alpha)| > |N(\alpha - 1)| = |N(\sigma - \phi^k)|.$$

If $\sigma > \phi^k$, repeat the process. Thus we obtain $\sigma = \phi^{k_1} + \phi^{k_2} + \cdots + \phi^{k_m}$ after a finite number of steps since $|N(\alpha)|$ is a positive integer for all $\alpha \neq 0$, and

$$|N(\sigma)| > |N(\sigma - \phi^{k_1})| > |N(\sigma - \phi^{k_1} - \phi^{k_2})| > \cdots \ge 0.$$

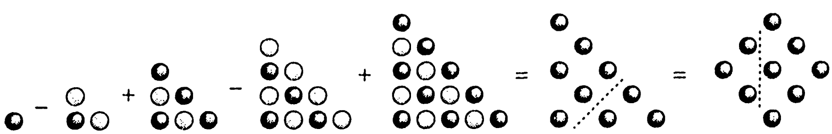
The integers $k_1, k_2, \ldots, k_m$ are distinct; indeed, $k_i - k_{i+1} \ge 2$ for $i = 1, 2, \ldots, m-1$ since $\phi^k \le \sigma < \phi^{k+1}$ implies $0 \le \sigma - \phi^k < \phi^{k-1}$. Thus, according to the result mentioned earlier, this procedure generates the unique "base ϕ" expansion in which no two consecutive digits are 1's.

REFERENCES

1. G. Bergman, A number system with an irrational base, this MAGAZINE, 31 (1957), 98–110.
2. D. Knuth, *Fundamental Algorithms, Vol. 1 of The Art of Computer Programming*, Addison-Wesley, Reading, MA, 1973.
3. H. E. Rose, *A Course in Number Theory*, Clarendon Press, Oxford, 1988.

Proof without Words: Alternating Sums of Triangular Numbers

$$T_k = 1 + 2 + \cdots + k \Rightarrow \sum_{k=1}^{2n-1} (-1)^{k+1} T_k = n^2$$



NOTE: For "proofs without words" of a dual statement—alternating sums of squares are triangular numbers—see this MAGAZINE, December 1987, p. 291; and April 1992, p. 90.

—ROGER B. NELSEN
LEWIS AND CLARK COLLEGE
PORTLAND, OR 92719

$$|N(\alpha)| = \alpha(b\sqrt{5} - \alpha) > (\alpha - 1)(b\sqrt{5} - (\alpha - 1)) = |N(\alpha - 1)|.$$

Similarly, $b < 1$ yields $b\sqrt{5} < 2\alpha - 1$ and

$$|N(\alpha)| = \alpha(\alpha - b\sqrt{5}) > (\alpha - 1)((\alpha - 1) - b\sqrt{5}) = |N(\alpha - 1)|.$$

Hence $|N(\alpha)| > |N(\alpha - 1)|$ holds in all cases.

The proof that every positive number in $\mathbb{Z}[\phi]$ can be expressed as a finite sum of integral powers of ϕ now proceeds by the greedy method. Let σ be a positive number in $\mathbb{Z}[\phi]$. Choose $k \in \mathbb{Z}$ so that $\phi^k \le \sigma < \phi^{k+1}$ and set $\alpha = \phi^{-k}\sigma$. Then $1 \le \alpha < \phi < \sqrt{5}$ and by the lemma,

$$|N(\sigma)| = |N(\alpha)| > |N(\alpha - 1)| = |N(\sigma - \phi^k)|.$$

If $\sigma > \phi^k$, repeat the process. Thus we obtain $\sigma = \phi^{k_1} + \phi^{k_2} + \cdots + \phi^{k_m}$ after a finite number of steps since $|N(\alpha)|$ is a positive integer for all $\alpha \ne 0$, and

$$|N(\sigma)| > |N(\sigma - \phi^{k_1})| > |N(\sigma - \phi^{k_1} - \phi^{k_2})| > \cdots \ge 0.$$

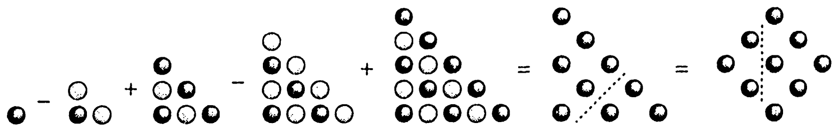
The integers $k_1, k_2, \ldots, k_m$ are distinct; indeed, $k_i - k_{i+1} \ge 2$ for $i = 1, 2, \ldots, m-1$ since $\phi^k \le \sigma < \phi^{k+1}$ implies $0 \le \sigma - \phi^k < \phi^{k-1}$. Thus, according to the result mentioned earlier, this procedure generates the unique "base ϕ" expansion in which no two consecutive digits are 1's.

REFERENCES

1. G. Bergman, A number system with an irrational base, this MAGAZINE, 31 (1957), 98–110.
2. D. Knuth, *Fundamental Algorithms, Vol. 1 of The Art of Computer Programming*, Addison-Wesley, Reading, MA, 1973.
3. H. E. Rose, *A Course in Number Theory*, Clarendon Press, Oxford, 1988.

Proof without Words: Alternating Sums of Triangular Numbers

$$T_k = 1 + 2 + \cdots + k \Rightarrow \sum_{k=1}^{2n-1} (-1)^{k+1} T_k = n^2$$



NOTE: For "proofs without words" of a dual statement—alternating sums of squares are triangular numbers—see this MAGAZINE, December 1987, p. 291; and April 1992, p. 90.

—ROGER B. NELSEN
LEWIS AND CLARK COLLEGE
PORTLAND, OR 92719

A Note on the Equality of the Column and Row Rank of a Matrix

GEORGE MACKIW
Loyola College in Maryland
Baltimore, MD 21210

A basic result in linear algebra is that the row and column spaces of a matrix have dimensions that are equal. In this note we derive this result using an approach that is elementary yet different from that appearing in most current texts. Moreover, we do not rely on the echelon form in our arguments.

The column and row ranks of a matrix A are the dimensions of the column and row spaces, respectively, of A. As is often noted, it is enough to establish that

$$\textit{row rank } A \leq \textit{column rank } A \tag{1}$$

for any matrix A. For, applying the result in (1) to A^t, the transpose of A, would produce the desired equality, since the row space of A^t is precisely the column space of A. We present the argument for real matrices, indicating at the end the modifications necessary in the complex case and the case of matrices over arbitrary fields.

We take full advantage of the following two elementary observations:

1) for any vector x in R^n, Ax is a linear combination of the columns of A, and
2) vectors in the null space of A are orthogonal to vectors in the row space of A, relative to the usual Euclidean product.

Both of these remarks follow easily from the very definition of matrix multiplication. For example, if the vector x is in the null space of A then $Ax = 0$, and thus the inner product of x with the rows of A must be zero. Since these rows span the row space of A, remark 2) follows. An immediate consequence is that only the zero vector can be common to both the null space and row space of A, since 2) requires such a vector to be orthogonal to itself.

Given an m by n matrix A, let the vectors $x_1, x_2, \dots, x_r$ in R^n form a basis for the row space of A. Then the r vectors $Ax_1, Ax_2, \dots, Ax_r$ are in the column space of A and, further, we claim they are linearly independent. For, if $c_1 Ax_1 + c_2 Ax_2 + \cdots + c_r Ax_r = 0$ for some real scalars $c_1, c_2, \dots, c_r$, then $A(c_1 x_1 + c_2 x_2 + \cdots + c_r x_r) = 0$ and the vector $v = c_1 x_1 + c_2 x_2 + \cdots + c_r x_r$ would be in the null space of A. But, v is also in the row space of A, since it is a linear combination of basis elements. So, v is the zero vector and the linear independence of $x_1, x_2, \dots, x_r$ guarantees that $c_1 = c_2 = \cdots = c_r = 0$. The existence of r linearly independent vectors in the column space requires that $r \leq$ *column rank* A. Since r is the row rank of A, we have arrived at the desired inequality (1).

This approach also yields additional information. Let $y_1, y_2 \dots y_k$ form a basis for the null space of A. Since the row space of A and the null space of A intersect trivially, it follows that the set $\{x_1, x_2, \dots, x_r, y_1, y_2, \dots, y_k\}$ is linearly independent. Further, if z is a vector in R^n then Az is in the column space of A and hence expressible as a linear combination of basis elements. Thus, we can write $Az = \sum_{j=1}^{r} d_j Ax_j$, for scalars d_j. But then the vector $z - \sum_{j=1}^{r} d_j x_j$ is in the null space of A and can be written as a linear combination of $y_1, y_2, \dots, y_k$. Thus z can be expressed as a linear combination of the vectors in the set $\{x_1, x_2, \dots, x_r,$

$y_1, y_2, \dots, y_k\}$ and this set then forms a basis for R^n. A dimension count yields $r + k = n$, giving the rank and nullity theorem without use of the echelon form.

Note that the reliance on orthogonality in the arguments above is elementary and does not require the Gram-Schmidt process. The ideas used here can be readily applied to complex matrices with minor modifications. The hermitian inner product is used instead in the complex vector space C^n, as is the hermitian transpose. Observation 2) would note then that vectors in the null space of A are orthogonal to those in the row space $\bar{A}$.

The row and column rank theorem is a well-known result that is valid for matrices over arbitrary fields. The notion of orthogonal complement can be generalized using linear functionals and dual spaces, and the general structure of the arguments here can then be carried over to arbitrary fields. The text [**1**, pp. 97 ff.] contains such an approach.

Many authors base their discussion of rank on the echelon form. The fact that the non-zero rows of the echelon form are a basis for the row space, or that columns in the echelon form containing lead ones can be used to identify a basis for the column space in the original matrix A, are central to such a development of rank. Results such as these follow easily if it is established independently that row rank and column rank must be equal.

The author would like to thank the referee for several valuable suggestions.

REFERENCE

1. Kenneth Hoffman and Ray Kunze, *Linear Algebra*, 2nd edition, Prentice-Hall, Englewood Cliffs, NJ, 1971.

A One-Sentence Proof That $\sqrt{2}$ Is Irrational

DAVID M. BLOOM
Brooklyn College of CUNY
Brooklyn, NY 11210

If $\sqrt{2}$ were rational, say $\sqrt{2} = m/n$ in lowest terms, then also $\sqrt{2} = (2n - m)/(m - n)$ in *lower* terms, giving a contradiction.

(The three verifications needed—that the second denominator is less than the first and still positive, and that the two fractions are equal—are straightforward.)

The argument is not original; it's the algebraic version of a geometric argument given in [**1**, p. 84], and it was presented (in slightly different form) by Ivan Niven at a lecture in 1985. Note, though, that the algebraic argument, unlike the geometric, easily adapts to an arbitrary $\sqrt{k}$ where k is any positive integer that is not a perfect square. Indeed, let j be the integer such that $j < \sqrt{k} < j + 1$. If we had $\sqrt{k} = m/n$ in lowest terms ($m, n \in \mathbf{Z}^+$), then also $\sqrt{k} = (kn - jm)/(m - jn)$ in lower terms, a contradiction.

REFERENCE

1. H. Eves, *An Introduction to the History of Mathematics*, 6th edition, W. B. Saunders Co., Philadelphia, 1990.

$y_1, y_2, \ldots, y_k\}$ and this set then forms a basis for R^n. A dimension count yields $r+k=n$, giving the rank and nullity theorem without use of the echelon form.

Note that the reliance on orthogonality in the arguments above is elementary and does not require the Gram-Schmidt process. The ideas used here can be readily applied to complex matrices with minor modifications. The hermitian inner product is used instead in the complex vector space C^n, as is the hermitian transpose. Observation 2) would note then that vectors in the null space of A are orthogonal to those in the row space $\bar{A}$.

The row and column rank theorem is a well-known result that is valid for matrices over arbitrary fields. The notion of orthogonal complement can be generalized using linear functionals and dual spaces, and the general structure of the arguments here can then be carried over to arbitrary fields. The text [**1**, pp. 97 ff.] contains such an approach.

Many authors base their discussion of rank on the echelon form. The fact that the non-zero rows of the echelon form are a basis for the row space, or that columns in the echelon form containing lead ones can be used to identify a basis for the column space in the original matrix A, are central to such a development of rank. Results such as these follow easily if it is established independently that row rank and column rank must be equal.

The author would like to thank the referee for several valuable suggestions.

REFERENCE

1. Kenneth Hoffman and Ray Kunze, *Linear Algebra*, 2nd edition, Prentice-Hall, Englewood Cliffs, NJ, 1971.

A One-Sentence Proof That $\sqrt{2}$ Is Irrational

DAVID M. BLOOM
Brooklyn College of CUNY
Brooklyn, NY 11210

If $\sqrt{2}$ were rational, say $\sqrt{2}=m/n$ in lowest terms, then also $\sqrt{2}=(2n-m)/(m-n)$ in *lower* terms, giving a contradiction.

(The three verifications needed—that the second denominator is less than the first and still positive, and that the two fractions are equal—are straightforward.)

The argument is not original; it's the algebraic version of a geometric argument given in [**1**, p. 84], and it was presented (in slightly different form) by Ivan Niven at a lecture in 1985. Note, though, that the algebraic argument, unlike the geometric, easily adapts to an arbitrary $\sqrt{k}$ where k is any positive integer that is not a perfect square. Indeed, let j be the integer such that $j<\sqrt{k}<j+1$. If we had $\sqrt{k}=m/n$ in lowest terms $(m, n \in \mathbf{Z}^+)$, then also $\sqrt{k}=(kn-jm)/(m-jn)$ in lower terms, a contradiction.

REFERENCE

1. H. Eves, *An Introduction to the History of Mathematics*, 6th edition, W. B. Saunders Co., Philadelphia, 1990.

On Groups of Order pq

JOSEPH A. GALLIAN
University of Minnesota, Duluth
Duluth, MN 55812

DAVID MOULTON
University of California, Berkeley
Berkeley, CA 94720

The characteristic that distinguishes finite group theory from other branches of algebra is the arithmetical nature of many important theorems. Indeed, proofs in finite group theory often amount to sophisticated counting; Lagrange's theorem and the Sylow theorems are quintessential examples of the genre.

Another highly appealing "arithmetical" result that is typically presented as an application of the Sylow theorems is that a group of order pq, where p and q are primes, $q < p$ and q does not divide $p-1$, is cyclic (see, for example, [**1**, p. 245], [**2**, p. 358] or [**3**, p. 495]). The standard proof of this statement is as follows. By the Sylow theorems there are subgroups H and K of orders p and q that are normal. From the normality of H and K and the fact that $H \cap K = \{e\}$, it follows that elements of H and K commute with each other (see [**2**, p. 358]). Thus, if $|h| = p$ and $|k| = q$, then $|hk| = pq$ and the group is cyclic.

Since the Sylow theorems occur late in many undergraduate abstract algebra textbooks, most students who take only one semester of algebra will not encounter them. In this note we offer a proof of the "pq" result that does not use the Sylow theorems or normality. It can be given as an application of cosets and Lagrange's theorem. Our argument is an excellent illustration of "sophisticated counting."

The proof utilizes the following easy-to-prove facts.

Fact 1. For any group, the conjugacy classes $\text{cl}(a) = \{xax^{-1} | x \in G\}$ partition the group.

Fact 2. For any a in a finite group G, $|\text{cl}(a)| = |G|/|C(a)|$. (Here, $C(a)$ is the subgroup $\{x \in G | xa = ax\}$.)

Fact 3. $|xax^{-1}| = |a|$.

THEOREM. *Suppose p and q are primes with $q < p$ and q does not divide $p-1$. Then any group G of order pq is cyclic.*

Proof. Assume the statement is false. By Lagrange's Theorem we may assume that every nonidentity element of G has order p or q. Let $a \neq e$ belong to G. If $|C(a)| = pq$, then for any $b \notin \langle a \rangle$ we have $|b| \neq |a|$, for otherwise $|\langle a, b \rangle| = p^2$ or q^2. But $|b| \neq |a|$ and $b \in C(a)$ implies $|ab| = pq$. So, we may assume that for all nonidentity elements a in G, $|C(a)| = p$ or q.

We now count the elements of order p and q. Since $|a| = p$ implies $|\text{cl}(a)| = |G|/|C(a)| = q$, the number of elements of order p is a multiple of q. Moreover, because $|a| = p$ implies $|a^i| = p$ for $i = 2, \ldots, p-1$, the number of elements of order p is also a multiple of $p-1$. Since $\gcd(q, p-1) = 1$, the number of elements of order p is a multiple of $q(p-1)$. Analogously, the number of elements of order q is a multiple of $p(q-1)$. Since neither $q(p-1)$ nor $p(q-1)$ divides $pq-1$, not all the nonidentity elements of G can have the same order. Thus, there must be at least $q(p-1) + p(q-1) > pq$ elements in G. This contradiction finishes the proof.

REFERENCES

1. John B. Fraleigh, *A First Course in Abstract Algebra*, 5th edition, Addison Wesley, Reading, 1994.
2. Joseph A. Gallian, *Contemporary Abstract Algebra*, 3rd edition, D. C. Heath, Lexington, 1994.
3. W. Keith Nicholson, *Introduction to Abstract Algebra*, PWS-Kent Publishing, Boston, 1993.

Finding Time

Points chase points
around the circle,
anti-clockwise,
fighting time.
You know time's a circle,
rather than a line.

Make a line a circle!
Pick a center.
Wrap and wrap and wrap
the line around the rim.
How do the ends
get tucked in?

Cut a circle open,
stretch into a line.
Does the cut destroy
a point or fit
between a pair?
If the cut's midway

from now to Tuesday,
how do I get there?
Do I move on
by going back,
or may I
skip a space?

A square is neither line
nor circle; it is timeless.
Points don't chase around
a square. Firm, steady,
it sits there and knows
its place. A circle
won't be squared.

—JoAnne Growney
Department of Mathematics and Computer Science
Bloomsburg University, Bloomsburg, PA 17815

REFERENCES

1. John B. Fraleigh, *A First Course in Abstract Algebra*, 5th edition, Addison Wesley, Reading, 1994.
2. Joseph A. Gallian, *Contemporary Abstract Algebra*, 3rd edition, D. C. Heath, Lexington, 1994.
3. W. Keith Nicholson, *Introduction to Abstract Algebra*, PWS-Kent Publishing, Boston, 1993.

Finding Time

Points chase points
around the circle,
anti-clockwise,
fighting time.
You know time's a circle,
rather than a line.

Make a line a circle!
Pick a center.
Wrap and wrap and wrap
the line around the rim.
How do the ends
get tucked in?

Cut a circle open,
stretch into a line.
Does the cut destroy
a point or fit
between a pair?
If the cut's midway

from now to Tuesday,
how do I get there?
Do I move on
by going back,
or may I
skip a space?

A square is neither line
nor circle; it is timeless.
Points don't chase around
a square. Firm, steady,
it sits there and knows
its place. A circle
won't be squared.

—JoAnne Growney
Department of Mathematics and Computer Science
Bloomsburg University, Bloomsburg, PA 17815

Thinning Out the Harmonic Series

G. HOSSEIN BEHFOROOZ
Utica College of Syracuse University
Utica, NY 13502

Introduction It is well known that the harmonic series

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots \tag{1}$$

is divergent, and its partial sum

$$S_n = \sum_{k=1}^{n} \frac{1}{k}$$

is close to $\log n$; this shows that the harmonic series diverges very, very slowly. For example, it takes more than 1.5×10^{43} terms for its partial sums to reach 100; see **[3]**, **[4]**, and **[5]**. In this paper, we refine and thin out the harmonic series to show that the divergence of this series depends on some of its specific terms and without those terms the remaining subseries is convergent. Then, an upper bound and an approximate value will be given to the value of the convergent subseries. Finally, by use of the derived results, it will be shown that the divergence of the Euler series $\sum \frac{1}{p}$ (over prime numbers p) depends on specific prime numbers.

We begin with two interesting results about subseries of (1). For simplicity, it is convenient to introduce a little terminology here. Let $r = 0, 1, 2, \ldots, 9$. We say that the positive integer n *contains* r, if r is one of the digits of n, *i.e.*, the decimal representation of n contains at least one r, and we write $r \in n$; otherwise n is called an $r-$*free* integer and we will denote this by $r \notin n$. For example, 16 contains 1 and 6, but it is obviously a $2-$ free number, and we can also write $2 \in 289$, $5 \in 1059$, and $6 \notin 281$.

PROPOSITION 1. *The subseries*

$$\sum_{9 \in n} \frac{1}{n} = \frac{1}{9} + \frac{1}{19} + \frac{1}{29} + \cdots + \frac{1}{89} + \frac{1}{90} + \frac{1}{91} + \cdots + \frac{1}{99} + \cdots, \tag{2}$$

of the harmonic series (1), *which consists of all terms of* (1) *with one or more nines, is a divergent series.*

Proof. It is clear that the series (2) with positive terms is greater than the divergent series

$$\sum_{k=0}^{\infty} \frac{1}{10k+9},$$

and so it is a divergent series.

Now, if all the terms of the subseries (2) are deleted (omitted) from the original harmonic series (1), the remaining series

$$\sum_{9 \notin n} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{8} + \frac{1}{10} + \cdots, \tag{3}$$

consisting of all $9-$ free terms of (1), appears to be divergent. If we compare the terms of the two series (2) and (3), we see that, between 1 and 9 there is only one term from (2) but there are eight terms from (3), and similarly between 10 and 99, there are 18 terms from (2), but 72 terms from (3), and so on. For these reasons, it seems that the series (3) is divergent. In fact, though, the series (3) is convergent, and this has been known for a long time! In 1914, Kempner **[8]** proved this by using mathematical induction, and he called (3) "A Curious Convergent Series"; see also **[1]**, **[2]**, **[3]**, **[7]**, **[9]**, **[10]**, **[12]** and **[13]**. Here, we will give a different proof, which gives us a sharper upper bound for the sum of (3). Then by extending this idea we will find further interesting and new surprising results about other subseries of the harmonic series (1).

PROPOSITION 2. *The series* (3) *is convergent.*

Proof. There are only 8×9^k different $(k+1)$-digit positive integers with their digits belonging to $\{0, 1, 2, \ldots, 8\}$. So, for any natural number k, there are only 8×9^k 9-free positive integers between 10^k and $10^{k+1} - 1$. The series (3) can be rearranged and rewritten as:

$$\sum_{9 \notin n} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{8} + \frac{1}{10} + \cdots$$

$$= \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{8}\right) + \left(\frac{1}{10} + \cdots + \frac{1}{88}\right)$$

$$+ \left(\frac{1}{100} + \cdots + \frac{1}{888}\right) + \cdots + \left(\frac{1}{10^k} + \cdots + \frac{1}{\underbrace{88\cdots8}_{k \text{ times}}}\right) + \cdots$$

$$< \left(\overbrace{1 + 1 + \cdots + 1}^{8 \times 9^0 \text{ times}}\right) + \left(\overbrace{\frac{1}{10} + \frac{1}{10} + \cdots + \frac{1}{10}}^{8 \times 9^1 \text{ times}}\right) + \cdots + \left(\overbrace{\frac{1}{10^k} + \cdots + \frac{1}{10^k}}^{8 \times 9^k \text{ times}}\right) + \cdots$$

$$= \left(8 \times 9^0 \times \frac{1}{10^0}\right) + \left(8 \times 9^1 \times \frac{1}{10^1}\right) + \left(8 \times 9^2 \times \frac{1}{10^2}\right)$$

$$+ \cdots + \left(8 \times 9^k \times \frac{1}{10^k}\right) + \cdots = 8 \sum_{k=0}^{\infty} \left(\frac{9}{10}\right)^k = 80.$$

This proves that (3) converges.

For any positive integer k, let N_k be the number of all 9-free positive integers that are less than 10^k, and M_k be the number of positive integers that are less than 10^k and contain at least one digit 9. From the above proof we have $N_k = \sum_{i=0}^{k-1} 8 \times 9^i = 9^k - 1$ and $M_k = (10^k - 1) - (9^k - 1) = 10^k - 9^k$. These results imply that $\lim_{k \to \infty} N_k / M_k = 0$.

The series (3) converges very, very slowly. The value of its partial sum with 10^5 terms reaches 12.0908, which is relatively very far from the actual value of the series, 22.921. After 10^5 terms from this convergent series, the partial sum and the exact value do not have even a single digit in common!

Obviously, the digit 9 does not play a critical role in Propositions 1 and 2, so we have the following theorem.

THEOREM 1. *For any number* $r = 0, 1, \ldots, 8, 9$,

$$\text{(I) the series } \sum_{r \in n} \frac{1}{n} \text{ is divergent,} \tag{4}$$

$$\text{(II) the series } \sum_{r \notin n} \frac{1}{n} \text{ is convergent.} \tag{5}$$

Recently, Baillie [2] has calculated (to 20 decimal places) the value of the convergent series (5) for $r = 0, 1, \ldots, 9$. Table 1 gives three decimal-place values of these sums $T_0, T_1, T_2, \ldots, T_9$; see also [3] and [11].

TABLE 1

$T_0 = 23.103$	$T_1 = 16.177$	$T_2 = 19.257$	$T_3 = 20.570$	$T_4 = 21.327$
$T_5 = 21.835$	$T_6 = 22.206$	$T_7 = 22.493$	$T_8 = 22.726$	$T_9 = 22.921$

We can extend the above result by changing the base from base 10 to any particular base b. For example, if we consider base 100 then we obtain a convergent series by deleting any particular digit in base 100; see [3]. Also, if we look at Table 1, we see that the values of the T_i are increasing for $i \geq 1$. This suggests the following *open question* (prove or disprove).

Open Question. *For any base b, is $T_m < T_n$, whenever $1 \leq m < n$?*

New surprising results Let us consider two special cases in Theorem 1, with $r = 9$ and $r = 8$:

$$\text{(I) the series } \sum_{9 \in n} \frac{1}{n} \text{ is divergent,} \tag{6}$$

$$\text{(II) the series } \sum_{8 \notin n} \frac{1}{n} \text{ is convergent.} \tag{7}$$

It is obvious that the convergent series (7) contains all terms of the divergent series (6), except those terms that contain both digits 8 and 9 (like $89, 98, 189, \cdots$). So, the divergence of (6) and also the harmonic series (1) depends on such terms. This means that the series

$$\sum_{8,9 \in n} \frac{1}{n} = \frac{1}{89} + \frac{1}{98} - \frac{1}{189} + \cdots \tag{8}$$

is divergent, and its complement with respect to (1) is convergent. In general, we have the following theorem.

THEOREM 2. *For any two different integers $r, s = 0, 1, \ldots, 9$,*

$$\text{(I) the series } \sum_{r, s \in n} \frac{1}{n} \text{ is divergent,} \tag{9}$$

$$\text{(II) the series } \sum_{r \notin n \,\vee\, s \notin n} \frac{1}{n} \text{ is convergent.} \tag{10}$$

If we sum the 10 convergent series with their sums given in Table 1, then the resulting series

$$T = \sum_{0 \notin n} \frac{1}{n} + \sum_{1 \notin n} \frac{1}{n} + \cdots + \sum_{9 \notin n} \frac{1}{n} \tag{11}$$

will be a convergent series with sum less than 213. Obviously, there are still some other terms in the harmonic series that do not belong to the convergent series (11). All such terms contain each of the digits $0, 1, 2, \ldots, 9$, at least once, and $\frac{1}{1023456789}$ is the first of these terms. Now we state the main theorems of this paper.

THEOREM 3. *Let D be the set of positive integers that contain each of the digits $0, 1, \dots, 9$ at least once; then*

$$\text{(I) the series } \sum_{n \in D} \frac{1}{n} \text{ is divergent,} \tag{12}$$

$$\text{(II) the series } \sum_{n \notin D} \frac{1}{n} \text{ is convergent.} \tag{13}$$

THEOREM 4. *For any positive integer k, let N_k be the number of positive integers n that are less than 10^k and $n \in D$, and M_k be the number of positive integers m that are less than 10^k and $m \notin D$. Then $\lim_{k \to \infty} N_k / M_k = 0$.*

Proof. This theorem is a special case of the following theorem.

THEOREM 5. *Suppose that C is a subset of positive integers and $\sum_{n \in C} 1/n$ is convergent. For any positive integer k, let N_k be the number of elements in C that are $\leq 10^k$, and $M_k = 10^k - N_k$. Then $\lim_{k \to \infty} N_k / M_k = 0$.*

Proof. It suffices to show that

$$\lim_{k \to \infty} \frac{N_k}{10^k} = 0.$$

Since $N_k - N_{k-1}$ is the number of elements of $C_k = \{n \in C \mid 10^{k-1} < n \leq 10^k\}$, we have

$$\sum_{n \in C_k} \frac{1}{n} \geq \sum_{n \in C_k} \frac{1}{10^k} = \frac{N_k - N_{k-1}}{10^k}$$

and so

$$\sum_{k=2}^{\infty} \frac{N_k - N_{k-1}}{10^k} \leq \sum_{n \in C} \frac{1}{n} < \infty.$$

Thus

$$\lim_{k \to \infty} \frac{N_k - N_{k-1}}{10^k} = 0.$$

If we set $a_k = N_k / 10^k$, then for all k, we have $a_k \leq 1$ and $\lim_{k \to \infty} (a_k - a_{k-1})/10 = 0$. Now we show that $\lim_{k \to \infty} a_k = 0$. Let $\epsilon > 0$. There exists an integer n, such that $(a_k - a_{k-1})/10 < \epsilon$, for all $k \geq n$. By induction on m, it can be shown that for any $m \geq 1$,

$$a_{n+m} \leq \epsilon \left(\sum_{k=0}^{m-1} \frac{1}{10^k} \right) + \frac{1}{10^m}.$$

Letting $m \to \infty$, this will imply that $\limsup a_k \leq (10/9)\epsilon$, and since $\epsilon > 0$ is arbitrary, we conclude that $\lim_{k \to \infty} a_k = 0$.

A note on the Euler series Now we consider the divergent Euler series $\sum \frac{1}{p}$, over prime numbers p (see, Hardy and Wright; [**6**; pp. 17, 120, and 349]). Again, the sum of ten convergent series (11) contains all the terms of $\sum \frac{1}{p}$, except for those prime numbers that belong to the set D. Hence the divergence of the Euler series $\sum \frac{1}{p}$ is due to such prime numbers. For example, 1012345789, 10123465789, and 10123465897 are the first three such primes. When these primes are deleted from the Euler series $\sum \frac{1}{p}$, the remaining series is convergent.

THEOREM 6. *Let E be the set of all primes that contain each of the digits* 0, 1, ..., 9 *at least once. Then*

$$\text{(I) } \textit{the series } \sum_{p \in E} \frac{1}{p} \textit{ is divergent,} \tag{14}$$

$$\text{(II) } \textit{the series } \sum_{p \notin \bar{E}} \frac{1}{p} \textit{ is convergent.} \tag{15}$$

The divergence of series (14) implies that the set E is infinite.

Acknowledgment. I would like to thank the referees for their comments.

REFERENCES

1. T. M. Apostol, *Mathematical Analysis*, Addison-Wesley, Reading, MA, 1960, p. 384, Ex. 12–16.
2. R. Baillie, Sums of reciprocals of integers missing a given digit, *Amer. Math. Monthly*, 86 (1979) 372–374.
3. R. P. Boas, Jr., Some remarkable sequences of integers, *Mathematical Plums*, MAA, Washington, DC, 1979.
4. R. P. Boas, Jr., Partial sums of infinite series and how they grow, *Amer. Math. Monthly*, 84 (1977) 237–258.
5. R. P. Boas, Jr., and J. W. Wrench, Jr., Partial sums of the harmonic series, *Amer. Math Monthly*, 78 (1971) 864–870.
6. G. H. Hardy and E. M. Wright, *An Introduction to Number Theory*, Oxford University Press, Oxford, 1968.
7. F. Irwin, A curious convergent series, *Amer. Math. Monthly*, 23 (1916) 149–152.
8. A. J. Kempner, A curious convergent series, *Amer. Math. Monthly*, 21 (1914) 48–50.
9. E. J. Moulton, Solution of Problem #453, *Amer. Math. Monthly*, 23 (1916) 302–303.
10. G. Polya and G. Szego, *Problems and Theorems in Analysis*, Vol. 1, part I, problem 124, Springer-Verlag, New York, 1972.
11. A. C. Segal, A limit problem, elementary problems and solutions #E 2204, *Amer. Math. Monthly*, 77 (1970) 1009–1010.
12. A. D. Wadhwa, An interesting subseries of the harmonic series, *Amer. Math. Monthly*, 82 (1975) 931–933.
13. A. D. Wadhwa, Some convergent subseries of the harmonic series, *Amer. Math. Monthly*, 85 (1978) 661–663.

Limitless Integrals and a New Definition of the Logarithm

DAVID SHELUPSKY
Department of Physics
The City College of New York
New York, NY 10031

The integrals we consider are

$$\int_a^b x^n\,dx \quad \text{for integers } n \neq -1. \tag{1}$$

We show by a clever choice of the evaluation points in the partition intervals used to define the Riemann sums for these integrals that we can arrange for these sums to agree exactly with the value of the integral. Thus the integrals can be evaluated without having to take limits. We also show how the inequalities used here lead to a new definition of the natural logarithm that does not involve limits, series or integrals, or even the exponential function.

A Riemann sum for the integral in (1) is

$$\sum_{i=1}^{m} u_i^n(x_i - x_{i-1}) \tag{2}$$

where $(x_0, x_1, x_2, \dots x_m)$ is an ordered partition of $[a,b]$ and $x_{i-1} < u_i < x_i$ for $i = 1, 2, \dots m$. For convenience we suppose that $0 \le a < b$, and that $a > 0$ when n is negative. Since the case $n = 0$ is trivial, we assume that $n \neq 0, -1$. Let $u_i = M_n(x_i, x_{i-1})$, where $M_n(u,v)$ is defined for $u, v > 0$ by

$$M_n(u,u) = u \tag{3}$$

and

$$M_n(u,v) = \left[\frac{u^{n+1} - v^{n+1}}{(n+1)(u-v)}\right]^{1/n}$$

for $u \neq v$. If $0 \le v < u$ then, as we show below, $v < M_n(u,v) < u$ and so $x_{i-1} < u_i < x_i$. With this choice of u_i we have

$$u_i^n(x_i - x_{i-1}) = \left(x_i^{n+1} - x_{i-1}^{n+1}\right)/(n+1)$$

and so the Riemann sum (2) telescopes and has the value

$$(b^{n+1} - a^{n+1})/(n+1). \tag{4}$$

The mesh size $\max_i(x_i - x_{i-1})$ may be made arbitrarily small, so, invoking the usual theorems on the existence of the integral (see **[1]**, for example) we see that the value of the integral (1) must be the expression (4). For n negative we only need these results for $v > 0$.

What makes the method work is the fact that each expression M_n defined by (3) is an average. We note in fact that $M_1(u,v) = (u+v)/2$ and $M_{-2}(u,v) = (uv)^{1/2}$ are the arithmetic and geometric means respectively. For n a positive integer the mean-value property of M_n follows from the formula

$$M_n(u,v) = \left[(u^n + u^{n-1}v + \cdots + v^n)/(n+1)\right]^{1/n} \tag{5}$$

and the fact that the radicand is the arithmetic average of $n+1$ non-negative numbers, all not equal, the largest of which is u^n and the smallest of which is v^n. Since we have ruled out $n=0$ or -1, and $n=-2$ is the easy case of the geometric mean, we only have to consider $n \leq -3$ among the negative integers. Some easy algebra shows that the inequality $M_{-n}(u,v) > v$ is equivalent to the inequality

$$M_{n-2}\left(\frac{v}{u}, 1\right) < \left(\frac{u}{v}\right)^{1/(n-2)},$$

and this follows from $u/v > 1$ and the result already established that M_{n-2} is an average. The inequality $M_n(u,v) < u$ for n negative is proved similarly. Various generalizations of the M_n are discussed in **[2]**.

When $n=-1$ in (1) this method seems incapable of giving an evaluation of the integral (1) that does not depend on dealing with limits because of a lack of a suitable elementary definition of the logarithm. Below we make it work by giving a new definition of the logarithm. We note first that the usual tactic in this situation is to define the logarithm by setting

$$L(x) = \int_1^x \frac{1}{t}\,dt. \tag{6}$$

The functional equation of the logarithm, $L(xy) = L(x) + L(y)$, is then obtained from this definition (see **[3]**, p. 93, problems 3 and 4), and this functional equation is easily transformed into Cauchy's functional equation, $f(x+y) = f(x) + f(y)$, by introducing $f(x) = L(a^x)$, $a > 1$. From (6) the continuity of L follows as does the fact that the monotone function L takes on every real value exactly once. From the form of the continuous solutions of Cauchy's functional equation ($f(x) = cx$) it follows that $L(x) = \log_e x$, where e is defined by $L(e) = 1$.

We suggest another method of attack: We note that L defined by (6) must satisfy the inequality

$$u^{-1} < \frac{L(u) - L(v)}{u - v} < v^{-1} \tag{7}$$

for all positive u and v, $v < u$. We assert that this inequality defines the logarithm up to an additive constant. With the added condition $L(1) = 0$, we make this our definition of the logarithm. The unique solution of (7) is then $L(x) = \log_e x$ (with e now defined as usual).

We give two proofs of this. Our first proof depends on the calculus and runs as follows (we omit the details): We show in succession that any solution of (7) is necessarily strictly increasing, continuous, differentiable, and satisfies $L'(x) = x^{-1}$. The assertion now follows. For a similar elementary definition of the exponential function, see **[4]**, part (a).

The following treatment of (7) is more in the spirit of this paper and depends only on the existence of the integral. Let $n=-1$ in (1) and (2) and let u_i in the Riemann sum be defined by

$$u_i^{-1} = \frac{L(x_i) - L(x_{i-1})}{x_i - x_{i-1}},$$

where L is any solution of (7). By (7) $x_{i-1} < u_i < x_i$, and the value of the Riemann sum is $L(b) - L(a)$. It follows as before that

$$L(b) - L(a) = \int_a^b \frac{1}{t}\,dt.$$

Thus any solution of (7) together with $L(1) = 0$ is necessarily given by (6).

REFERENCES

1. A. E. Taylor, *Advanced Calculus*, Ginn and Co., Boston, 1955, Chapter 16.
2. G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, Cambridge, 1959.
3. Z. Nehari, *Conformal Mapping*, McGraw-Hill, New York, 1952.
4. D. Shelupsky, Problem E3127, *Amer. Math. Monthly*, 95 (1988), 457–458.

A Note on Cauchy Sequences

JOSÉ A. FACENDA AGUIRRE[1]
Universidad de Sevilla
41080 Sevilla, Spain

In undergraduate courses we teach that a sequence of real numbers converges if, and only if, it is a Cauchy sequence. Usually, the students have no problems with the necessary condition. Our aim in this note is to clarify the sufficient condition with the use of pairs of monotone sequences. The notation is standard; we follow **[1]**.

First of all, we recall the notion of pairs of monotone sequences:

Let $\{a_n\}_{n \in \mathbb{N}}$ and $\{b_n\}_{n \in \mathbb{N}}$ be sequences of real numbers. We say that $(\{a_n\}_{n\in\mathbb{N}}, \{b_n\}_{n\in\mathbb{N}})$ is a *pair of monotone sequences* if:

(1) $\{a_n\}_{n \in \mathbb{N}}$ is nondecreasing and $\{b_n\}_{n \in \mathbb{N}}$ is nonincreasing, and
(2) $a_n \leq b_n \ \forall n \in \mathbb{N}$.

Plainly, $a_p \leqslant b_q$, for all $p, q \in \mathbb{N}$ so there exist $a = \sup\{a_n\colon n \in \mathbb{N}\} \in \mathbb{R}$ and $b = \inf\{b_n\colon n \in \mathbb{N}\} \in \mathbb{R}$, and $a \leq b$.

When can we assure that $a = b$? If the following condition is fulfilled:

$$\forall \varepsilon > 0 \ \exists k \in \mathbb{N} \quad \text{such that } b_k - a_k < \varepsilon, \qquad a = b. \tag{$*$}$$

We recall the definition of *limit superior* and *limit inferior* of a bounded sequence of real numbers. Let $\{x_n\}_{n \in \mathbb{N}} \subset \mathbb{R}$ be a bounded sequence. For each $n \in \mathbb{N}$ define

$$y_k = \inf\{x_n\colon n \geq k\} \quad \text{and} \quad z_k = \sup\{x_n\colon n \geq k\}.$$

Let $y = \lim_{k\to\infty} y_k = \sup\{y_k\colon k \in \mathbb{N}\}$ and $z = \lim_{k\to\infty} z_k = \inf\{z_k\colon k \in \mathbb{N}\}$. The numbers $y, z \in \mathbb{R}$ are called the *limit inferior* of $\{x_n\}_{n \in \mathbb{N}}$ and the *limit superior* of $\{x_n\}_{n \in \mathbb{N}}$ respectively. (Note that $(\{y_n\}_{n \in \mathbb{N}}, \{z_n\}_{n \in \mathbb{N}})$ is a pair of monotone sequences). Furthermore, we recall that every Cauchy sequence is bounded and a sequence of real numbers is convergent if, and only if, the limits superior and inferior exist in $\mathbb{R}$ and are equal.

With these remarks, we are ready. Suppose that $\{x_n\}_{n \in \mathbb{N}} \subset \mathbb{R}$ is a Cauchy sequence. Since $\{x_n\}_{n \in \mathbb{N}}$ is bounded, $y = \liminf x_n \in \mathbb{R}$, $z = \limsup x_n \in \mathbb{R}$. To show $\{x_n\}_{n \in \mathbb{N}}$ converges, we only have to prove that $y = z$.

We know that

$$\forall \varepsilon > 0 \ \exists k \in \mathbb{N} \quad \text{such that } \forall m, n \geq k, \qquad |x_m - x_n| < \varepsilon.$$

[1] Supported by D.G.I.C.Y.T., grant no. PB 93-0926.

Thus any solution of (7) together with $L(1)=0$ is necessarily given by (6).

REFERENCES

1. A. E. Taylor, *Advanced Calculus*, Ginn and Co., Boston, 1955, Chapter 16.
2. G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, Cambridge, 1959.
3. Z. Nehari, *Conformal Mapping*, McGraw-Hill, New York, 1952.
4. D. Shelupsky, Problem E3127, *Amer. Math. Monthly*, 95 (1988), 457–458.

A Note on Cauchy Sequences

JOSÉ A. FACENDA AGUIRRE[1]
Universidad de Sevilla
41080 Sevilla, Spain

In undergraduate courses we teach that a sequence of real numbers converges if, and only if, it is a Cauchy sequence. Usually, the students have no problems with the necessary condition. Our aim in this note is to clarify the sufficient condition with the use of pairs of monotone sequences. The notation is standard; we follow **[1]**.

First of all, we recall the notion of pairs of monotone sequences:

Let $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ be sequences of real numbers. We say that $(\{a_n\}_{n\in\mathbb{N}}, \{b_n\}_{n\in\mathbb{N}})$ is a *pair of monotone sequences* if:

(1) $\{a_n\}_{n\in\mathbb{N}}$ is nondecreasing and $\{b_n\}_{n\in\mathbb{N}}$ is nonincreasing, and
(2) $a_n \le b_n \ \forall n \in \mathbb{N}$.

Plainly, $a_p \leqslant b_q$, for all $p, q \in \mathbb{N}$ so there exist $a = \sup\{a_n\colon n \in \mathbb{N}\} \in \mathbb{R}$ and $b = \inf\{b_n\colon n \in \mathbb{N}\} \in \mathbb{R}$, and $a \le b$.

When can we assure that $a = b$? If the following condition is fulfilled:

$$\forall \varepsilon > 0\ \exists k \in \mathbb{N} \quad \text{such that } b_k - a_k < \varepsilon, \qquad a = b. \tag{$*$}$$

We recall the definition of *limit superior* and *limit inferior* of a bounded sequence of real numbers. Let $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ be a bounded sequence. For each $n \in \mathbb{N}$ define

$$y_k = \inf\{x_n\colon n \ge k\} \quad \text{and} \quad z_k = \sup\{x_n\colon n \ge k\}.$$

Let $y = \lim_{k\to\infty} y_k = \sup\{y_k\colon k \in \mathbb{N}\}$ and $z = \lim_{k\to\infty} z_k = \inf\{z_k\colon k \in \mathbb{N}\}$. The numbers $y, z \in \mathbb{R}$ are called the *limit inferior* of $\{x_n\}_{n\in\mathbb{N}}$ and the *limit superior* of $\{x_n\}_{n\in\mathbb{N}}$ respectively. (Note that $(\{y_n\}_{n\in\mathbb{N}}, \{z_n\}_{n\in\mathbb{N}})$ is a pair of monotone sequences). Furthermore, we recall that every Cauchy sequence is bounded and a sequence of real numbers is convergent if, and only if, the limits superior and inferior exist in $\mathbb{R}$ and are equal.

With these remarks, we are ready. Suppose that $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ is a Cauchy sequence. Since $\{x_n\}_{n\in\mathbb{N}}$ is bounded, $y = \liminf x_n \in \mathbb{R}$, $z = \limsup x_n \in \mathbb{R}$. To show $\{x_n\}_{n\in\mathbb{N}}$ converges, we only have to prove that $y = z$.

We know that

$$\forall \varepsilon > 0\ \exists k \in \mathbb{N} \quad \text{such that } \forall m, n \ge k, \qquad |x_m - x_n| < \varepsilon.$$

[1] Supported by D.G.I.C.Y.T., grant no. PB 93-0926.

Hence,

$$\forall \varepsilon > 0\ \exists k \in \mathbb{N} \quad \text{such that } \sup\{|x_m - x_n|: m, n \geq k\} < \varepsilon.$$

Now then

$$\begin{aligned}\sup\{|x_m - x_n|: m, n \geq k\} &= \sup\{x_m - x_n: m, n \geq k\} \\ &= \sup\{x_m: m \geq k\} - \inf\{x_n: n \geq k\} \\ &= z_k - y_k.\end{aligned}$$

Clearly,

$$\forall \varepsilon > 0\ \exists k \in \mathbb{N} \quad \text{such that } z_k - y_k < \varepsilon.$$

Because the pair of monotone sequences $(\{y_n\}_{n\in\mathbb{N}}, \{z_n\}_{n\in\mathbb{N}})$ fulfills the condition $(*)$, we can assure that $y = z$ and so, the sequence $\{x_n\}_{n\in\mathbb{N}}$ is convergent.

In the case when $\{x_n\}_{n\in\mathbb{N}}$ is a sequence of complex numbers, $x_n = a_n + ib_n$, we apply the result to the sequences of real numbers $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$.

Acknowledgement. The author would like to thank the referees for their suggestions.

REFERENCE

1. K. R. Stromberg, *An Introduction to Classical Real Analysis*, Wadsworth, Belmont, CA, 1981.

Bell's Conjecture*

For math, the Oscar envelope
(Assured by Price and Waterhouse)
Would list a three-way tie, I'd hope:
Archimedes, Newton, Gauss.
fine

Archimedes' *modern* mind
(Narrowly he bounded pi)
Impelled to seek and swift to find,
Defined the Hellenistic high.

Newton's fluxions formed the frame
That fit the Universal Law.
Even Leibniz spread his fame:
"We know the Lion by his claw."

Many Magi graced the scene
But Gauss was greater than all since.
If Number Theory is the Queen,
Carl Friedrich is its freshest Prince.
D.C.

—J. D. MEMORY
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NC 275515-2688

*E. T. Bell, *Men of Mathematics*, New York, 1965.

Hence,

$$\forall \varepsilon > 0\ \exists k \in \mathbb{N} \quad \text{such that } \sup\{|x_m - x_n|: m, n \geq k\} < \varepsilon.$$

Now then

$$\begin{aligned}\sup\{|x_m - x_n|: m, n \geq k\} &= \sup\{x_m - x_n: m, n \geq k\}\\ &= \sup\{x_m: m \geq k\} - \inf\{x_n: n \geq k\}\\ &= z_k - y_k.\end{aligned}$$

Clearly,

$$\forall \varepsilon > 0\ \exists k \in \mathbb{N} \quad \text{such that } z_k - y_k < \varepsilon.$$

Because the pair of monotone sequences $(\{y_n\}_{n\in\mathbb{N}}, \{z_n\}_{n\in\mathbb{N}})$ fulfills the condition $(*)$, we can assure that $y = z$ and so, the sequence $\{x_n\}_{n\in\mathbb{N}}$ is convergent.

In the case when $\{x_n\}_{n\in\mathbb{N}}$ is a sequence of complex numbers, $x_n = a_n + ib_n$, we apply the result to the sequences of real numbers $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$.

Acknowledgement. The author would like to thank the referees for their suggestions.

REFERENCE

1. K. R. Stromberg, *An Introduction to Classical Real Analysis*, Wadsworth, Belmont, CA, 1981.

Bell's Conjecture*

For math, the Oscar envelope
(Assured by Price and Waterhouse)
Would list a three-way tie, I'd hope:
Archimedes, Newton, Gauss.
fine

Archimedes' *modern* mind
(Narrowly he bounded pi)
Impelled to seek and swift to find,
Defined the Hellenistic high.

Newton's fluxions formed the frame
That fit the Universal Law.
Even Leibniz spread his fame:
"We know the Lion by his claw."

Many Magi graced the scene
But Gauss was greater than all since.
If Number Theory is the Queen,
Carl Friedrich is its freshest Prince.
D.C.

—J. D. MEMORY
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NC 275515-2688

*E. T. Bell, *Men of Mathematics*, New York, 1965.

Condensing a Slowly Convergent Series

BRUCE CHRISTIANSON
School of Information Sciences
University of Hertfordshire
England AL10 9AB

The condensation test for convergence says that, for any monotone decreasing positive sequence $(a_n)_{n\geq 1}$, the convergence of the corresponding series $\Sigma_n a_n$ is equivalent to the convergence of the condensed series $\Sigma_n 2^n a_{2^n}$. This test (sometimes called Cauchy's condensation test) used to be known to every undergraduate, but has lately rather fallen out of fashion. Probably the best account of the condensation test and its numerous generalizations and refinements is still the classic book by Konrad Knopp **[3]**, now republished by Dover. There is also a good treatment in **[2]**.

We can establish that the condensation test works by the following argument. Define $l(1)=l(2)=1$ and $l(n)=k$ for $2^{k-1}<n\leq 2^k$, $k>1$. We can regard $l(n)$ as the integer ceiling of $\log_2 n$. By hypothesis, $a_{2^k}\leq a_n\leq a_{2^{k-1}}$ for all the 2^{k-1} values of n for which $l(n)=k$. This means that

$$a_2+\frac{1}{2}\sum_k 2^k a_{2^k}=2a_2+\sum_k 2^k a_{2^{k+1}}\leq\sum_n a_n\leq a_1+\sum_k 2^k a_{2^k}$$

whence $\Sigma_n a_n$ and $\Sigma_n 2^n a_{2^n}$ converge or diverge together, as asserted.

The test condenses the series by dividing it into chunks, the nth chunk consisting of 2^n terms, and because the terms of the series are decreasing, the test need consider only one representative from each chunk to establish convergence.

Here is an example of the use of the condensation test. If $a_n=1/n$ then $2^n a_{2^n}=2^n\cdot(1/2^n)=1$ for all n. Since we know that $\Sigma_n 1$ diverges, the condensation test tells us that $\Sigma_n 1/n$ diverges. Repeated use of the condensation test tells us that

$$\sum_n\frac{1}{n},\ \sum_n\frac{1}{n\cdot l(n)},\ \sum_n\frac{1}{n\cdot l(n)\cdot l(l(n))},\ \sum_n\frac{1}{n\cdot l(n)\cdot l(l(n))\cdot l(l(l(n)))},\ \text{etc.}$$

all diverge. Similarly, if $a_n=1/n^2$ then $2^n a_{2^n}=2^n\cdot(1/2^n)^2=2^{-n}$. Since we know that $\Sigma_n 2^{-n}$ converges, repeated use of the condensation test tells us that

$$\sum_n\frac{1}{n^2},\ \sum_n\frac{1}{n\cdot[l(n)]^2},\ \sum_n\frac{1}{n\cdot l(n)\cdot[l(l(n))]^2},$$

$$\sum_n\frac{1}{n\cdot l(n)\cdot l(l(n))\cdot[l(l(l(n)))]^2},\ \text{etc.}$$

all converge.

This makes it natural to consider the limiting case $\Sigma_n b_n$ where

$$b_n=\frac{1}{n\cdot\pi(n)}$$

and we define $\pi(1)=1$, $\pi(n)=l(n)\cdot\pi(l(n))$ for $n>1$, so that for example $b_{65536}=1/(65536\cdot 16\cdot 4\cdot 2)$. We assert that the sequence (b_n) shrinks to zero more slowly than do the terms in any of the convergent series given above, but more rapidly than those in any of the corresponding divergent series. For example, it is clear that

$l(n)\cdot l(l(n))/\pi(n) = 1/\pi(l(l(n))) \to 0$ with n, and similarly $l(n)\cdot[l(l(n))]^2/\pi(n) = l(l(n))/\pi(l(l(n))) \to \infty$ with n provided we can show $n/\pi(n) \to \infty$. This can be shown by induction using the fact that $2^n/\pi(2^n) = [2^n/n^2][n/\pi(n)]$.

Given our results so far, it is natural to ask whether $\Sigma_n b_n$ converges or diverges. The condensation test appears to give no help, since $2^n b_{2^n} = 1/\pi(2^n) = 1/(n\cdot\pi(n)) = b_n$, so the series is transformed into itself. But we shall see that by using a refinement of the condensation test we can determine not only that the series converges, but that it does so sufficiently rapidly for us to get a tight estimate of the limit using a geometric series.

To prove convergence of $\Sigma_n b_n$ we argue as follows. Let c_n be the sum of the first 2^n terms, so that $c_n = b_1 + \cdots + b_{2^n}$. It is clear from the definition of b_n that

$$c_n - c_{n-1} = \left[\frac{1}{2^{n-1}+1} + \frac{1}{2^{n+1}+2} + \cdots + \frac{1}{2^n}\right] b_n.$$

Now divide the interval $[1,2]$ into 2^{n-1} equal parts and consider the corresponding upper and lower Riemann sums for $L = \int_1^2 1/x\,dx = \log_e 2 = 0.693147\ldots$. This shows that

$$2^{-n} \geq \int_1^2 \frac{dx}{x} - \left[\frac{1}{2^{n-1}+1} + \frac{1}{2^{n-1}+2} + \cdots + \frac{1}{2^n}\right] \geq 0.$$

Consequently $c_n - c_{n-1} \leq Lb_n$, and summing this over a range of n gives

$$c_{2^n} - c_1 \leq Lb_2 + \cdots + Lb_{2^n} = Lc_n - L \leq Lc_{2^n} - L,$$

whence $c_{2^n} \leq (1.5 - L)/(1-L) \leq 3$ for all n, which establishes convergence of $\Sigma_n b_n$.

To obtain a closer bound on the value of this sum, argue as follows. From the Riemann estimate above we have (dividing through by $n\cdot\pi(n)$)

$$\frac{2^{-n}}{n\cdot\pi(n)} \geq \frac{L}{n\cdot\pi(n)} - [b_{2^{n-1}+1} + b_{2^{n-1}+2} + \cdots + b_{2^n}] \geq 0$$

i.e.,

$$b_{2^n} \geq Lb_n - (c_n - c_{n-1}) \geq 0.$$

Choose n_0, and set $n_1 = 2^{n_0}$, $n_2 = 2^{n_1}$, etc. Summing the previous inequality from $n = n_1 + 1$ to n_2 and noting that $b_{2^{n+1}} \leq \frac{1}{2} b_{2^n}$ gives

$$\tfrac{1}{2}b_{n_2} + \tfrac{1}{4}b_{n_2} + \tfrac{1}{8}b_{n_2} \geq Lb_{n_1+1} + Lb_{n_1+2} + \cdots + Lb_{n_2} + (c_{n_2} - c_{n_1+1}) \geq 0,$$

so we have

$$b_{n_2} \geq L(c_{n_1} - c_{n_0}) - (c_{n_2} - c_{n_1}) \geq 0.$$

Similarly,

$$b_{n_3} \geq L(c_{n_2} - c_{n_1}) - (c_{n_3} - c_{n_2}) \geq 0,$$
$$b_{n_4} \geq L(c_{n_3} - c_{n_2}) - (c_{n_4} - c_{n_3}) \geq 0,$$

and so on. Successively multiplying by L and adding the next inequality gives the following sequence of inequalities

$$b_{n_2} \geq L(c_{n_1} - c_{n_0}) - (c_{n_2} - c_{n_1}) \geq 0,$$
$$b_{n_3} + Lb_{n_2} \geq L^2(c_{n_1} - c_{n_0}) - (c_{n_3} - c_{n_2}) \geq 0,$$

$$b_{n_4} + Lb_{n_3} + L^2 b_{n_2} \geq L^3(c_{n_1} - c_{n_0}) - (c_{n_4} - c_{n_3}) \geq 0,$$

and so on. Summing these and noting that $1 + L + L^2 + \cdots = 1/(1-L)$ gives

$$\frac{b_{n_2} + b_{n_3} + b_{n_4} + \cdots}{1-L} \geq \frac{L}{1-L}(c_{n_1} - c_{n_0}) - \left(\sum_n b_n - c_{n_1}\right) \geq 0,$$

whence

$$\left(1 + \frac{2}{n_3}\right) \cdot \frac{b_{n_2}}{1-L} \geq \frac{c_{n_1} - Lc_{n_0}}{1-L} - \sum_n b_n \geq 0.$$

So, for example, setting $n_0 = 4$ so that $n_2 = 65536$ gives the value of the sum of $\Sigma_n b_n = 2.403448\ldots$ with an accuracy of over six decimal places. The same methods can also be used to evaluate the convergent series introduced earlier, for example

$$\sum_n \frac{1}{n^2} = 1.644934\ldots; \sum_n \frac{1}{n \cdot [l(n)]^2} = 1.910214\ldots;$$

$$\sum_n \frac{1}{n \cdot l(n) \cdot [l(l(n))]^2} = 2.068641\ldots$$

If further places are required for $\Sigma_n b_n$ (so that we need accurate values for c_{32}, c_{64}, etc.) then we can estimate $c_{n+1} - c_n$ to within 2^{-3n} as follows. Recall that

$$c_n - c_{n-1} = \left[\frac{1}{2^{n-1}+1} + \frac{1}{2^{n-1}+2} + \cdots + \frac{1}{2^n}\right] b_n.$$

By using Simpson's rule

$$\begin{aligned} 3L &= \left[\frac{1}{2^{n-1}} + \frac{4}{2^{n-1}+1} + \frac{2}{2^{n-1}+2} + \frac{4}{2^{n-1}+3} + \cdots + \frac{4}{2^n - 1} + \frac{1}{2^n}\right] \\ &= \frac{1}{2^{n-1}} + 4\left[\frac{1}{2^{n-1}+1} + \frac{1}{2^{n-1}+2} + \cdots + \frac{1}{2^n}\right] \\ &\quad - \left[\frac{1}{2^{n-2}+1} + \frac{1}{2^{n-2}+2} + \cdots + \frac{1}{2^{n-1}}\right] - \frac{1}{2^n} \end{aligned}$$

to within order 2^{-3n}, so we derive the recurrence relation

$$c_n - c_{n-1} = \frac{b_n}{4}\left[\frac{c_{n-1} - c_{n-2}}{b_{n-1}} + 3L - 2^{-n}\right]$$

to the order $2^{-3n}n^{-1}$. For example, if we require the value of $\Sigma_n b_n$ to 12 decimal places, we can calculate c_{13}, c_{14} directly, then use the recurrence relations to calculate $c_{32} = 0.000722028313\ldots$, then use the previous estimate with $n_0 = 5$ to give the value of the sum to the required accuracy.

It is worth noting that partitioning the series $\Sigma 1/n^2$ into chunks according to the rule $n_1 = 2n_0$, $n_2 = 2n_1$, etc. and summing each of the individual chunks also produces an (approximately) geometrically decreasing sequence of chunk sums. The condensation test corresponds to exponential increase in chunk length. The series Σb_n considered here requires more drastic (i.e. super exponential) growth of chunk lengths in order to obtain a sustainable geometric rate of decay for the sequence of chunk sums. This is a slightly counter-intuitive form of the general observation that the more slowly convergent the series, the higher the order of condensation required to reduce it to the geometric case, and is the basis of some of the results in the analysis of Tsierelsen spaces (see [**1**]).

The fact that $2 < e$ (so that $L < 1$) is vital to convergence. If we define $l_a(n)$ to be the integer ceiling of $\log_a n$ and $\pi_a(n) = l_a(n) \cdot \pi_a(l_a(n))$ then methods similar to those used earlier show that $\Sigma_n[n \cdot \pi_a(n)]^{-1}$ converges if, and only if, $a < e$. The gentle reader is invited to set $a = e$ and modify the denominator so as to find a series that requires an even higher order of condensation to establish convergence.

REFERENCES

1. Peter Casazza and Thaddeus Shura, *Tsirelson's Space*, Lecture Notes in Mathematics #1363, Springer Verlag, New York, 1989.
2. W. L. Ferrar, *A Text-book of Convergence*, Clarendon, Oxford, 1938.
3. Konrad Knopp, *Theory and Application of Infinite Series*, Dover Publications, Inc., Mineola, NY, 1990.

On Characterizations of the Gamma Function

YUAN-YUAN SHEN
Tunghai University
Taichung, Taiwan 40704

1. Introduction It is well known that the gamma function $\Gamma(x) > 0$ on $(0, \infty)$ satisfies the functional equation $\Gamma(x+1) = x\Gamma(x)$ and the initial condition $\Gamma(1) = 1$. However, these two properties do not characterize the gamma function. Rather surprisingly, the additional assumption of the convexity of $\log \Gamma(x)$ is sufficient for a characterization, a fact discovered by Bohr and Mollerup [**1**]. For a proof, see Artin's book [**4**, **5**] or Rudin's book [**6**], or the last section of this paper. Note that the initial condition in the characterization is not essential, for if f is a positive function on $(0, \infty)$ such that $f(x+1) = xf(x)$ then $g(x) = f(1)^{-1}f(x)$ is a positive function that satisfies the same functional equation and $g(1) = 1$.

A second characterization formulated and proved by Laugwitz and Rodewald [**2**] says that the convexity of $\log \Gamma(x)$ can be replaced by the property, call it property (L), that the function $L(x) = \log \Gamma(x+1)$ satisfies

$$L(n+x) = L(n) + x\log(n+1) + r_n(x), \quad \text{where } r_n(x) \to 0 \text{ as } n \to \infty. \qquad \text{(L)}$$

However, they did not show how this property is related to the convexity of $\log \Gamma(x)$. The original idea of the second characterization goes back to Euler [**3**].

In the present paper we give a third characterization of the gamma function and then show how these three characterizations are related.

2. A third characterization In property (L), the use of logarithms is not essential and without logarithms the expression on the right-hand side becomes a product instead of a sum. We might therefore expect that a modified property (L) will give us a characterization that is closer to the product expression of the gamma function. With this in mind, we modify property (L) as follows: The gamma function satisfies the following property

$$\Gamma(x+n) = \Gamma(n)n^x t_n(x), \quad \text{where } t_n(x) \to 1 \text{ as } n \to \infty.$$

it to the geometric case, and is the basis of some of the results in the analysis of Tsierelsen spaces (see [1]).

The fact that $2 < e$ (so that $L < 1$) is vital to convergence. If we define $l_a(n)$ to be the integer ceiling of $\log_a n$ and $\pi_a(n) = l_a(n) \cdot \pi_a(l_a(n))$ then methods similar to those used earlier show that $\Sigma_n [n \cdot \pi_a(n)]^{-1}$ converges if, and only if, $a < e$. The gentle reader is invited to set $a = e$ and modify the denominator so as to find a series that requires an even higher order of condensation to establish convergence.

REFERENCES

1. Peter Casazza and Thaddeus Shura, *Tsirelson's Space*, Lecture Notes in Mathematics #1363, Springer Verlag, New York, 1989.
2. W. L. Ferrar, *A Text-book of Convergence*, Clarendon, Oxford, 1938.
3. Konrad Knopp, *Theory and Application of Infinite Series*, Dover Publications, Inc., Mineola, NY, 1990.

On Characterizations of the Gamma Function

YUAN-YUAN SHEN
Tunghai University
Taichung, Taiwan 40704

1. Introduction It is well known that the gamma function $\Gamma(x) > 0$ on $(0, \infty)$ satisfies the functional equation $\Gamma(x+1) = x\Gamma(x)$ and the initial condition $\Gamma(1) = 1$. However, these two properties do not characterize the gamma function. Rather surprisingly, the additional assumption of the convexity of $\log \Gamma(x)$ is sufficient for a characterization, a fact discovered by Bohr and Mollerup [1]. For a proof, see Artin's book [4, 5] or Rudin's book [6], or the last section of this paper. Note that the initial condition in the characterization is not essential, for if f is a positive function on $(0, \infty)$ such that $f(x+1) = xf(x)$ then $g(x) = f(1)^{-1} f(x)$ is a positive function that satisfies the same functional equation and $g(1) = 1$.

A second characterization formulated and proved by Laugwitz and Rodewald [2] says that the convexity of $\log \Gamma(x)$ can be replaced by the property, call it property (L), that the function $L(x) = \log \Gamma(x+1)$ satisfies

$$L(n+x) = L(n) + x \log(n+1) + r_n(x), \quad \text{where } r_n(x) \to 0 \text{ as } n \to \infty. \quad \text{(L)}$$

However, they did not show how this property is related to the convexity of $\log \Gamma(x)$. The original idea of the second characterization goes back to Euler [3].

In the present paper we give a third characterization of the gamma function and then show how these three characterizations are related.

2. A third characterization In property (L), the use of logarithms is not essential and without logarithms the expression on the right-hand side becomes a product instead of a sum. We might therefore expect that a modified property (L) will give us a characterization that is closer to the product expression of the gamma function. With this in mind, we modify property (L) as follows: The gamma function satisfies the following property

$$\Gamma(x+n) = \Gamma(n) n^x t_n(x), \quad \text{where } t_n(x) \to 1 \text{ as } n \to \infty.$$

THEOREM 1. *There exists a unique function $f(x) > 0$ on $(0, \infty)$ that satisfies the following three properties*:

(a) $f(1) = 1$;
(b) $f(x+1) = xf(x)$;
(c) $f(x+n) = f(n)n^x t_n(x)$, *where* $t_n(x) \to 1$ *as* $n \to \infty$.

DEFINITION. *For each positive integer n, we define the function Γ_n on $(0, \infty)$ by*

$$\Gamma_n(x) = \frac{n^x n!}{x(x+1)\cdots(x+n)}, \quad x > 0.$$

LEMMA. *The sequence $(\Gamma_n(x))$ of functions on $(0, \infty)$ converges for any $x > 0$.*

Proof. Taking logarithms, we have

$$\begin{aligned}
\log \Gamma_n(x) &= x \log n + \sum_{k=1}^{n} \log k - \log x - \sum_{k=1}^{n} \log(x+k) \\
&= x \log n - \log x - \sum_{k=1}^{n} \log\left(1 + \frac{x}{k}\right) \\
&= -\log x - x\left[\sum_{k=1}^{n} \frac{1}{k} - \log n\right] + \sum_{k=1}^{n} \left[\frac{x}{k} - \log\left(1 + \frac{x}{k}\right)\right] \\
&= -\log x - x\gamma_n + c_n(x),
\end{aligned}$$

where $\gamma_n = \sum_{k=1}^{n} \frac{1}{k} - \log n$, and $c_n(x) = \sum_{k=1}^{n} [\frac{x}{k} - \log(1 + \frac{x}{k})]$. It is well known that (γ_n) converges to Euler's constant $\gamma \approx 0.577\ldots$. Also, the sequence $(c_n(x))$ converges, since for $k > x > 0$,

$$0 < \frac{x}{k} - \log\left(1 + \frac{x}{k}\right) = \frac{x}{k} - \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i}\left(\frac{x}{k}\right)^i \leq \frac{x^2}{2k^2}.$$

Thus, the sequence $(\log \Gamma_n(x))$ converges and hence so does the sequence $(\Gamma_n(x))$ for $x > 0$. This completes the proof of the lemma.

Remark. In fact, the limit function of the above sequence is the product expression of the gamma function (see [4, **5**]). Therefore we have

$$\Gamma(x) = \lim_{n \to \infty} \frac{n^x n!}{x(x+1)\cdots(x+n)}, \quad \text{for } x > 0. \tag{1}$$

Proof of Theorem 1. First we prove that $\Gamma(x)$ in (1) satisfies (a)–(c).
(a) $\Gamma(1) = 1$, since

$$\Gamma(1) = \lim_{n \to \infty} \frac{n^1 n!}{1(1+1)\cdots(1+n)} = \lim_{n \to \infty} \frac{n}{1+n} = 1.$$

(b) Γ satisfies the functional equation, since

$$\begin{aligned}
\Gamma(x+1) &= \lim_{n \to \infty} \frac{n^{x+1} n!}{(x+1)(x+2)\cdots(x+1+n)} \\
&= \lim_{n \to \infty} \frac{nx}{x+1+n} \cdot \lim_{n \to \infty} \frac{n^x n!}{x(x+1)\cdots(x+n)} \\
&= x\Gamma(x).
\end{aligned}$$

As a consequence of these two properties, we get $\Gamma(n) = (n-1)!$.

(c) Let $s_n(x) = \Gamma(x)/\Gamma_n(x)$. Then $\Gamma(x) = \Gamma_n(x)s_n(x)$, and $\lim_{n\to\infty} s_n(x) = 1$.

For natural n and real $x > 0$, we apply (b) n times to get

$$\begin{aligned}\Gamma(x+n) &= [(x+n-1)\cdots(x+1)x]\cdot\Gamma(x)\\ &= \frac{(x+n)\cdots(x+1)x}{x+n}\cdot\frac{n^x n!}{x(x+1)\cdots(x+n)}\cdot s_n(x)\\ &= n^x\Gamma(n)t_n(x),\end{aligned}$$

where $t_n(x) = (n/(x+n))\cdot s_n(x)$. Thus, $\Gamma(x+n) = n^x\Gamma(n)t_n(x)$ and $t_n(x) \to 1$ as $n \to \infty$.

To show uniqueness, we assume $f(x)$ is a function that satisfies (a)–(c). From properties (a) and (b), we have

$$f(n) = (n-1)!, \tag{2}$$

$$f(x+n) = (x+n-1)(x+n-2)\cdots(x+1)xf(x). \tag{3}$$

Combining (3), property (c), and (2) together, we have

$$f(x) = \frac{x^n(n-1)!}{x(x+1)\cdots(x+n-1)}\cdot t_n(x) = \Gamma_n(x)\cdot s_n(x),$$

where $s_n(x) = ((x+n)/n)t_n(x) \to 1$ as $n \to \infty$. Therefore $f(x) = \Gamma(x)$ and hence f is uniquely determined. This completes the proof of the third characterization of the gamma function.

3. How are these characterizations related? To simplify our discussion, we adopt the following terminology.

DEFINITION. *By a PG function (pre-gamma function), we mean a positive function f on $(0,\infty)$ that satisfies the functional equation $f(x+1) = xf(x)$.*

Remark. For a PG function f, we may assume $f(1) = 1$, since if g is a PG function then $f(x) = g(1)^{-1}g(x)$ is also a PG function such that $f(1) = 1$. Now we can rephrase what we have so far on characterizations of the gamma function.

CHARACTERIZATIONS. *If f is a PG function such that*

(C) $\log f$ *is convex on* $(0,\infty)$,

or

(L) $L(n+x) = L(n) + x\log(n+1) + r_n(x)$,

where $L(x) = \log f(x+1)$ and $r_n(x) \to 0$ as $n \to \infty$, or

(P) $f(n+x) = f(n)n^x t_n(x)$,

where $t_n(x) \to 1$ as $n \to \infty$, then $f(x) = c\Gamma(x)$, for some constant c.

Remarks. (a) It is easy to see that the constant c in each characterization is simply $f(1)$. In other words, any PG function f with $f(1) = 1$ that satisfies either (C) or (L) or (P) must be the gamma function.

(b) In the previous section we showed that (P) characterizes the gamma function. Therefore it suffices to show that properties (C), (L), and (P) are equivalent to one another for a PG function. To do this, we need three basic facts about convex functions (see [**4**, **5**]).

(1) *If f is convex on (a,b) and if $x<y$, $x, y \in (a,b)$, then*

$$\frac{f(x)-f(c)}{x-c} \le \frac{f(y)-f(c)}{y-c}$$

for any $c \in (a,b)$.

(2) *The limit function of a convergent sequence of convex functions is convex.*

(3) *Let g be a twice-differentiable function on (a,b). Then g is convex on (a,b) if, and only if, $g''(x) > 0$ for all $x \in (a,b)$.*

THEOREM 2. *For a PG function f, the properties* (C), (L), *and* (P) *are equivalent.*

Proof. (a) (P) $\Leftrightarrow$ (L). We have

$$\begin{aligned}
&(\mathrm{P})\\
&\Leftrightarrow f(x+(n+1)) = f(n+1)(n+1)^x t_{n+1}(x),\ t_{n+1}(x) \to 1\\
&\Leftrightarrow \log f((x+n)+1) = \log f(n+1) + x\log(n+1) + \log t_{n+1}(x),\ t_{n+1}(x) \to 1\\
&\Leftrightarrow L(x+n) = L(n) + x\log(n+1) + r_n(x),\ r_n(x) \to 0\\
&\Leftrightarrow (\mathrm{L}).
\end{aligned}$$

(b) (C) $\Rightarrow$ (P). Let $m < x \le m+1$, where $m = 0,1,2\dots$. For any natural n, $n+m-1 < n+m < n+x \le n+m+1$. The convexity of $\log f$ and (1) above give us (we write $L_m = \log f(n+m)$)

$$\frac{L_m - L_{m-1}}{n+m-(n+m-1)} \le \frac{\log f(n+x) - \log f(n+m)}{(n+x)-(n+m)} \le \frac{L_{m+1}-L_m}{(n+m+1)-(n+m)}$$

$$\Leftrightarrow (x-m)\log(n+m-1) \le \log\left(\frac{f(n+x)}{f(n+m)}\right) \le (x-m)\log(n+m)$$

$$\Leftrightarrow (n+m-1)^{x-m} \le \frac{f(n+x)}{(n+m-1)(n+m-2)\cdots nf(n)} \le (n+m)^{x-m}$$

$$\Leftrightarrow \left(1+\frac{m-1}{n}\right)^x T_{m-1} \le \frac{f(n+x)}{f(n)n^x} \le \left(1+\frac{m}{n}\right)^x T_m,$$

where

$$T_m = \left(1-\frac{1}{n+m}\right)\left(1-\frac{2}{n+m}\right)\cdots\left(1-\frac{m}{n+m}\right).$$

Therefore, we have

$$\lim_{n\to\infty} \frac{f(n+x)}{f(n)n^x} = 1,$$

by the squeezing theorem. If we let

$$t_n(x) = \frac{f(n+x)}{f(n)n^x},$$

then

$$f(n+x) = f(n)n^x t_n(x),$$

where $t_n(x) \to 1$ as $n \to \infty$. This proves that f satisfies (P).

(c) (P) $\Rightarrow$ (C). From the uniqueness part of the proof of Theorem 1 we have

$$f(x)=f(1)\lim_{n\to\infty}\Gamma_n(x).$$

By (2) above, it suffices to show that $\log\Gamma_n(x)$ is convex. Now

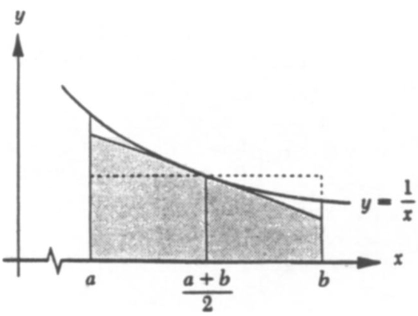
$$(\log\Gamma_n(x))''=\frac{1}{x^2}+\frac{1}{(x+1)^2}+\cdots+\frac{1}{(x+n)^2}>0.$$

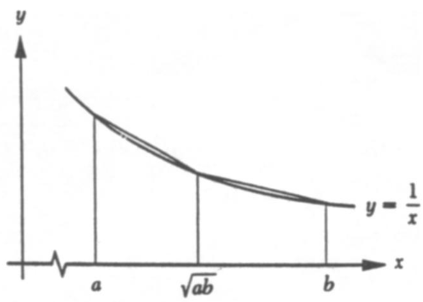
By (3) above, $\log\Gamma_n(x)$ is convex and so is $\log f$. This completes the proof.

REFERENCES

1. H. Bohr and J. Mollerup, *Laerebog i matematisk Analyse*, Kopenhage (1922), Vol. III, pp. 149–164.
2. Detlef Laugwitz and Bernd Rodewald, A simple characterization of the gamma function, *Amer. Math. Monthly*, 94 (1987), 534–536.
3. Leonhard Euler, *Institutiones calculi differentialis*, Teubner, 1980; Leonhardi Euleri opera omnia, 10.
4. Emil Artin, *The Gamma Function*, Holt, Rinehart & Winston, Inc., New York, 1964.
5. Emil Artin, *Einführung in die Theorie der Gammafunktion*, Teubner, 1931.
6. Walter Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill Book Co., New York, 1976.

Proof without Words: The Arithmetic-Logarithmic-Geometric Mean Inequality





$$\ln b-\ln>\frac{2}{a+b}(b-a) \qquad \ln b-\ln a<\frac{ab-a}{2a\sqrt{ab}}+\frac{a-ab}{2b\sqrt{ab}}=\frac{b-a}{\sqrt{ab}}$$

$$\frac{a+b}{2}>\frac{b-a}{\ln b-\ln a} \qquad \sqrt{ab}<\frac{b-a}{\ln b-\ln a}$$

$$b>a>0\Rightarrow\frac{a+b}{2}>\frac{b-a}{\ln b-\ln a}>\sqrt{ab}$$

NOTE: Approximating the integral by inscribed and circumscribed rectangles yields *Napier's Inequality* [*College Mathematics Journal* V. 24, no. 2 (March 1993), p. 165].

—ROGER B. NELSEN
LEWIS AND CLARK COLLEGE
PORTLAND, OR 92719

(c) (P) ⇒ (C). From the uniqueness part of the proof of Theorem 1 we have

$$f(x) = f(1) \lim_{n\to\infty} \Gamma_n(x).$$

By (2) above, it suffices to show that $\log \Gamma_n(x)$ is convex. Now

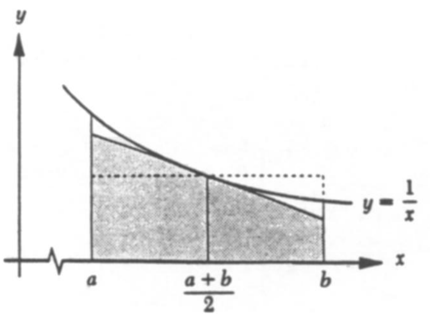
$$(\log \Gamma_n(x))'' = \frac{1}{x^2} + \frac{1}{(x+1)^2} + \cdots + \frac{1}{(x+n)^2} > 0.$$

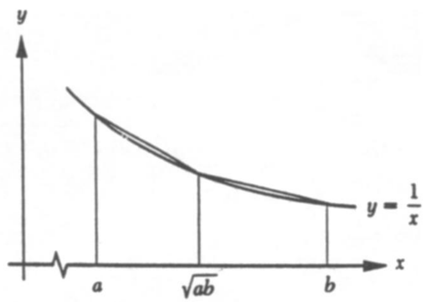
By (3) above, $\log \Gamma_n(x)$ is convex and so is $\log f$. This completes the proof.

REFERENCES

1. H. Bohr and J. Mollerup, *Laerebog i matematisk Analyse*, Kopenhage (1922), Vol. III, pp. 149–164.
2. Detlef Laugwitz and Bernd Rodewald, A simple characterization of the gamma function, *Amer. Math. Monthly*, 94 (1987), 534–536.
3. Leonhard Euler, *Institutiones calculi differentialis*, Teubner, 1980; Leonhardi Euleri opera omnia, 10.
4. Emil Artin, *The Gamma Function*, Holt, Rinehart & Winston, Inc., New York, 1964.
5. Emil Artin, *Einführung in die Theorie der Gammafunktion*, Teubner, 1931.
6. Walter Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill Book Co., New York, 1976.

Proof without Words:
The Arithmetic-Logarithmic-Geometric Mean Inequality





$$\ln b - \ln > \frac{2}{a+b}(b-a)$$

$$\frac{a+b}{2} > \frac{b-a}{\ln b - \ln a}$$

$$\ln b - \ln a < \frac{ab-a}{2a\sqrt{ab}} + \frac{a-ab}{2b\sqrt{ab}} = \frac{b-a}{\sqrt{ab}}$$

$$\sqrt{ab} < \frac{b-a}{\ln b - \ln a}$$

$$b > a > 0 \Rightarrow \frac{a+b}{2} > \frac{b-a}{\ln b - \ln a} > \sqrt{ab}$$

NOTE: Approximating the integral by inscribed and circumscribed rectangles yields *Napier's Inequality* [*College Mathematics Journal* V. 24, no. 2 (March 1993), p. 165].

—ROGER B. NELSEN
LEWIS AND CLARK COLLEGE
PORTLAND, OR 92719

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 March 1, 1996.

1479. *Proposed by Donald E. Knuth, Stanford University, Stanford, California.*

Let m_n be the maximum value of the quantity

$$\frac{x_1}{(1+x_1+x_2+\cdots+x_n)^2}+\frac{x_2}{(1+x_2+\cdots+x_n)^2}+\cdots+\frac{x_n}{(1+x_n)^2}$$

over all nonnegative real numbers $(x_1,\ldots,x_n)$. At what point(s) does the maximum occur? Express m_n in terms of m_{n-1}, and find $\lim_{n\to\infty} m_n$.

1480. *Proposed by Ron Rietz and John Holte, Gustavus Adolphus College, St. Peter, Minnesota.*

Prove that

$$\sum_{i_1=0}^{n}\sum_{i_2=0}^{i_1}\cdots\sum_{i_k=0}^{i_{k-1}} r^{i_1+i_2+\cdots+i_k}=\prod_{j=1}^{k}\frac{1-r^{n+j}}{1-r^j}$$

for $r\neq\pm1$, $k=1,2,3,\ldots$, and $n=0,1,2,\ldots$.

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 sent to George T. Gilbert, Problems Editor-Elect, Department of Mathematics, Box 32903, Texas Christian University, Fort Worth, TX 76129 or mailed electronically to g.gilbert@tcu.edu. Electronic submission of TeX input files is acceptable. Readers who use e-mail should also provide an e-mail address.

1481. *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada.*

It is known that if a point moves on a straight line with constant acceleration and s_1, s_2, s_3 are its positions at time t_1, t_2, t_3, respectively, then the constant acceleration is given by

$$2\left(\frac{(s_2-s_3)t_1+(s_3-s_1)t_2+(s_1-s_2)t_3}{(t_1-t_2)(t_2-t_3)(t_3-t_1)}\right).$$

Show that this property characterizes uniformly accelerated motion; that is, if a particle moves on a straight line and s_1, s_2, s_3 are its positions at any times t_1, t_2, t_3, respectively, then if

$$\frac{(s_2-s_3)t_1+(s_3-s_1)t_2+(s_1-s_2)t_3}{(t_1-t_2)(t_2-t_3)(t_3-t_1)}=\text{constant},$$

the motion is one of constant acceleration.

1482. *Proposed by C. F. Eaton, Pepperell, Massachusetts.*

Show that all even perfect numbers, $P>6$, are of the form $P=1+9T_n$, where T_n is a triangular number of the form $T_n=n(n+1)/2$, $n=8j+2$.

1483. *Proposed by Alexandru Teodorescu-Frumosu, student, Boston University, Boston, Massachusetts.*

Let ABC be an arbitrary triangle, and let a, b, c, be the lengths of the sides BC, AC, AB, respectively. Let M be the midpoint of the segment BC, let $\alpha=\angle BAM$, $\beta=\angle CAM$ and $x=\angle AMB$. Show that

$$\frac{b}{\sin\alpha}=\frac{a\cos x}{\sin(\alpha-\beta)}.$$

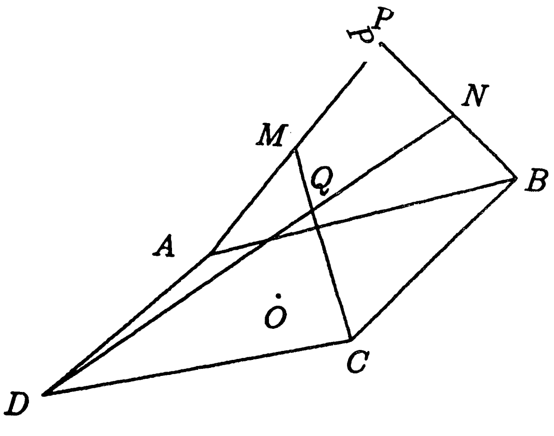
Quickies

Answers to the Quickies are on page 318.

Q838. *Proposed by Ismor Fischer, University of Wisconsin, Madison, Wisconsin.*

Let O and P be fixed points in the plane, and $ABCD$ an arbitrary parallelogram centered at O. Let $\overline{CM}$ and $\overline{DN}$ bisect $\overline{PA}$ and $\overline{PB}$, respectively, and call their point of intersection Q (see figure).

Show that points O, P, and Q are collinear, and that the location of Q on $\overline{OP}$ is fixed, independent of parallelogram $ABCD$.



Q839. *Proposed by J. C. Binz, University of Bern, Switzerland.*

Let $\mathbf{Z}_p$ be the field of integers modulo a prime number p. Given $a_0, a_1 \in \mathbf{Z}_p$ and $r + s = 1$, the recurrence $a_n = ra_{n-1} + sa_{n-2}$, $n \geq 2$, defines a periodic sequence (a_n) with period of length $L \leq p^2 - 1$. Show that, in fact, $L \leq p$.

Q840. *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada.*

If $A_1 A_2 A_3 A_4$ is a plane cyclic quadrilateral and H_i is the orthocenter of triangle $A_{i+1} A_{i+2} A_{i+3}$ where $A_{i+4} = A_i$ for $i = 1, 2, 3, 4$, prove that

(i) $\text{Area}(A_1 A_2 A_3 A_4) = \text{Area}(H_1 H_2 H_3 H_4)$, and

(ii) the lines $A_1 H_1$, $A_2 H_2$, $A_3 H_3$, $A_4 H_4$ are concurrent.

Solutions

Names drawn from a hat **October 1994**

1454. *Proposed by Barry Cipra, Northfield, Minnesota.*

Suppose n people put their names in a hat (on slips of paper), and the names are redistributed at random (using the uniform probability distribution on the space of permutations). Those who receive their own name drop out, while the rest repeat the procedure. On average, how many rounds will it take until all have gotten their own name back?

I. *Solution by Michael Reid, student, University of California Berkeley, Berkeley, California.*

On average, it will take n rounds.

Note that the probability that the first person doesn't receive his own name during the first k rounds is $\leq ((n - 1/n))^k$. Thus the probability that this process lasts more than k rounds, which is the probability that some person doesn't receive his own name during the first k rounds, is $\leq n((n - 1/n))^k$. From this, it is easy to see that this process terminates with probability 1, and that the expected number of rounds exists and is finite.

LEMMA. *The expected number of people who receive their own names in any round is* 1.

Proof. This expected value is just

$$\sum_{i=1}^{n} (\text{probability that the } i\text{th person receives his own name}) = \sum_{i=1}^{n} \frac{1}{n} = 1.$$

PROPOSITION. *The expected number of rounds is* n.

Proof. We'll prove this by induction on n. This is immediate for $n = 0$ and $n = 1$. Now suppose that it holds for all $n < m$. Let E_k denote the expected number of rounds for k people, and let p_i denote the probability that i out of m people receive their own names. Then we have

$$E_m = 1 + p_0 E_m + p_1 E_{m-1} + \cdots + p_m E_0,$$

or,

$$(1 - p_0) E_m = 1 + p_1 E_{m-1} + p_2 E_{m-2} + \cdots + p_m E_0.$$

From the induction hypothesis, this becomes

$$\begin{aligned}(1-p_0)E_m &= 1+p_1(m-1)+p_2(m-2)+\cdots+p_m(m-m)\\ &= 1+m(p_1+p_2+\cdots+p_m)-(p_1+2p_2+\cdots+mp_m)\\ &= 1+m(1-p_0)-(p_1+2p_2+\cdots+mp_m).\end{aligned}$$

From the Lemma, we have $p_1+2p_2+\cdots+mp_m=1$, so the above becomes

$$(1-p_0)E_m = m(1-p_0).$$

Since $1-p_0 \neq 0$, we have $E_m = m$, as desired.

II. *Solution by Jerrold W. Grossman, Oakland University, Rochester, Michigan.*

Let x_n be the expected number of rounds until all people have received their own names. We claim that $x_n = n$ and prove this by induction on n, the result being trivial for $n=0$.

The expected length of the game with n players is the weighted average, over all outcomes of the first round that leave i players still in the game, of $1+x_i$. Let q_i be the number of ways to leave i players (that is, q_i is the number of permutations of n objects that fix $n-i$ of them). Then we have

$$\sum_{i=0}^{n} \frac{q_i}{n!}(1+x_i) = x_n.$$

The claim therefore follows from the inductive hypothesis once we prove the identity

$$\sum_{i=0}^{n} q_i(1+i) = n\cdot n!.$$

Both sides count the number of ways to choose a permutation π of $\{1,2,\ldots,n\}$ and optionally single out a particular nonfixed point of π. Indeed, the left side counts this by focusing on the number of fixed points. The right side equals $n!+n(n-1)(n-1)!$: There are $n!$ ways to specify a permutation and choose not to distinguish a nonfixed point, and there are $n(n-1)(n-1)!$ ways to select a nonfixed point k, choose $\pi(k)\neq k$, and then select $\pi(j)$ for all $j\neq k$.

Also solved by Robert A. Agnew, D. M. Bloom, Michael H. Andreoli, David Callan, Con Amore Problem Group (two solutions, Denmark), Curtis Cooper, Robert L. Doucette, Peter Flusser, Herbert Gintis, Richard Holzsager, Barthel Wayne Huff, R. Daniel Hurwitz, Kee-Wai Lau (Hong Kong), Robert A. Leslie, O. P. Lossers (The Netherlands), Rick Mabry, F. C. Rembis, Gordon Rice, Wolfgang Rolke (two solutions, Puerto Rico), Gregg Rosenkranz (student), Nicholas C. Singer, Paul K. Stockmeyer, Frank Thorne, Michael Vowe (two solutions, Switzerland), Dennis Walsh, WMC Problems Group, Homer White, David Zhu, and the proposer. There was one incomplete solution.

Equilateral triangles in a hexagon **October 1994**

1455. *Proposed by Jiro Fukuta, Gifu-ken, Japan.*

In a hexagon, $A_1A_2A_3A_4A_5A_6$ inscribed in a circle with center O, let M_i, $i=1,2,\ldots,6$, be the midpoints of the sides A_iA_{i+1}, where $A_7=A_1$. Prove that if $\Delta M_1M_3M_5$ and $\Delta M_2M_4M_6$ are equilateral, $\Delta A_1A_3A_5$ and $\Delta A_2A_4A_6$ are also equilateral.

Solution by Ted Zerger, Kansas Wesleyan University, Salina, Kansas.

The condition that the hexagon be cyclic is unnecessary.

Let the A_i's and M_i's be represented by complex numbers.
Since $M_i = (A_i + A_{i+1})/2$, it follows that

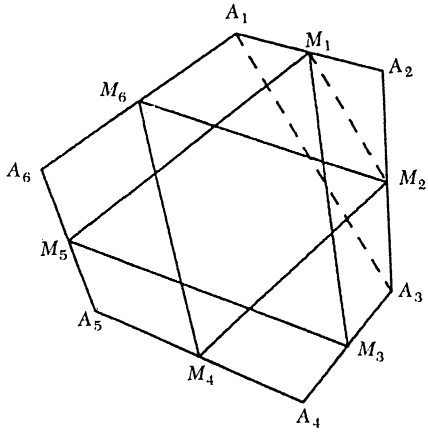
$$M_1 + M_3 + M_5 = \tfrac{1}{2}(A_1 + A_2 + A_3 + A_4 + A_5 + A_6)$$

and

$$M_2 + M_4 + M_6 = \tfrac{1}{2}(A_1 + A_2 + A_3 + A_4 + A_5 + A_6),$$

and therefore $\frac{1}{3}(M_1 + M_3 + M_5) = \frac{1}{3}(M_2 + M_4 + M_6)$. Thus, $\Delta M_1 M_3 M_5$ and $\Delta M_2 M_4 M_6$ have the same centroid, and, since the triangles are equilateral, their circumcircles are concentric.

The symmetry of the figure implies that $|M_2 - M_1| = |M_4 - M_3| = |M_6 - M_5|$. Moreover, $|M_2 - M_1| = \frac{1}{2}|A_3 - A_1|$ (see figure).
In the same way, $|M_4 - M_3| = \frac{1}{2}|A_5 - A_3|$ and $|M_6 - M_5| = \frac{1}{2}|A_5 - A_1|$. Thus $|A_3 - A_1| = |A_5 - A_3| = |A_5 - A_1|$, and $\Delta A_1 A_3 A_5$ is equilateral. Similarly, $\Delta A_2 A_4 A_6$ is equilateral.



Also solved by Con Amore Problem Group (Denmark), Hans Kappus (Switzerland), Murray Klamkin (Canada), O. P. Lossers (The Netherlands), Michael Reid (student), Szabó Szila'rd (Hungary), Robert L. Young, David Zhu, and the proposer.

Klamkin *generalized the result to an arbitrary* $2n$*-gon in the plane. Several correspondents noted that the labels must be placed consecutively in a clockwise or counterclockwise manner.*

Recurrence in which positivity implies uniqueness **October 1994**

1456. *Proposed by Howard Morris, Chatsworth, California.*

Show that the only sequence of numbers (α_i) that satisfies the conditions

(i) $\alpha_i > 0$ for all $i \geq 1$, and

(ii) $$\alpha_{i-1} = \frac{i\alpha_i + 1}{\alpha_i + i} \quad \text{for all } i > 0,$$

is the sequence $\alpha_i = 1$ for all i.

I. *Solution by O. P. Lossers, Technical University Eindhoven, Eindhoven, The Netherlands.*

A straightforward calculation shows that

$$\frac{\alpha_i - 1}{\alpha_i + 1} = \left(\frac{i+1}{i-1}\right)\left(\frac{\alpha_{i-1} - 1}{\alpha_{i-1} + 1}\right) \quad \text{for all } i > 1.$$

Hence for all α_1, except $\alpha_1 = 1$, there exists an i such that $|\alpha_i - 1| > |\alpha_i + 1|$, which is equivalent to $\Re(\alpha_i) < 0$.

II. *Solution by David Zhu, Jet Propulsion Laboratory, Pasadena, California.*

It is clear that if $\alpha_i = 1$ for some $i > 0$, then $\alpha_i = 1$ for all i.

From (ii) we find that

$$\alpha_i = \frac{i\alpha_{i-1} - 1}{i - \alpha_{i-1}}, \quad \text{for } i > 1.$$

a. Suppose that $\alpha_{i-1} < 1$ for some $i > 1$. Then

$$\alpha_{i-1} - \alpha_i = \frac{1 - \alpha_{i-1}^2}{i - \alpha_{i-1}} > \frac{1 - \alpha_{i-1}^2}{i} > 0,$$

and similarly,

$$\alpha_i - \alpha_{i+1} > \frac{1 - \alpha_i^2}{i+1} > 0,$$

$$\vdots$$

$$\alpha_{k-1} - \alpha_k > \frac{1 - \alpha_{k-1}^2}{k} > 0.$$

Thus, for all $k \geq i$,

$$\begin{aligned}\alpha_k &< \alpha_{i-1} - \frac{1 - \alpha_{i-1}^2}{i} - \frac{1 - \alpha_i^2}{i+1} - \cdots - \frac{1 - \alpha_{k-1}^2}{k} \\ &< \alpha_{i-1} - \left(1 - \alpha_{i-1}^2\right)\left(\frac{1}{i} + \frac{1}{i+1} + \cdots + \frac{1}{k}\right).\end{aligned}$$

It follows that for large k, $\alpha_k < 0$, which contradicts (i).

b. Suppose that $\alpha_i > 1$ for all $i > 0$. Let $\beta_i = 1/\alpha_i$. Then

$$\beta_i = \frac{i\beta_{i-1} - 1}{i - \beta_{i-1}}, \quad \text{for } i > 1,$$

and the same argument as in (a) shows that $\beta_k < 0$ for sufficiently large k, a contradiction.

We must conclude that $\alpha_i = 1$ for all i.

III. *Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada.*

If we let $\alpha_n = x_n + 1$, (ii) becomes

$$\frac{n+1}{x_n} - \frac{n-1}{x_{n-1}} = -1.$$

Then, with $x_n = n(n+1)F_n$, we get the telescoping difference equation

$$\frac{1}{F_n} - \frac{1}{F_{n-1}} = -n.$$

Hence,

$$F_n = \frac{2F_1}{2 - (n+2)(n-1)F_1},$$

and thus

$$\alpha_n = \frac{2 + (n^2 + n + 2)F_1}{2 - (n^2 + n - 2)F_1}.$$

Since $2F_1 = \alpha_1 - 1$, (i) implies that $F_1 > -1/2$. Since α_n is positive for $n = 2, 3, \ldots,$ it must be the case that $F_1 = 0$, or equivalently, $\alpha_n = 1$ for $n = 1, 2, \ldots$. Finally, from (ii), with $i = 1$, we get $\alpha_0 = 1$.

Also solved by Sinefakopoulos Achilleas (Greece), Michael H. Andreoli, Michael Bertrand, J. C. Binz (Switzerland), Martin Burger (student, Austria), David Callan, D. K. Cohoon, Con Amore Problem Group (Denmark), Paul Deiermann, Robert L. Doucette, Mathew Foss, Herbert Gintis, Joe Howard, Pavlos B. Konstadinidis (student), Kee-Wai Lau (Hong Kong), Detlef Laugwitz (two solutions, Germany), Henry S. Lieberman, David E. Manes, Beatriz Margolis (France), Hugh McGuire (student), Kandasamy Muthuvel, Jeremy Ottenstein (Israel), Michael Reid, F. C. Rembis, Michael Rierson (student), Robert Schneck (student), Heinz-Jürgen Seiffert (Germany), Nicholas C. Singer, Allan Swett, Michael Vowe (Switzerland), WMC Problems Group, Yan-Loi Wong (Singapore), Yongzhi Yang, Paul J. Zwier, and the proposer.

Cohoon *generalized the result to the recursion* $\alpha_{i-1} = (c_i\alpha_i + 1)/(\alpha_i + c_i)$, *where the* c_i *satisfy* $1 < c_i < \kappa i$ *for some constant* κ. Lossers' *solution extends to this case with only superficial changes.*

The center of a sliced pizza — October 1994

1457. *Proposed by Larry Carter, IBM Watson Research Center, Yorktown Heights, New York, John Duncan, University of Arkansas, Fayetteville, Arkansas, and Stan Wagon, Macalester College, St. Paul, Minnesota.*

a. For a point P inside a circle draw three chords through P making six 60° angles at P and form two regions by coloring the six "pizza slices" alternately black and white. Prove that the region containing the center has the larger area.

b*. Prove that if five chords make ten 36° angles at P, then the region containing the center has the lesser area.

Solution by Paul Deiermann and Rick Mabry, Louisiana State University in Shreveport, Shreveport, Louisiana.

a. Our pizza is the unit disk. Without loss of generality, let P have polar coordinates (r, θ), with $0 \le \theta < \pi/3$, and let one of the cuts be parallel to the x-axis. (Any other configuration can be rotated into one like this.) We prove the result by showing that the region containing the origin, call it $\mathscr{R}(|||)$ (the shaded region in FIGURE 1), comprises no less than half the area of the pizza. (We will not always explicitly distinguish regions from their areas, it being clear from the context.) Let P^* denote the intersection of the x-axis with the cut in the direction of $\pi/3$. Now take a fresh pizza and make cuts parallel to those on the first, but through P^*. Then $\mathscr{R}(\equiv)$, the region on the second pizza corresponding to $\mathscr{R}(|||)$, is exactly one half of the pizza. ("Pie over two," see FIGURE 2.) If $\mathscr{R}(|||)$ is bigger than or equal to $\mathscr{R}(\equiv)$, we're done.

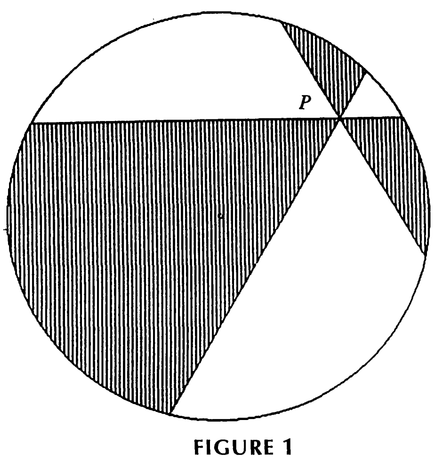


FIGURE 1

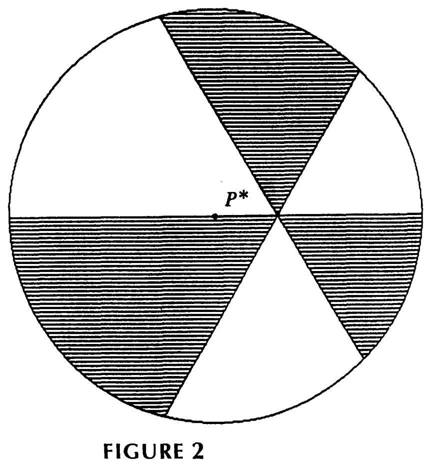


FIGURE 2

Overlay the two pizzas, as in FIGURE 3. (This picture is the proof.) We disregard the doubly shaded regions, since they occur in both pizzas, and compare what is left, the two "strips," $ABCD$ and $A'B'C'D'$, minus their intersection. But the intersection, being common to both, can be ignored and we are left comparing the filled strips. These strips are of equal thickness, but $ABCD$ is clearly the larger, since $A'B'C'D'$ sits at a greater distance from the center.

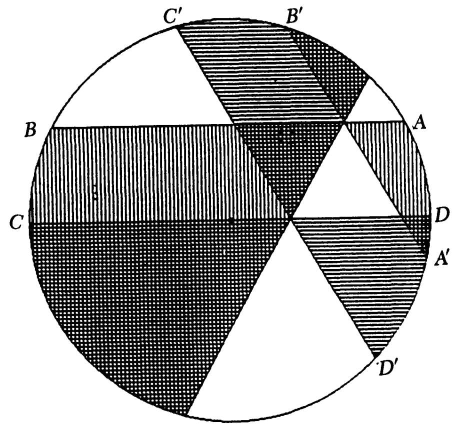


FIGURE 3

b*. Begin the same way as above, but restrict P so that $0 \leq \theta < \pi/5$, and let P^* denote the intersection of the x-axis with the cut in the direction of $\pi/5$. Form the two pizzas and overlay them, as in FIGURE 4.

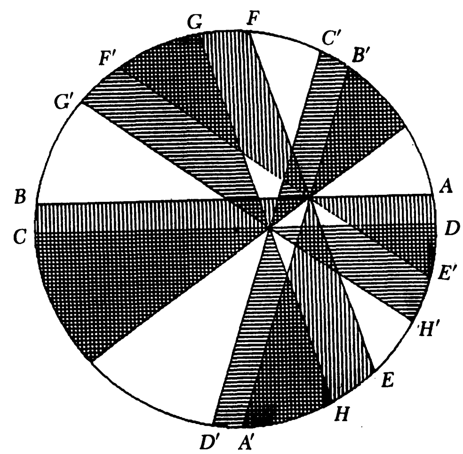


FIGURE 4

As above, the problem amounts to comparing the sum $ABCD + EFGH$ of the areas of the strips on the first pizza with the sum $A'B'C'D' + E'F'G'H'$ of the areas of the strips on the second. This comparison is no longer trivial, as in part a., for although $ABCD$ and $A'B'C'D'$ have the same "width," as $EFGH$ and $E'F'G'H'$, the "lengths" of these strips do not cooperate—$ABCD$ is larger than $A'B'C'D'$ while $EFGH$ is smaller than $E'F'G'H'$.

To show that $ABCD + EFGH \leq A'B'C'D' + E'F'G'H'$, we need to find the area of a strip, by which we mean the difference of two segments of the unit circle. Consider a strip at perpendicular distance D from the center of the circle, and having width W, as in FIGURE 5. Its area is easily found to be $s(D) - s(D+W)$, where

$s(x) = \cos^{-1}(x) - x\sqrt{1-x^2}$ is the area of the segment at distance x. (This is the sector OPQ minus the triangle OPQ in FIGURE 6.)

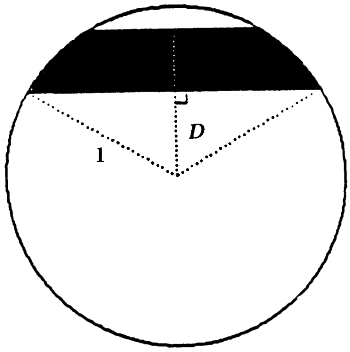


FIGURE 5

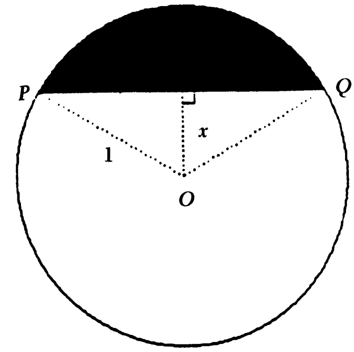


FIGURE 6

Returning to FIGURE 4, let d denote the distance from P to P^* and w the distance from P^* to the center, and let $d_j = d\sin(\pi j/5)$ and $w_j = w\sin(\pi j/5)$, for $j = 1, 2$. It is easy to see that

$$ABCD = s(0) - s(d_1), \qquad A'B'C'D' = s(w_2) - s(d_1 + w_2),$$
$$EFGH = s(w_2) - s(d_2 + w_2), \qquad E'F'G'H' = s(w_1) - s(d_2 + w_1),$$

and so $(ABCD + EFGH) - (A'B'C'D' + E'F'G'H')$ equals

$$s(0) + s(d_1 + w_2) + s(d_2 + w_1) - s(d_2 + w_2) - s(w_1) - s(d_1). \tag{1}$$

Next we observe that $s'(x) = -2\sqrt{1-x^2}$, so we have, with the aid of the binomial series expansion,

$$\begin{aligned} s(x) &= s(0) + \int_0^x s'(t)\,dt \\ &= \frac{\pi}{2} - \int_0^x \left(1 - \frac{t^2}{2} - \frac{t^4}{8} - \frac{t^6}{16} + \dots\right) dt \\ &= \frac{\pi}{2} - 2x + \frac{x^3}{3} + \frac{x^5}{20} + \frac{x^7}{56} + \dots, \end{aligned} \tag{2}$$

where the coefficients of all the odd powers beginning with x^3 are positive. Applying this to the expression in (1), we see that the contributions of the constant term and the x term cancel in the sum. If each of the remaining terms just happened to yield non-positive contributions, we would be done. This is exactly what happens! Set $g = 2\cos(\pi/5) = 1.681\dots$ (the golden ratio). Then we have $d_2 = gd_1$ and $w_2 = gw_1$, so that the contribution of x^n in (2) to the expression in (1) is

$$\begin{aligned} &(d_1 + w_2)^n + (d_2 + w_1)^n - (d_2 + w_2)^n - w_1^n - d_1^n \\ &= (d_1 + gw_1)^n + (gd_1 + w_1)^n - (gd_1 - gw_1)^n - w_1^n - d_1^n \\ &= \sum_{k=0}^{n} \binom{n}{k} \left[(gd_1)^k w_1^{n-k} + d_1^k (gw_1)^{n-k} - (gd_1)^k (gw_1)^{n-k} \right] - w_1^n - d_1^n \\ &= \sum_{k=1}^{n-1} \binom{n}{k} d_1^k w_1^{n-k} \left[g^k + g^{n-k} - g^n \right]. \end{aligned}$$

We're done if

$$g^k + g^{n-k} - g^n \le 0, \quad \text{for all } n \ge 3 \text{ and } 0 < k < n. \tag{3}$$

In fact, for $n = 3$ we have *equality* for $k = 1, 2$ (since the golden ratio satisfies $g^2 = g + 1$), and we proceed by induction. Assuming that (3) holds for some $n \ge 3$ and whenever $0 < k < n$, we let $0 < k < n + 1$. Then for $0 < k < n$

$$\begin{aligned} g^k + g^{n+1-k} - g^{n+1} &= g(g^{k-1} + g^{n-k} - g^n) \\ &= g(g^k + g^{n-k} - g^n) + g^k - g^{k+1} \\ &\le g^k - g^{k+1} \le 0. \end{aligned}$$

In the case that $k = n$, we get

$$g^k + g^{n+1-k} - g^{n+1} = g^n + g - g^{n+1} = g - g^{n-1} \le 0,$$

completing the induction and establishing (3).

The pizza is done and dinner is served.

Part a also solved by Con Amore Problem Group (Denmark), Jiro Fukuta (Japan), Western Maryland College Problem Group, and the proposers.

Bill James' "Pythagorean Model" October 1994

1458. *Proposed by Dave Trautman, The Citadel, Charleston, South Carolina.*

Let $(a_i)_{i=1}^n$ and $(b_i)_{i=1}^n$ be sequences of positive integers with $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$.

a. Evaluate

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

b. Same as (a) under the additional restriction that for all i, $A \le (a_i/b_i) \le B$, where A and B are fixed positive constants.

(Note: This problem has its genesis in the work of the baseball statistician Bill James. James constructs what he calls the "Pythagorean Model," which states that if a team scores a runs in a season and gives up b runs in the same season, then its winning percentage should be approximately $a^2/(a^2 + b^2)$. The sum of the winning percentages of n teams in a league is $n/2$. This problem explores how far off this Pythagorean Model could be for the sum of the predicted winning percentages.)

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

a. Our task is to find sup E where

$$E = \left\{ \left(\sum_{i=1}^n \frac{a_i^2}{a_i^2 + b_i^2} \right) - \frac{n}{2} \; : \; a_i, b_i \in \mathbf{N}, i = 1, 2, \ldots, n \right\},$$

and $\mathbf{N}$ is the set of positive integers. The set E will not change if we allow the a_i's and b_i's to be arbitrary rationals, for then there are x_i, y_i, $D \in \mathbf{N}$ such that $a_i = x_i/D$ and $b_i = y_i/D$, for $i = 1, 2, \ldots, n$, and therefore $a_i^2/(a_i^2 + b_i^2) = x_i^2/(x_i^2 + y_i^2)$. So, redefine E as

$$E = \left\{ \left(\sum_{i=1}^n \frac{a_i^2}{a_i^2 + b_i^2} \right) - \frac{n}{2} \; : \; a_i, b_i \in \mathbf{Q}^+, i = 1, 2, \ldots, n \right\},$$

where $\mathbf{Q}^+$ is the set of positive rationals. Note that with the notation used above, $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ if, and only if, $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$.

Let $a_i, b_i \in \mathbf{Q}^+$, $i = 1, 2, \ldots, n$, and put

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

Then

$$s = \sum_{i=1}^{n} \left(\frac{a_i^2}{a_i^2 + b_i^2} - \frac{1}{2} \right) = \frac{1}{2} \sum_{i=1}^{n} \frac{a_i^2 - b_i^2}{a_i^2 + b_i^2} = \frac{1}{2} \sum_{i=1}^{n} \frac{c_i^2 - 1}{c_i^2 + 1} = \sum_{i=1}^{n} \left(\frac{1}{2} - \frac{1}{1 + c_i^2} \right)$$

where

$$c_i = \frac{a_i}{b_i}, \qquad i = 1, 2, \ldots n. \tag{1}$$

Of course, $c_i \in \mathbf{Q}^+$, $i = 1, 2, \ldots, n$. We now ask under what conditions $(c_1, c_2, \ldots, c_n) \in (\mathbf{Q}^+)^n$ is such that there are $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n \in \mathbf{Q}^+$ satisfying (1) and

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i. \tag{2}$$

As far as (1) is concerned the b_i's may be chosen freely in $\mathbf{Q}^+$ and the a_i's are then determined as

$$a_i = c_i b_i, \qquad i = 1, 2, \ldots, n. \tag{3}$$

This places some restrictions on the c_i's. But let us first note that we *may* have

$$c_i = 1 \qquad i = 1, 2, \ldots, n. \tag{4}$$

Then (3) is satisfied however the b_i's are chosen.

There are two cases that we *cannot* have. One is that

$$c_i \le 1, \qquad i = 1, 2, \ldots, n \quad \text{and not all } c_i = 1.$$

For in that case $\sum_{i=1}^n c_i b_i < \sum_{i=1}^n b_i$ for all choices of $b_1, b_2, \ldots, b_n \in \mathbf{Q}^+$.

Similarly, we cannot have

$$c_i \ge 1, \qquad i = 1, 2, \ldots, n \quad \text{and not all } c_i = 1.$$

If these cases and (for the moment) (4) are excluded, there remains the case that

$$\text{there exists } i, j \in \{1, 2, \ldots, n\} \text{ such that } c_i < 1 \text{ and } 1 < c_j. \tag{5}$$

We show that in this case there are $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n \in \mathbf{Q}^+$ such that (1) and (2) are satisfied.

We may assume that the indexing of the c_i's is such that there are $n_1, n_2 \in \mathbf{N}$ with $n_1 + n_2 = n$, $c_1, c_2, \ldots, c_{n_1} \le 1$, and $c_{n_1+1}, c_{n_1+2}, \ldots, c_{n_1+n_2} = c_n > 1$. Now put

$$\frac{c_1 + c_2 + \cdots + c_{n_1}}{n_1} = p \quad \text{and} \quad \frac{c_{n_1+1} + c_{n_1+2} + \cdots + c_n}{n_2} = q.$$

Then

$$p < 1 \quad \text{and} \quad 1 < q.$$

Also, let

$$b_i = \begin{cases} (q-1)/n_1 & \text{for } i = 1, 2, \dots, n_1 \\ (1-p)/n_2 & \text{for } i = n_1+1,\ n_1+2, \dots, n. \end{cases}$$

Then $b_i \in \mathbf{Q}^+$, $i = 1, 2, \dots, n$, and

$$\sum_{i=1}^{n} c_i b_i = \left(\frac{c_1 + c_2 + \cdots + c_{n_1}}{n_1} \right)(q-1) + \left(\frac{c_{n_1+1} + c_{n_1+2} + \cdots + c_n}{n_2} \right)(1-p)$$

$$= p(q-1) + q(1-p) = q - p = \left(\frac{q-1}{n_1} \right) n_1 + \left(\frac{1-p}{n_2} \right) n_2$$

$$= b_1 + b_2 + \cdots + b_{n_1} + b_{n_1+1} + b_{n_1+2} + \cdots + b_n = \sum_{i=1}^{n} b_i.$$

Putting $a_i = c_i b_i$, $i = 1, 2, \dots, n$, (1) and (2) are satisfied.

It is now easy to see that

$$E = \left\{ \sum_{i=1}^{n} \left(\frac{1}{2} - \frac{1}{1 + c_i^2} \right) \ :\ c_i \in \mathbf{Q}^+,\ i = 1, 2, \dots n \text{ and} \right.$$

$$\left. \left(c_i = 1 \text{ for } i = 1, 2, \dots n, \text{ or, for some } i, j, c_i < 1 \text{ and } c_j > 1 \right) \right\}.$$

Now if we choose c_1 very near 1, and $c_2, \dots, c_n$ very big, we can get values of s near as we please to $0 + (n-1)\frac{1}{2}$, and no 'admissible' choice will produce an s greater than this number. Therefore

$$\sup E = \frac{n-1}{2}.$$

b. We are given positive numbers A and B with $A \le B$ and the extra condition that

$$A \le c_i \le B,\ i = 1, 2, \dots, n.$$

We divide into three cases.

Case 1. $1 < A$ or $B < 1$. Then, neither (4) nor (5) can be satisfied, and E is the empty set. In this case, it is sometimes said that $\sup E = -\infty$.

Case 2. $A = 1$ or $B = 1$. Then $(c_1, c_2, \dots, c_n)$ can only be $(1, 1, \dots, 1)$, so $E = \{0\}$, and $\sup E = 0$.

Case 3. $A < 1 < B$. Then, by choosing c_1 just below 1, and $c_2, \dots, c_n$, just below B, we can make s as close as we please to

$$(n-1)\left(\frac{1}{2} - \frac{1}{1 + B^2} \right),$$

and we cannot make s greater than this number. Thus

$$\sup E = (n-1)\left(\frac{1}{2} - \frac{1}{1+B^2} \right) = \left(\frac{n-1}{2} \right)\left(\frac{B^2 - 1}{B^2 + 1} \right).$$

Also solved by Michael Golomb, F. C. Rembis, and the proposer.

Answers

Solutions to the Quickies on page 307.

A838. We have $|\overline{AM}| = |\overline{MP}|$ and since the diagonals of any parallelogram bisect each other, $|\overline{AO}| = |\overline{OC}|$. Therefore $\overline{CM}$ and $\overline{PO}$ are medians of triangle APC, intersecting in its centroid. Likewise, medians $\overline{DN}$ and $\overline{PO}$ intersect in the centroid of triangle BPD. Thus $\overline{PO}$ is a *fixed* median common to both triangles, so that the two centroids must coincide at point Q on $\overline{PO}$, with $|OQ| = \frac{1}{3}|\overline{PO}|$ independent of $ABCD$.

A839. For $n \geq 2$ we have

$$a_n - a_{n-1} = (r-1)a_{n-1} + sa_{n-2} = -s(a_{n-1} - a_{n-2}) = \cdots = (-s)^{n-1}(a_1 - a_0).$$

If $a_1 = a_0$ then (a_n) is a constant sequence, and $L = 1 \leq p$.

If $a_1 \neq a_0$ and $s = -1$, then $a_n = a_{n-1} + (a_1 - a_0) = a_0 + n(a_1 - a_0)$. In this case $\{a_0, a_1, \ldots, a_{p-1}\} = \mathbf{Z}_p$, $a_p = a_0$, $a_{p+1} = a_1$, and $L = p$.

Finally, if $a_1 \neq a_0$ and $s \neq -1$, then

$$a_n = a_0 + (a_1 - a_0)\sum_{i=1}^{n-1}(-s)^i = a_0 + \left(\frac{a_1 - a_0}{1+s}\right)(1 - (-s)^n).$$

By Fermat's Little Theorem, $a_{p-1} = a_0$, $a_p = a_0 + (a_1 - a_0) = a_1$, and $L \leq p - 1$.

A840. A vector proof is particularly apt here since if $\mathbf{A}_i$ is a vector from the circumcenter of $A_1A_2A_3A_4$ to the vertex A_i, the orthocenter of $A_{i+1}A_{i+2}A_{i+3}$ is given simply by $\mathbf{H}_i = \mathbf{S} - \mathbf{A}_i$, $i = 1,2,3,4$, where $\mathbf{S} = \mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3 + \mathbf{A}_4$ (note for example that $(\mathbf{H}_1 - \mathbf{A}_2)\cdot(\mathbf{A}_3 - \mathbf{A}_4) = \mathbf{A}_3^2 - \mathbf{A}_4^2 = 0$). It follows that $H_1H_2H_3H_4$ is congruent to $A_1A_2A_3A_4$, which establishes (i).

For (ii), the lines H_iA_i are given by $\mathbf{A}_i + \lambda_i(\mathbf{H}_i - \mathbf{A}_i)$, where λ_i are scalar parameters. Letting $\lambda_i = 1/2$, the four lines are concurrent at the point given by $\mathbf{S}/2$.

REVIEWS

PAUL J. CAMPBELL, *editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles and books are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of mathematics literature. Readers are invited to suggest items for review to the editors.

Macpherson, Kitta, Last theorem, first in heart of Princetonians, *Newark Star-Ledger* (17 May 1995) 1, 3. Cipra, Barry, Princeton mathematician looks back on Fermat proof, *Science* 268 (26 May 1995) 1133–1134. Faltings, Gerd, The proof of Fermat's Last Theorem by R. Taylor and A. Wiles, *Notices of the AMS* 42 (7) (July 1995) 743–746.

There's nothing quite like the championship team coming back to play the season-ending game at home. That's what it must have been like for Andrew Wiles, now recognized as having proved Fermat's Last Theorem, when he spoke to the Princeton academic community in May about his personal experiences attacking the problem since age 10. "You have to have tremendous faith that you can do it, tremendous faith that you are right and you have to never, ever give up." (Is such advice always healthy for those who do not eventually succeed?) Wiles's proof occupies the entire May 1995 issue of *Annals of Mathematics*.

Nowak, Martin A., Robert M. May, and Karl Sigmund, The arithmetics of mutual help, *Scientific American* (June 1995) 76–81. Lloyd, Alun L., Computing bouts of the Prisoner's Dilemma, 110–115.

"[H]ow can reciprocal altruism work in the absence of [law and enforcement]?" Nowak et al. expand on previous computer simulations of repeated Prisoner's Dilemma by determining the response for the next trial based on the outcome of the previous one in a stochastic way. Dominant strategies include the well-known Tit-for-Tat and also Pavlov ("win-stay, lose-shift"). When players are constrained to interact only with neighbors in a planar grid, the long-term average proportion of cooperators is predictable; Lloyd gives a Basic program that lets readers explore a version of this variation.

Macpherson, Kitta, The fruits of knowledge: Bell labs staff marvel at clone of Newton's "gravity tree," *Newark Star-Ledger* (29 June 1995) 52.

How did the legend of Newton and the apple tree start? Newton never wrote about apple trees. A tree at Bell Labs, descended from the genetic line ("the Pride of Kent") of the tree in Newton's garden, has borne its first two small apples. It's 3 ft high; "I don't think you could sit under it and think. But it will inspire people," says Arno Penzias, Nobel laureate and vice-president of Bell Labs. And the legend? A contemporary of Newton's, William Stukeley, while walking with the elderly Newton through an apple orchard, was told by Newton that he was "in the ... same situation ... when ... the notion of gravitation came into his mind. It was occasion'd by the fall of an apple, as he sat in a contemplative mood."

Gallian, Joseph A., The Duluth undergraduate research program in discrete mathematics, *Council on Undergraduate Research Quarterly* (March 1995) 142–144.

Author Gallian has run a highly successful summer program of undergraduate research for almost twenty years. Through 1994, the 56 participants (from 28 institutions) had published 39 papers written in the program, with 13 more accepted and 21 more submitted; 43 of the 48 who had completed undergraduate work went to graduate school, 36 won graduate fellowships, and 15 now have Ph.D.'s. "[C]arefully selected undergraduates ... have the ability, the desire and the time [to do publishable research in mathematics]. They do not know how to start or how to finish. To begin, they need problems and guidance. To end, they need assistance with manuscript preparation and the publication process. In between, they need encouragement and reassurance." (Thanks to Anant Godbole, Michigan Technological University.)

Penrose, Roger, Various puzzles based on Penrose tiles. Kites and Darts (plastic pieces in five colors, in packages with equal numbers of each or in the ratio 5 Kites to 3 Darts); Diamonds and Rhombi (same); Collidescape (uses two different isosceles triangle shapes); kits start at $36. Distributed in the U.S. by Kadon Enterprises (1227 Lorene Dr., Suite 16, Pasadena, MD 21122). Perplexing Poultry (skinny chickens, fat chickens, and romping dogs, 195 pieces), $99.95 (color), $79.95 (black and white); Game Birds (a smaller version with 45 pieces), $39.95; circular jigsaw puzzles (Perplexing Poultry, Perplexing Pisces, Cat Amongst the Pigeons, Pentaplexity, 500 pieces each), $17.95 each. Distributed in the U.S. by Kadon Enterprises and also by World of Escher (14542 Brook Hollow Blvd #250, San Antonio, TX 78232–3810; (800) 237-2232). Detailed review: Schattschneider, Doris, Penrose puzzles, *SIAM News* 28 (6) (July 1995) 8, 14.

Despite the appeal and popularity (among mathematicians) of Penrose's aperiodic tiling of the plane with two tiles, sets of the pieces have not been widely available. At last, they are, in several deluxe editions with puzzle frameworks.

Kolata, Gina, A vat of DNA may become the fast computer of the future, *New York Times* (11 April 1995) (National Edition) B5, B8. Pearson, Peter, Biochemical techniques take on combinatorial problems, *Dr. Dobb's Journal* (August 1995) 127–131. Linial, Michael, et al., On the potential of molecular computing, *Science* 268 (28 April 1995) 481–484. Lipton, Richard J., DNA solution of hard computational problems, *Science* 268 (28 April 1995) 542–545. Gonick, Larry, The solution [comic strip explanation of solution of the traveling sales representative problem via DNA computation], *Discover* 16 (April 1995) 36–37.

Late in 1994, Leonard Adleman (USC) reported how he found hamiltonian paths in a directed graph by modeling the problem with molecules, allowing the molecules to interact chemically, and observing the production of a particular molecule (*Science* 266 (11 November 1994) 1021–1024). His work has excited many computer scientists, who foresee the ability to use the massive parallelism of molecules to solve problems that cannot be done efficiently on current digital machines. For example, one researcher has estimated that the 56-bit key to a message encrypted using the Data Encryption Standard could be found in four months of lab work. Letters to *Science* suggest that Adleman's article is overly optimistic and that extending his method from a graph with 7 vertices to one with 70 would require 10^{25} kg of molecules. Adleman's response suggests that "it is too early for either great optimism or pessimism" and notes that " Devices become 'computers' when we learn how to interpret their behavior appropriately."

Morton, Carol Cruzan, Double bubbles designed by nature are best containers, mathematicians say, UC Davis News Release (3 August 1995). Text available from kholmay@nas.edu or ccmorton@ucdavis.edu.

Joel Hass (UC–Davis) and Roger Schlafly (Real Software, Davis, CA) report solving a 2,000-year-old problem. They proved that the "double bubble," formed by forcing together two bubbles into a compound bubble with a wall separating two spherical pieces, is the most efficient way (minimizing surface area) to enclose two equal volumes. They presented their results at the MAA's 1995 Mathfest in Burlington, VT. There are potential practical applications, as well as possible applications of the mathematical techniques to other problems in global optimization theory. The problem had languished until about five years ago, when Frank Morgan (Williams College) brought it up with a group of undergraduate researchers. After other investigators narrowed the candidates for most efficient to two families of bubbles, based on the double bubble and a "torus" bubble, Hass and Schlafly reduced the problem to a 20-minute computer search among feasible minimizing surfaces.

McBeath, M.K., D.M. Shaffer, and M.K. Kaiser, How baseball outfielders determine where to run to catch fly balls, *Science* 268 (28 April 1995) 569ff. Hilts, Philip J., A theory on how outfielders make catches, *New York Times* (28 April 1995) (National Edition) B16. Stewart, Ian, Owzat! ... how outfielders make a perfect catch, *New Scientist* 146 (6 May 1995) 18.

Experiments show that baseball outfielders catch a fly ball by using a strategy different from the hitherto prevailing theory. They do not run in a straight line at constant speed to where they think the ball will be—that would make the path of the ball appear curved to them—but by taking a slightly curved path and varying speed so that the ball appears to follow a linear path. This new theory explains why outfielders collide with walls and other players: They don't really know where the ball is going to come down, as they are adapting all through its trajectory. (For baseball aficionados, author McBeath has an earlier article of interest, "The rising fastball: Baseball's impossible pitch," *Perception* 19 (1990) 545ff.)

Parker, Marla (ed.), *She Does Math! Real-Life Problems from Women on the Job*, MAA, 1995; xvi + 253 pp, $24 ($18.50 to MAA members) (P).

This is a unique and very welcome collection of autobiographical sketches, real-life problems from the work world, accompanying exercises (154) with solutions (which occupy 30% of the book), and photos of 38 women who use mathematics in a great variety of lines of work. This book belongs in every high-school library, after passing through the hands of the guidance counselors.

Calinger, Ronald (ed.), *Classics of Mathematics*, 2nd ed., Prentice-Hall, 1995; xxi + 793 pp, (P).

This splendid collection brings together short excerpts from more than 130 mathematical classics by 62 authors, together with biographies of the authors and substantial introductions to some of the nine historical periods (two of these run to 35 pp each). The fact that this is a second edition from a different publisher (1st ed: Moore Publ. Co., 1982) is barely evident from the publication information, and nothing indicates what differences there are from the first edition. In fact, there are four minor additions: two pages on "Poincaré and topology," by P.S. Alexandrov; Euclid's *Elements* Book IX, Propositions 25–30 (even plus even gives even, etc.); from Book X of the *Elements*, Propositions 1–3 (commensurable and incommensurable magnitudes); and a short essay on the mathematical notation of the ancient Maya, by Michael Closs.

NEWS AND LETTERS

LETTERS TO THE EDITOR

We have had lots of letters in response to the article "The Product of Chord Lengths of a Circle," by Mazzoleni and Shen (February 1995). We've already published one in June, suggesting other proofs that the product of the chord lengths from the point z=1 to the other roots of unity is equal to n. Here are some other ideas and comments.

***Zalman Usiskin**, University of Chicago, writes that the result is also proved using trigonometry and properties of roots of polynomials in his article, "Products of Sines" [1].*

***Kurt Eisemann**, San Diego State University, points out that the result was published in 1954 [2], and that he had written an extension of the theorem in 1972 [3].*

*We conclude with **Tom P. Apostol**, California Institute of Technology, who remarked that the equation*

$$\prod_{k=1}^{n-1}\left(1 - e^{2\pi ik/n}\right) = n,$$

is equivalent to the well-known product formula (see [4]). --*Ed.*

REFERENCES

1. Z. Usiskin, Products of sines, *Two-Year College Math. J.*, November (1979).
2. M. Sichardt, Ein satz vom kries, *ZAMM* 34 (1954), 429.
3. K. Eisemann, Extension of a theorem about the circle, *ZAMM* 52 (1972), 496-7.
4. T.M. Apostol, *Mathematical Analysis*, 2nd edition, Addison-Wesley Publishing Co., Reading, MA, 1974.

Dear Editor:

"The Box Problem: To Switch or Not To Switch," by Brams and Kilgour (February 1995), makes a mountain out of a molehill, albeit an instructive mountain. The resolution of the problem, which is more a joke than a paradox, has nothing to do with conditional probabilities of peculiar distributions. In fact, *the* weakness of the expected-value argument is that it is nonsense. Having picked a box containing $\$n$, calculating that the expected value of the other box is $(5/4)n$ involves evaluating an expression in n that simultaneously assumes that n is equal to both b and $2b$. It is not surprising that this leads to a peculiar result. Substituting the correct amounts for the occurrences of n in the averaging expression gives $(3/2)b$ as the expected value of the amount in the other box, exactly equal, of course, to the expected value of the box you've already picked.

Bill Mixon
14045 North Green Hills Loop
Austin, TX 78737

Bill Mixon incorrectly interprets the box problem to be one in which the amount in the box you pick is "equal to both b and 2b," so on average the amount you find is 3b/2, and this is the same for both boxes. But in fact the problem states that you begin by choosing one box and find the amount, x, in it, with x=b and x=2b equally probable. This equiprobability assumption is not the same as assuming that these two events are simultaneously true (a physical impossibility)--or, equivalently, that you choose one of the two boxes, containing either b or 2b at random, and don't open it--which would, indeed, give you no reason to switch to the second box. Rather, you open the first box and learn its contents, although, as we showed, the 5x/4 expected-value calculation about the contents of the second box is erroneous. Instead, your expectation about the second box's contents depends on your prior distribution. For some prior distributions, you always do better switching--whatever amount you find in the first box--which, if *not a paradox, is not a joke either.*

--Steven J. Brams and D. Marc Kilgour

Dear Editor:

The "Math Bite: Normality of the Commutator Subgroup," by Myers (February 1995), can be easily strengthened--twice.

Since every element x of G defines, via $a \to xax^{-1}$, an automorphism φ of G [onto itself], the argument in [1] also proves that the commutator subgroup is invariated by all automorphisms defined in this manner ("interior automorphisms"). Indeed, with the operator in postfixed position,

$$(aba^{-1}b^{-1})\varphi = a\varphi \cdot b\varphi \cdot a^{-1}\varphi \cdot b^{-1}\varphi =$$
$$= a\varphi \cdot b\varphi \cdot (a\varphi)^{-1} \cdot (b\varphi)^{-1} \in C. \quad (*)$$

It is obvious, in this formulation, that the result in no way depends on the "interiorness" of the automorphism (i.e. on the existence of a group element x such that for all a, $a\varphi = xax^{-1}$). Therefore, the commutator subgroup is invariated not only by interior automorphisms but also by all other automorphisms of G. In other words, the commutator of a group G, is not only a normal subgroup but also a characteristic subgroup of G.

Furthermore, nothing in (*) depends on the fact that φ, viewed as an endomorphism of G [into itself], is surjective. Any endomorphism φ of G, acting on the commutator of the elements a, b, yields a group element that is in the commutator group (being the commutator of the elements $a\varphi$ and $b\varphi$). Therefore, the commutator subgroup is invariated not only by all the automorphisms of G, but indeed by all the endomorphisms of G. In other words, the commutator of a group G, besides being a characteristic subgroup of G, is also a fully invariant subgroup of G. (See [2], Chap. IV, sec. 14.)

Kuroš also shows that the center of the group (defined as the subset [subgroup] consisting of all elements, each of which invariated by all interior automorphisms) is an example of a characteristic subgroup that is not fully invariant.

Robert Mandl
Box 397199 (MIT Br.)
Cambridge, MA 02139-0902

REFERENCES

1. Lawrence Myers, Math Bite: Normality of the commutator subgroup, this MAGAZINE, 68(1995), 49.
2. Aleksandr Gennedievič Kuroš, *Group Theory*, Moscow, 1944, 1953.

Dear Editor:

Here is an appetizer, a snack, and a seasoning to supplement "Math Bite: $\sum a_k b_k < (\sum a_k^2)^{1/2}(\sum b_k^2)^{1/2}$," by Peter Szüsz (April 1995).

$$\frac{\sum a_k b_k}{\left(\sum a_k^2\right)^{\frac{1}{2}}\left(\sum b_k^2\right)^{\frac{1}{2}}} = 1 - \frac{1}{2}\sum\left(\frac{a_k}{\left(\sum a_k^2\right)^{\frac{1}{2}}} - \frac{b_k}{\left(\sum b_k^2\right)^{\frac{1}{2}}}\right)^2$$

The appetizer: The following equality is similar to Szüsz's

$$\frac{\sum a_k b_k}{\left(\sum a_k^2\right)^{\frac{1}{2}}\left(\sum b_k^2\right)^{\frac{1}{2}}} = -1 + \frac{1}{2}\sum\left(\frac{a_k}{\left(\sum a_k^2\right)^{\frac{1}{2}}} + \frac{b_k}{\left(\sum b_k^2\right)^{\frac{1}{2}}}\right)^2$$

The snack: These equations imply

$$-1 \le \frac{\sum a_k b_k}{\left(\sum a_k^2\right)^{\frac{1}{2}}\left(\sum b_k^2\right)^{\frac{1}{2}}} \le 1$$

and

$$0 \le \sum\left(\frac{a_k}{\left(\sum a_k^2\right)^{\frac{1}{2}}} \pm \frac{b_k}{\left(\sum b_k^2\right)^{\frac{1}{2}}}\right)^2 \le 4$$

The seasoning: These equalities are applied by statisticians in interpreting the correlation coefficient, in which case $a_k = x_k - x$ and $b_k = y_k - y$ where x and y are the arithmetic means for bivariate observations (x_k, y_k). For example, see Gary G. Koch, A basic demonstration of the [-1,1] range for the correlation coefficient, *Amer. Statist.* 39 (1985) 201-202, and Joseph Lee Rogers and W. Alan Nicewander, Thirteen ways to look at the correlation coefficient, *Amer. Statist.* 42 (1988) 59-66.

David L. Farnsworth
Rochester Institute of Technology
Rochester, NY 14623

Dear Editor:

"Oscillating Sawtooth Functions" (June 1995) exhibits some derivatives that are continuous at 0. The following related reference might interest some readers: In the October 1963 *Monthly*, pp. 867-868, there is an example (attributed to William M. Myers) of a function g and a subset H of [0,1] such that H has measure ½, g is differentiable on [0,1], and g' is discontinuous at all points of H. The construction of H is somewhat analogous to that of the Cantor set.

David M. Bloom
Brooklyn College of CUNY
Brooklyn, NY 11210

Yongzhi (Peter) Yang, University of St. Thomas, St. Paul, MN, writes that there were several errors in "Perturbation of a Tridiagonal Stability Matrix," by M. Elizabeth Mayfield (April 1994). In particular, the theorem is not stated correctly at the beginning of the paper, although the theorem is correctly stated in the last sentence. In fact, equation (3) on page 127 should read

$$\mu_j = \lambda_j + (x_n^j)^2 + \alpha_j.$$

There is actually a different α_j for each eigenvalue, and the sum of the α_j's equals zero. Yang supplied a proof of the corrected theorem, pointing out an error in Mayfield's discussion.

In reponse, Mayfield notes that Che-Kao Fong, Carleton University, Ottawa, Canada, has also provided a proof. Fong suggests that we assume $Im\ \epsilon > 0$. Let μ be an eigenvalue of T' and y be a corresponding eigenvector: $T'y = \mu y$. Assume $\|y\| = 1$. Then $\mu = (T'y,y) = (Ty,y) + \epsilon\ (Uy,y) = (Ty,y) + \epsilon\ \|Uy\|$. Since T is a real, symmetric matrix, $(Ty,y) \in$ N and hence $Im\ \mu = (Im\ \epsilon)\ \|Uy\|^2$. Thus, $Im\ \mu > 0$ unless $Uy = 0$. But from $Ty + \epsilon Uy = T'y = \mu y$, we see that $Uy = 0$ implies that y is an eigenvector of T with its last component equal to zero. This is impossible as is stated in the article.

Dear Editor:

In "Using Complex Solutions to Aid in Graphing" (April 1993), C. Bandy used the complex roots $\alpha \pm \beta i$ to graph the quadratic $y = ax^2 + bx + c$. Assume that $a > 0$.

Measuring α is not difficult. Since

$$y(x) = x^2 + bx + c = a(x-\alpha-\beta i)(x-\alpha+\beta i) = a((x-\alpha)^2 + \beta^2) \geq a\beta^2,$$

y has its minimum at $x = \alpha$ (for $a > 0$). So, α is the x-coordinate of the minimum point or vertex.

Other intriguing questions are: Can the imaginary part be measured? And if it can be, how can it be measured? One way to see that the answer to the first question is "yes" is the following simple construction displayed in the FIGURE.

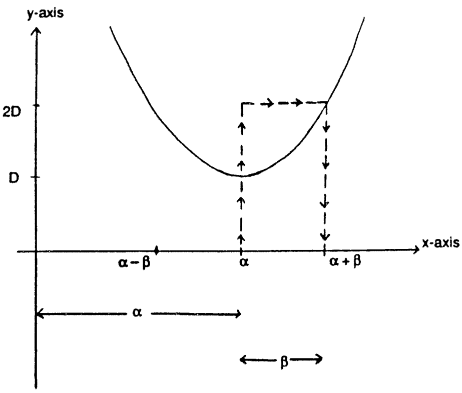


FIGURE 1
The construction for the imaginary part β

From the point $(\alpha,0)$, which lies directly below the minimum, draw a vertical line to $(\alpha,2D)$ where D is the y-coordinate of the minimum. Then draw a line parallel to the x-axis at $y = 2D$ to meet the curve, and drop a perpendicular line from the meeting point to the x-axis. The x-coordinate there is $\alpha + \beta$. So, β is the distance from $x = \alpha$ to $x = \alpha + \beta$.

David L. Farnsworth
Rochester Institute of Technology
Rochester, NY 14623

REFERENCE

1. C. Bandy, Using cmplex solutions to aid in graphing, this MAGAZINE, 66(1993), 125.

THE MATHEMATICAL ASSOCIATION OF AMERICA

COMPLEX NUMBERS AND GEOMETRY

Liang-shin Hahn

This book is an excellent source for enrichment topics for advanced students.
— The Mathematics Teacher

I recommend this book for anyone with an interest in solving problems involving geometric concepts. ***—AAAS Science Books and Films***

The purpose of this book is to demonstrate that complex numbers and geometry can be blended together beautifully, resulting in easy proofs and natural generalizations of many theorems in plane geometry—such as the Napoleon theorem, the Ptolemy-Euler theorem, the Simson theorem, and the Morley theorem.

Beginning with a construction of complex numbers, you are taken on a guided tour that includes something for everyone, even those with advanced degrees in mathematics. Yet, the entire book is accessible to a talented high-school student.

The book is self-contained—no backgound in complex numbers is assumed—and can be covered at a leisurely pace in a one-semester course. Over 100 exercises are included. The book would be suitable as a text for a geometry course, or for a problem solving seminar, or as enrichment for the student who wants to know more.

200 pp., Paperbound, 1993

ISBN 0-88385-510-0

Catalog Number: CNGE-JN

LIST: $28.00 MAA MEMBERS: $21.50

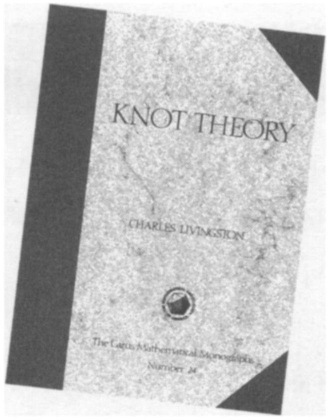
Send to: The MAA, 1529 18th, NW, Washington, DC 20036, 1-800-331-1622; FAX: (202) 265-2384

Name	Quantity	Catalog Number	Price
Address			
City State Zip Code			
❑ Check ❑ VISA ❑ MasterCard			
Card Number	Postage & Handling		
Signature Exp. Date	TOTAL		

Postage and Handling In the USA, please add $2.95 for the first book, and $.50 for each additional book. Canadian orders, please add $4.50 for the first book, and $1.50 for each additional book.

Knot Theory

Charles Livingston



I learned more about knots after an hour with the book than I thought I could, and I am glad that it is here on my desk so that I may spend more time with it and, I hope, learn more. —*Paul Halmos*

Knot Theory, a lively exposition of the mathematics of knotting, will appeal to a diverse audience from the undergraduate seeking experience outside the traditional range of studies to mathematicians wanting a leisurely introduction to the subject. Graduate students beginning a program of advanced study will find a worthwhile overview, and the reader will need no training beyond linear algebra to understand the mathematics presented.

Over the last century, knot theory has progressed from a study based largely on intuition and conjecture into one of the most active areas of mathematical investigation. **Knot Theory** illustrates the foundations of knotting as well as the remarkable breadth of techniques it employs—combinatorial, algebraic, and geometric.

The interplay between topology and algebra, known as algebraic topology, arises early in the book, when tools from linear algebra and from basic group theory are introduced to study the properties of knots, including the unknotting number, the braid index, and the bridge number. Livingston guides you through a general survey of the topic showing how to use the techniques of linear algebra to address some sophisticated problems, including one of mathematics' most beautiful topics, symmetry. The book closes with a discussion of high-dimensional knot theory and a presentation of some of the recent advances in the subject—the Conway, Jones and Kauffman polynomials. A supplementary section presents the fundamental group, which is a centerpiece of algebraic topology.

An extensive collection of exercises is included. Some problems focus on details of the subject matter; others introduce new examples and topics illustrating both the wide range of knot theory and the present borders of our understanding of knotting. All are designed to offer the reader the experience and pleasure of working in this fascinating area.

264 pp., Hardbound, 1993

ISBN 0-88385-027-3

List: $34.00 MAA Member: $27.50

Catalog Number CAM-24

ORDER FROM:

The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036
1-(800) 331-1622 (202)-387-5200

Membership Code
_ _ _ _ _ _

Name____________________

Address____________________

City____________________

State____ Zip Code__________

Qty. Catalog Number Price

Total $________

Payment ❑ Check ❑ VISA ❑ MASTERCARD

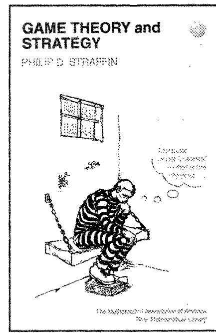
Credit Card No.____________________

Signature____________________

Exp. Date____________________

Game Theory and Strategy

Philip D. Straffin, Jr.



This valuable addition to the New Mathematical Library series pays careful attention to applications of game theory in a wide variety of disciplines. The applications are treated in considerable depth. The book assumes only high school algebra, yet gently builds to mathematical thinking of some sophistication. **Game Theory and Strategy** might serve as an introduction to both axiomatic mathematical thinking and the fundamental process of mathematical modelling. It gives insight into both the nature of pure mathematics, and the way in which mathematics can be applied to real problems.

Since its creation by John von Neumann and Oskar Morgenstern in 1944, game theory has contributed new insights to business, politics, economics, social psychology, philosophy, and evolutionary biology. In this book, the fundamental ideas of game theory share the stage with applications of the theory. How might strategic business decisions depend on information about a rival company, and how much would such information be worth? When is it advantageous to vote for a candidate who is not your favorite? What are the optimal strategies for teams in the football draft, and what paradoxes can result from following those strategies? What is a fair way to share the costs of a development project? What can we learn about the problem of "free will" by imagining playing a game with an omnipotent Being? How might natural selection lead to altruistic behavior in animal species? Game theory gives insight into all of these questions.

The book includes many exercises, with answers, which allow the reader to try out calculations, and explore alternative formulations of game-theoretic ideas.

200 pp., 1993, Paperbound

ISBN 0-88385-637-9

List: $29.50 MAA Member: $24.00

Catalog Number NML-36

ORDER FROM:

The Mathematical Association of America
1529 Eighteenth Street, NW
Washington, DC 20036
(202) 387-5200 Fax (800)331-1622

Membership Code ______

Name ____________________

Address ____________________

City ____________________

State ___ Zip Code ____________

Qty. Catalog Number Price

____

____

Total $ ________

Payment ❑ Check ❑ VISA ❑ MASTERCARD

Credit Card No. ____________________

Signature ____________________

Exp. Date ____________________

CONTENTS

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