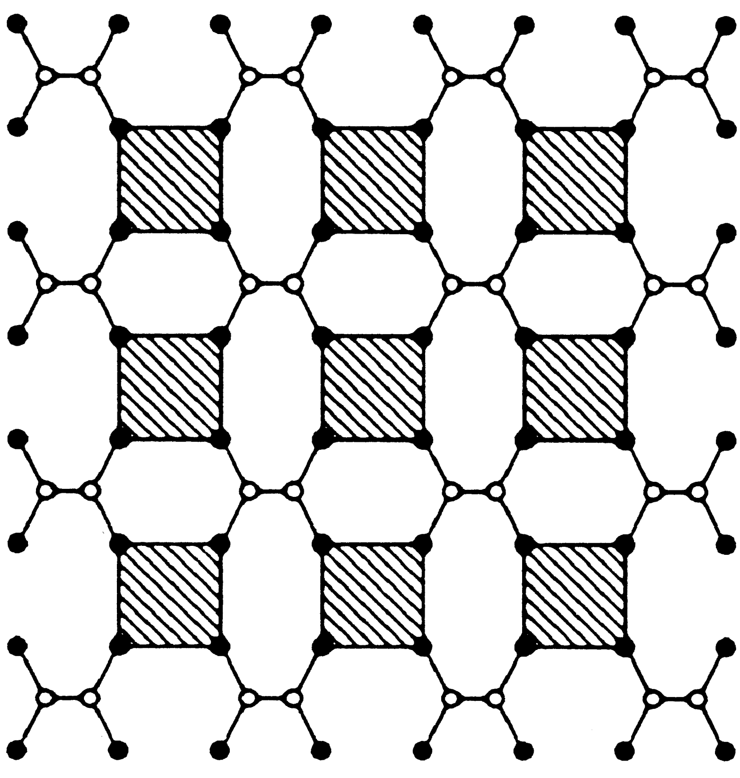
Vol. 62 No. 2 April 1989

MATHEMATICS MAGAZINE



- Steiner Trees on a Checkerboard
- Gunfight at the OK Corral
- How Small Is a Unit Ball?

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. 54, pp. 44–45, 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: G. L. Alexanderson, Editor, Mathematics Magazine, Santa Clara University, Santa Clara, CA 95053. 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 one copy and keep one copy. Illustrations should be carefully prepared on separate sheets in black ink, the original without lettering and two copies with lettering added.

AUTHORS

Fan Chung received her Ph.D. in mathematics from the University of Pennsylvania in 1974. Since then, she has done research in graph theory and combinatorics at Bell Laboratories and Bellcore. She is currently the Division Manager of the Mathematics, Information Sciences and Operations Research Division at Bellcore. She serves on the council of the AMS and on the editorial boards of several journals, and is the editor in chief of the *Journal of Graph Theory*. Among her research in extremal and algorithmic problems, one of her favorites is the Steiner tree problem. She has worked on the problem from time to time with both of her coauthors, one of whom is not her husband.

Martin Gardner wrote the Mathematical Games column in *Scientific American* for 25 years. Thirteen book collections of these columns have been published. He is the author of many other books on recreational mathematics, science, philosophy, and literature. His interest in Steiner trees was sparked by setting himself the task—apparently not attempted before—of finding a minimal network joining the 81 points of a chessboard.

Ron Graham has spent the past 25 years mainly at AT&T Bell Laboratories with occasional semester visits to Caltech, Princeton, Stanford, and UCLA. During this time he has been involved in a variety of mathematical activities which include teaching, research, juggling, editing, curriculum development, administration, and promoting the public understanding of mathematics. He has worked on various aspects of the subject of his article off and on over the past two decades with his coauthors, one of whom is his wife.

Vol. 62 No. 2 April 1989

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 $11 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $22. Student, unemployed and emeritus members receive a 50% discount; new members receive a 30% dues discount for the first two years of membership.) The nonmember/library subscription price is $28 per year. Bulk subscriptions (5 or more copies) are available to colleges and universities for classroom distribution to undergraduate students at a 41% discount ($6.50 per copy—minimum order $32.50).
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. 20077-9564.
PRINTED IN THE UNITED STATES OF AMERICA

ARTICLES

Steiner Trees on a Checkerboard

FAN CHUNG
Bellcore
Morristown, NJ 07960

MARTIN GARDNER
110 Glenbrook Drive
Hendersonville, NC 28739

RON GRAHAM
AT&T Bell Laboratories
Murray Hill, NJ 07974

1. Introduction

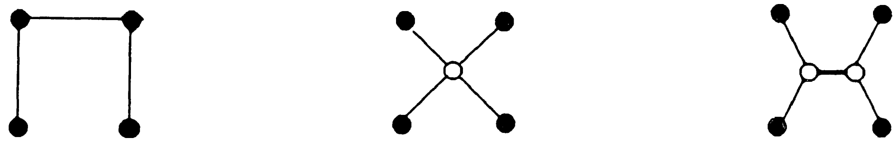
Suppose a finite set of n points are randomly scattered about in the plane. How can they be joined by a network of straight lines with the shortest possible total length? The solution to this problem has practical applications in the construction of a variety of network systems, such as roads, power lines, pipelines, and electrical circuits.

It is easy to see that the shortest network must be a tree, that is, a connected network containing no cycle. (A cycle is a closed path that allows one to travel along a connected path from a given point to itself without retracing any line.) If no *new* points can be added to the original set of points, the shortest network connecting them is called a *minimum spanning tree.*

A minimum spanning tree is not necessarily the shortest network spanning the original set of points. In most cases a shorter network can be found if one is allowed to add more points. For example, suppose you want to join three points which form the vertices of an equilateral triangle. Two sides of the triangle make up a minimum spanning tree. This spanning tree can be shortened by more than 13 percent by adding an extra point at the center and then making connections only between the center point and each corner (see FIGURE 1). Each angle at the center is 120°.

FIGURE 1

A less obvious example is the minimum network spanning the four vertices of a square. One might suppose one extra point in the center would give the minimum network, but it does not. The shortest network requires, in fact, *two* extra points (see FIGURE 2). Again all the angles around the extra points in the network are 120°. The



length 3 (a) — length $2\sqrt{2}$ (b) — length $1+\sqrt{3}$ (c)

FIGURE 2

network with one extra point in the center has length $2\sqrt{2}$, or about 2.828. The network with two extra points reduces the total length to $1+\sqrt{3}$, or about 2.732.

One of the first mathematicians to investigate such networks was Jacob Steiner, an eminent Swiss geometer who died in 1863. The extraneous points that minimize the length of the network are now called Steiner points. It has been proved that all Steiner points are junctions of three lines forming three 120° angles. The shortest network, allowing Steiner points, is called a *minimum Steiner tree*. Minimum Steiner trees are almost always shorter than minimum spanning trees, but the reduction in length usually depends on the shape of the original spanning tree. It has been conjectured [9] that for any given set of points in the plane, the length of the minimum Steiner tree cannot be less than a factor of $\sqrt{3}/2$, or about .866, times the length of the minimum spanning tree; the result has been proved, however, only for three, four, and five points [10], [12].

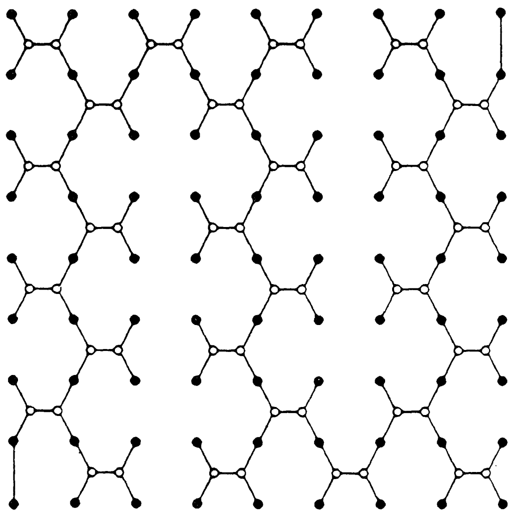
Many properties of minimum Steiner trees can be found in the excellent (but somewhat out-of-date) survey paper of E. N. Gilbert and H. O. Pollak [9]. The best current lower bound for the ratio of the minimum Steiner tree to the minimum spanning tree is .8241... (see [4]).

There are many ways to construct a minimum spanning tree. One of the simplest methods is known as a greedy algorithm, because at each step it bites off the most desirable piece. First find two points that are as close together as any other two and join them. If more than one pair of points are equally close, choose any such pair. Repeat this procedure with the remaining points in such a way that joining a pair never completes a circuit. The final result is a spanning tree of minimum length. This algorithm is due to Kruskal in a 1956 paper [11].

Given the simplicity of Kruskal's greedy algorithm for the construction of minimum spanning trees, one might suppose there would be correspondingly simple algorithms for finding minimum Steiner trees. Unfortunately, however, this is almost certainly not the case. This task belongs to a special class of "hard" problems known in computer science as NP-complete problems. When the number of points in a network is small, say 10 to 20, there are known algorithms [5], [13] for finding minimum Steiner trees in a reasonably short time. As the number of points grows, however, the computing time needed increases at a rapidly accelerating pace. Even for a relatively small number of points the best algorithms currently available could take thousands or even millions of years to terminate. Most mathematicians believe no efficient algorithms exist for constructing minimum Steiner trees on arbitrary sets of points in the plane [7], [8].

Imagine, however, that the points are arranged in a regular lattice of unit squares, like the points at the corners of the cells of a checkerboard. Is there a "good" algorithm for finding a minimum Steiner tree spanning the points of such regular patterns? In particular, what is the length of the minimum Steiner tree that joins the 81 points at the corners of a standard checkerboard? Is the tree in FIGURE 3 the solution?

Many problems involving paths through points in the plane, which are hard when



length 73.033

FIGURE 3

the points are arbitrary, become trivial when the points form regular lattices. One might expect that the task of spanning points in such arrays by minimum Steiner trees would be equally trivial. On the contrary, this problem seems to be surprisingly elusive. Up to now, only minimum Steiner trees for 2 by n rectangular arrays of points have been constructed [3]. Aside from this special case, very little seems to be known about how to find minimum Steiner trees for rectangular arrays when the number of points on each side is greater than 2.

In this paper, we will summarize various problems, conjectures, and some partial results on the minimum Steiner trees for rectangular arrays. Section 2 contains the shortest known trees for square lattices of small size. These trees consist of copies of the symmetrical tree on four points (see FIGURE 2(c)), which we call X from now on, together with a small number of "exceptional" pieces. For example, the conjectured solution for 64 points is a union of 21 X's (see FIGURE 5). In Section 3, we give a proof that a rectangular array can be spanned by a Steiner tree made up entirely of X's if and only if the array is a square and the order of the square is a power of 2. Further questions are proposed in Section 4.

2. Short Steiner trees on square lattices

In this section, we will first show the shortest Steiner trees we currently know for square lattices of size up to 14 by 14. We will then discuss a scheme for constructing Steiner trees for large square lattices from the small ones. Among all the constructions, only the patterns for the 2×2, 3×3 and 4×4 squares have been proved to be minimum Steiner trees (unpublished results of E. J. Cockayne). The constructions for square lattices of orders 2 to 9 were contained in the June 1986 issue of *Scientific American* [6]. The trees for square lattices of sizes 10 by 10 and 22 by 22 in the same article were soon improved by many readers. The current best tree for the 10 by 10 square lattice is due to one of the authors (RLG) and the best tree for the 22 by 22 square lattice is due to Eric Carlson [1]. His construction has the same total length as our general construction. Overall, the constructions fall naturally into six classes, depending on what n is modulo 6, with the rare (and remarkable) exceptions which occur when n is a power of 2. It seems that the Steiner trees for square lattices are always formed by attaching small minimum Steiner trees such as an edge E (for the 1 by 2 array), X (for the 2 by 2 array), a triangle T (formed from three vertices of a square) and L (for the 2 by 5 array). We will use the following notation:

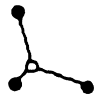
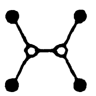
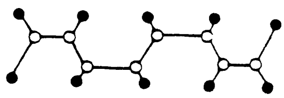
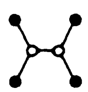
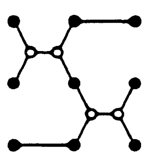
Tree	Symbol	Length
	E	$e=1$
	T	$t=\frac{1+\sqrt{3}}{\sqrt{2}}=1.93185\ldots$
	X	$x=1+\sqrt{3}=2.73205\ldots$
	L	$l=\sqrt{35+20\sqrt{3}}=8.34512\ldots$

FIGURE 4

The tree L is a minimum Steiner tree on the regular 2 by 5 array [3]. The minimum Steiner tree for the 2 by n array with n even is just made up of X's joined together by edges. On the other hand the minimum Steiner tree for the 2 by n array with n odd has length $\frac{1}{2}((n(2+\sqrt{3})-2)^2+1)^{1/2}$. The constructions for $n \times n$, $n \leqslant 14$, are illustrated in FIGURE 5.

n	Conjectured Minimum Steiner Tree	Length
2	 *	$x=2.73205\ldots$
3	 *	$2x+2=7.46410\ldots$

*known to be optimal

FIGURE 5

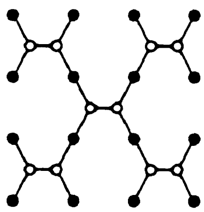
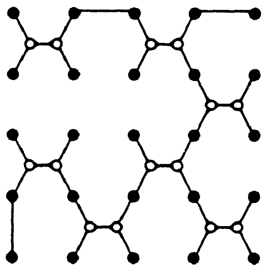
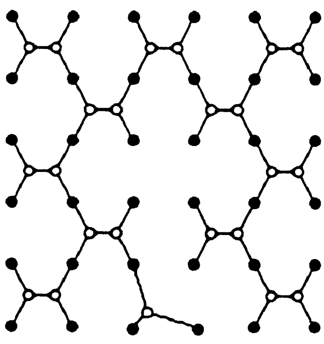
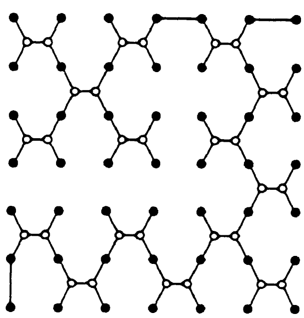
n	Conjectured Minimum Steiner Tree	Length
4	 *	$5x = 13.66025\ldots$
5		$7x + 3 = 22.12436\ldots$
6		$11x + t = 31.98441\ldots$
7		$15x + 3 = 43.98076\ldots$

FIGURE 5 (con't)

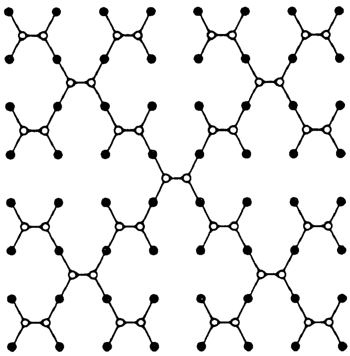
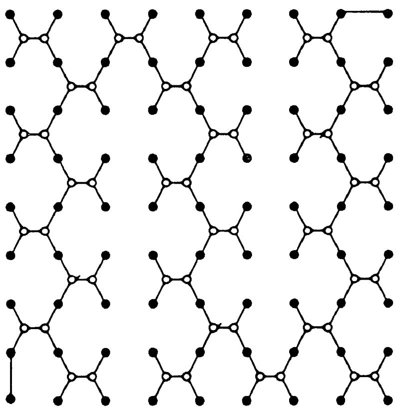
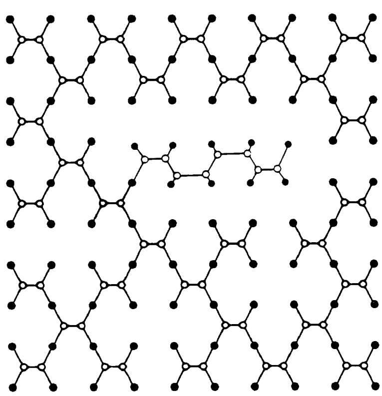
n	Conjectured Minimum Steiner Tree	Length
8		$21x = 57.373067\ldots$
9		$26x + 2 = 73.03332\ldots$
10		$30x + l = 90.30664\ldots$

FIGURE 5 (con't)

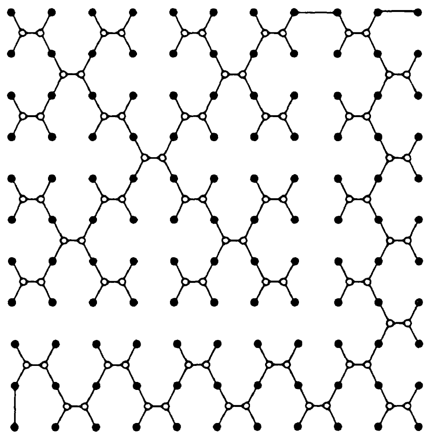
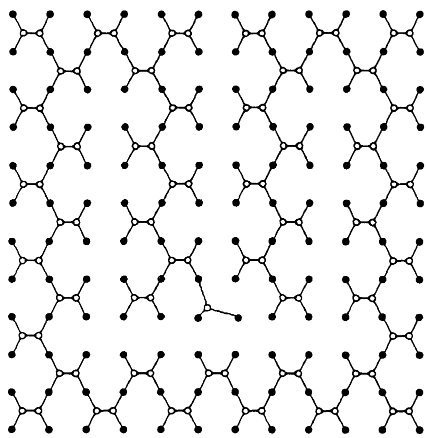
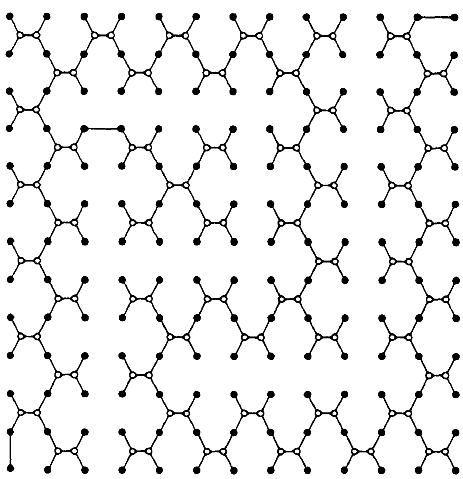
n	Conjectured Minimum Steiner Tree	Length
11		$39x + 3 = 109.54998$
12		$47x + t = 130.33824\ldots$
13		$55x + 3 = 153.26279$

FIGURE 5 (con't)

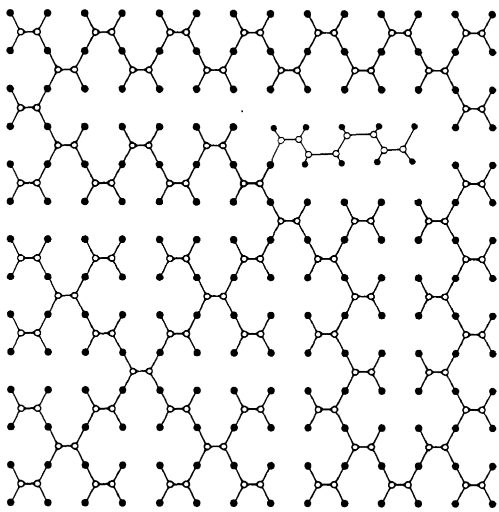
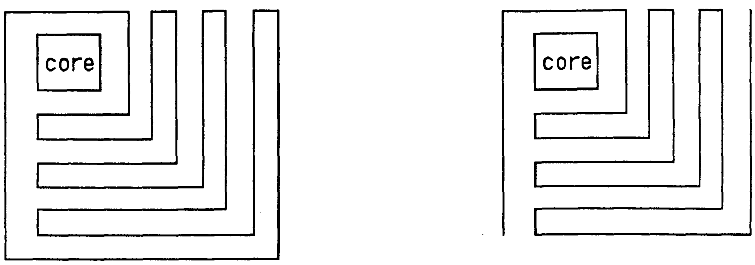
n	Conjectured Minimum Steiner Tree	Length
14		$62x + l = 177.73227$

FIGURE 5 (con't)

To construct Steiner trees for large square lattices with orders not equal to a power of 2, we will always use a "core" square with a "folded band of width 3" wrapped around it in various ways. (The only core squares we need are 6×6, 10×10, 14×14 for n even, and 0×0, 4×4 and 8×8 for n odd.) The general pattern looks like this:



n even $\qquad$ n odd

FIGURE 6

Each additional "fold" of the strip adds 6 to the size of the grid.

For example, for $n = 22$, we see that $22 \equiv 10 \pmod 6$ so we use a 10×10 core with 3 (doubled) folds of the band as shown in the picture.

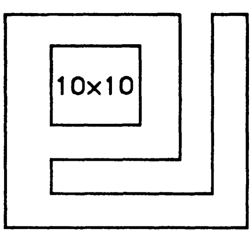
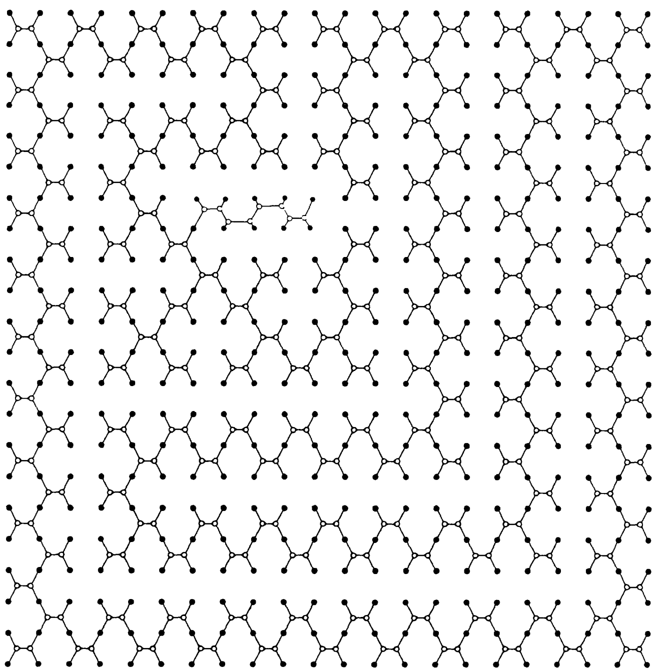


FIGURE 7

Of course, the strip must be broken and connected to the core. (See the detailed picture in FIGURE 8. We note that in order to make the connection to the strip, the corresponding place in the core must be "striplike".) When n is $6k$, the core is 0×0, i.e., empty, so we don't have to connect it to the band. In this case, we only have to break the band and reconnect the two isolated points with a T.



Conjectured minimum Steiner tree for 22×22 array.

FIGURE 8

For *odd* n, the band doesn't form a cycle but is open at each end, leaving two isolated points at the end. You can see this happening in the conjectured minimum Steiner trees for 9×9 and 13×13. In this case, when $n \neq 0 \pmod 6$ we only need an E to connect the core to the band. Summarizing these results we have: for $n \geqslant 15$, $n \neq 2^t$:

n	Length of conjectured minimum Steiner tree for G_n
$6k$	$(12k^2 - x + t)$
$6k+1$	$(12k^2 + 4k - 1)x + 3$
$6k+2$	$(12k^2 + 8k - 2)x + l$
$6k+3$	$(12k^2 + 12k + 2)x + 2$
$6k+4$	$(12k^2 + 16k + 2)x + l$
$6k+5$	$(12k^2 + 20k + 7)x + 3$

where $x = 1 + \sqrt{3}$, $l = \sqrt{35 + 20\sqrt{3}}$, and $t = (1 + \sqrt{3})/\sqrt{2}$. Of course, for $n = 2^t$, any right-thinking person would guess that the length of the minimum Steiner tree for G_{2^t} is just $(\frac{1}{3})(4^t - 1)x$ but unfortunately we can't even prove this for $t = 3$!

3. Square lattices for powers of 2

Here we will give the proof of the main result in this paper.

THEOREM. *If a rectangular array can be spanned by a Steiner tree made up entirely of X's, then the array is a square of size* 2^t *by* 2^t *for some* $t \geqslant 1$.

Proof. We start by giving each 2×2 "cell" a pair of coordinates (i, j) in the obvious way, where the lower left-hand cell has coordinates $(0,0)$.

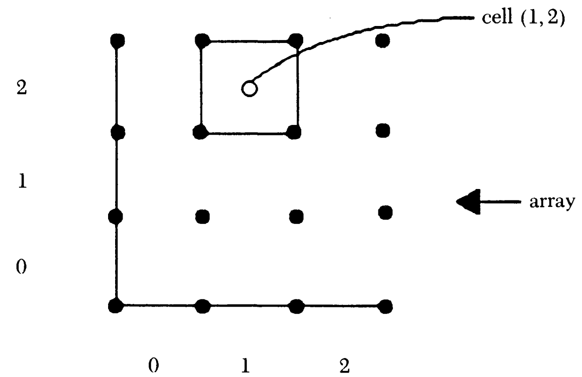


FIGURE 9

Let's call a cell (i, j) *even* if $i + j$ is even. Also, let's call a cell *occupied* if it has an X in it. *Suppose* our $a \times b$ array has an X-tree, that is a Steiner tree formed by X's.

Fact 1. Only *even* cells can be occupied.

Proof. The occupied cells must be connected. Furthermore, there can't be two *adjacent* occupied cells. Thus, occupied cells can only touch each other diagonally (as shown), in which case they are both even or both odd. However, the corner cell $(0,0)$ is occupied and even. Thus, *all* occupied cells are even.

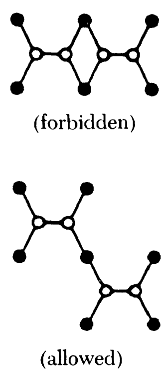


FIGURE 10

As a consequence, we see that a and b must both be even. Call an even cell (i, j) *doubly even* if i and j are both even. Otherwise, call it *doubly odd*.

Fact 2. Every doubly even cell must be occupied.

Proof. First note that *all* the even cells on the boundary must be occupied and are doubly even. Suppose we had some interior doubly even cell $(2i,2j)$ which was not occupied. Then, we must be able to draw a path P from the center of $(2i,2j)$ which goes to the outside of the array and doesn't pass through any occupied cell. (This is because the complement of our X-tree must be a connected set.) But if any even cell is *unoccupied* then *all four* of its even neighbors must be *occupied* (since, otherwise, one of its corner points would be disconnected.) This now implies that the only even cells P can pass through are doubly even ones, and never a doubly odd one.

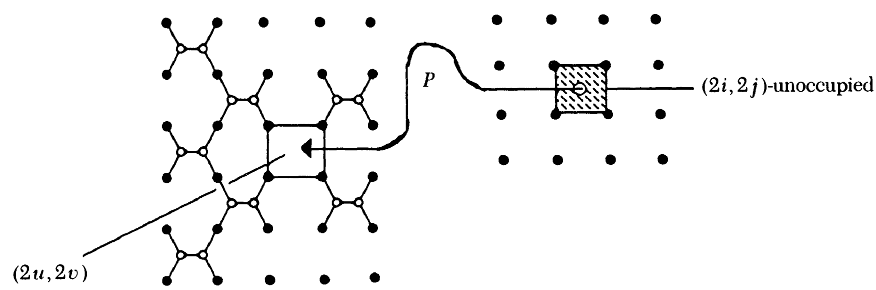


FIGURE 11

Now, focus on a doubly even cell $(2u,2v)$ just inside the boundary which P tries to pass through on its way to the outside. Since $(2u,2v)$ must be unoccupied (because P is going through it), all of its neighbors must be occupied. In particular, this forms a *barrier* with the adjacent occupied boundary cells which prevents P from going through to the outside here. But this happens wherever P tries to reach the outside since all the even cells on the boundary are occupied. Thus, P can never reach the outside, which is a contradiction. Hence, the hypothesis that there is an unoccupied doubly even cell is untenable, and the assertion is proved.

Therefore, in order to know what our X-tree is, we only have to know which (additional) *doubly odd* cells are occupied. Look at the picture for an 8×8 array in FIGURE 12.

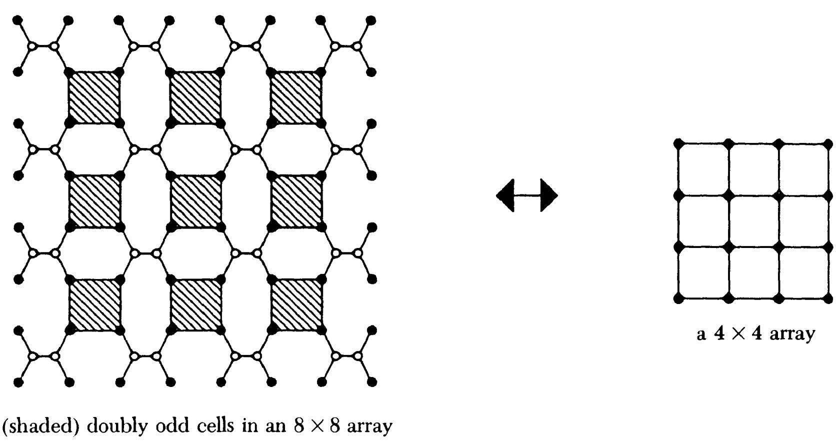


FIGURE 12

Next to the 8×8 array, a 4×4 array is drawn. Notice that there is natural correspondence between the array of 9 doubly odd cells (shaded) in the 8×8 and the array of 9 cells in the 4×4. The key observation now, which is not hard to check, is that *the set of occupied doubly odd cells in the* 8×8 *must correspond exactly to an X-tree* in the 4×4 array. Namely, consecutive doubly odd cells in the same row or column cannot both be occupied (or we get a cycle), and all points in the array must be joined together.

In effect, the doubly odd cells on the 8×8 array form a "stretched-out" version of all the cells of a 4×4 array.

More generally, this argument shows that if an $a \times b$ array has an X-tree then $a = 2A$, $b = 2B$, and, furthermore, the smaller $A \times B$ array must also have an X-tree.

We now can apply this repeatedly (similar to Fermat's method of infinite descent, except that we stop at 2×2) to get the conclusion that only $2^t \times 2^t$ arrays can have X-trees. This completes the proof of the Theorem.

Concluding Remarks

Of course, the main open problem is to determine the minimum Steiner trees for all (or even infinitely many) square lattices. It is embarrassing that even for 2^t by 2^t arrays, we still can't prove optimality for the "obviously" correct X-tree.

If we assume the distance between adjacent vertices in the lattice is 2, then it is not hard to show that any Steiner tree (minimum or not) has length of the form $\sqrt{a + b\sqrt{3}}$ where a and b are integers (possibly negative). Conceivably, this fact could be of help in proving optimality in some cases.

An interesting related question is to use only the 2×2 minimum Steiner trees and the smallest possible number of single edges (called E's) to form a spanning tree of $G(m, n)$. We know that for m and n both large enough, only a bounded number of E's are ever needed. When m is small however (where we can assume $m \leqslant n$), we may need arbitrarily many E's. Some examples of this are illustrated in FIGURE 13.

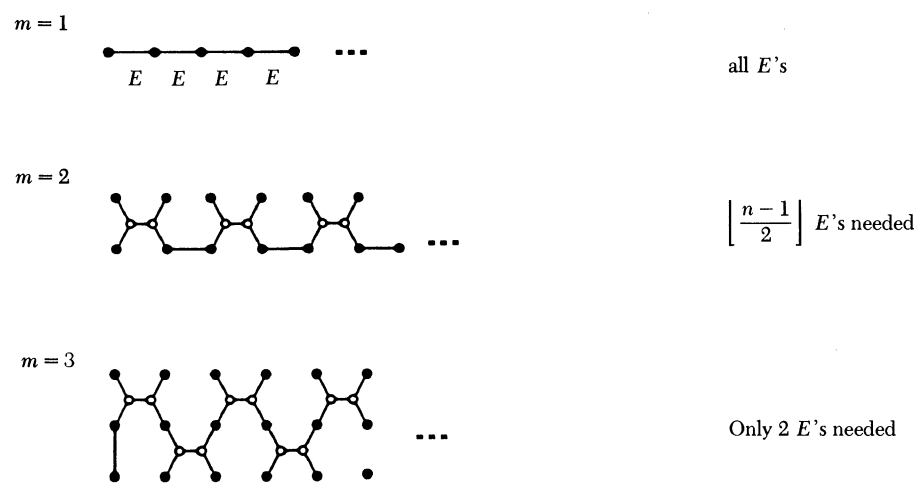
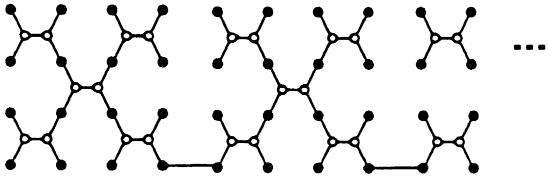


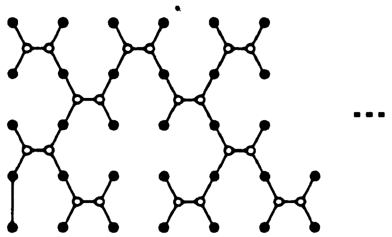
FIGURE 13

$m = 4$



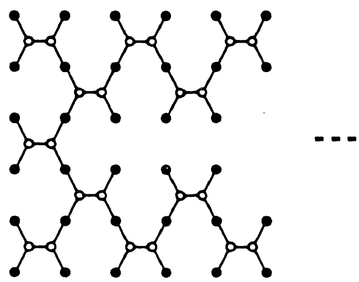
$\left\lfloor \frac{n-1}{4} \right\rfloor$ E's needed

$m = 5$



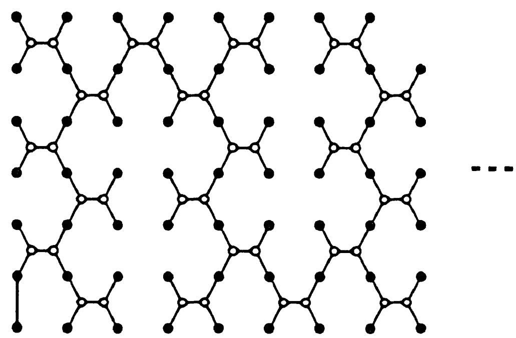
A *bounded* number of E's needed (at most 4)

$m = 6$



At most 2 E's needed

$m = 7$



At most 4 E's needed

FIGURE 13 (con't)

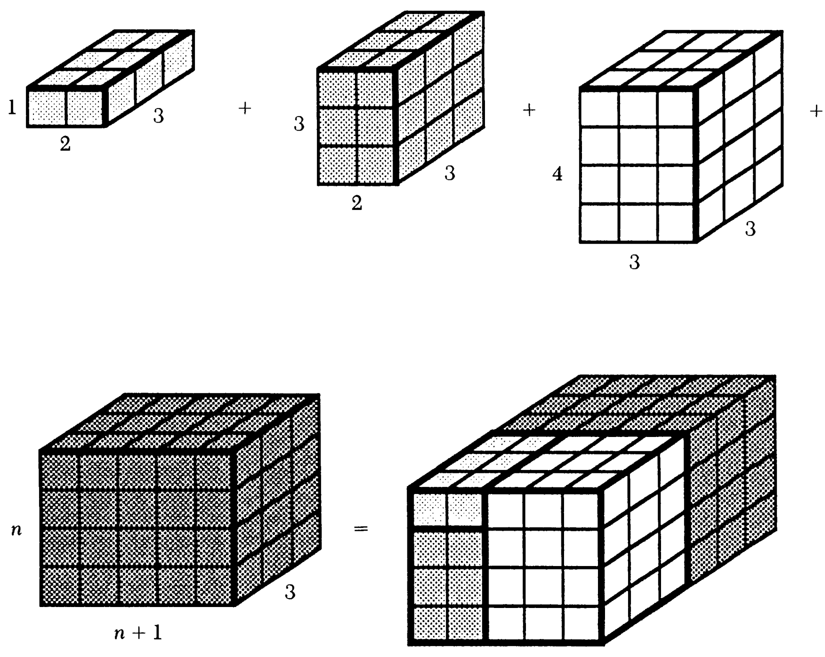
The authors would like to acknowledge the meticulous help of Nancy Davidson in preparing the figures for this paper.

REFERENCES

1. E. Carlson, private communication.
2. F. R. K. Chung and E. N. Gilbert, Steiner trees for the regular simplex, *Bull. Inst. Math. Acad. Sinica* 4 (1976), 313–325.
3. F. R. K. Chung and R. L. Graham, Steiner trees for ladders, *Annals of Discrete Math.* 2 (1978), 173–200.
4. ______, A new bound for Euclidean Steiner minimal trees, *Annals of the N.Y. Acad. Sci.* 440 (1985), 328–346.
5. E. J. Cockayne and D. E. Hewgill, Exact computation of Steiner minimal trees in the plane, *Inform. Proc. Letters* 22 (1986), 151–156.
6. M. Gardner, Mathematical games, *Scientific American* (June 1986), 16–22.
7. M. R. Garey, R. L. Graham and D. S. Johnson, The complexity of computing Steiner minimal trees, *SIAM J. Appl. Math.* 32 (1977), 835–859.
8. M. R. Garey and D. S. Johnson, *Computers and Intractability, a Guide to the Theory of NP-Completeness*, W. H. Freeman and Co., San Francisco, CA, 1979.
9. E. N. Gilbert and H. O. Pollak, Steiner minimal trees, *SIAM J. Appl. Math.* 16 (1968), 1–29.
10. R. Kallman, On a conjecture of Gilbert and Pollak on minimal trees, *Stud. Appl. Math.* 52 (1973), 141–151.
11. J. B. Kruskal, On the shortest spanning subtree of a graph and the traveling salesman problem, *Proc. Amer. Math. Soc.* 7 (1956), 48–50.
12. H. O. Pollak, Some remarks on the Steiner problem, *J. Combin. Theory Ser. A* 24(3) (1978), 278–295.
13. P. Winter, An algorithm for the Steiner problem in the Euclidean plane, *Networks* 15 (1985), 323–345.

Proof without Words: Sum of Special Products

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



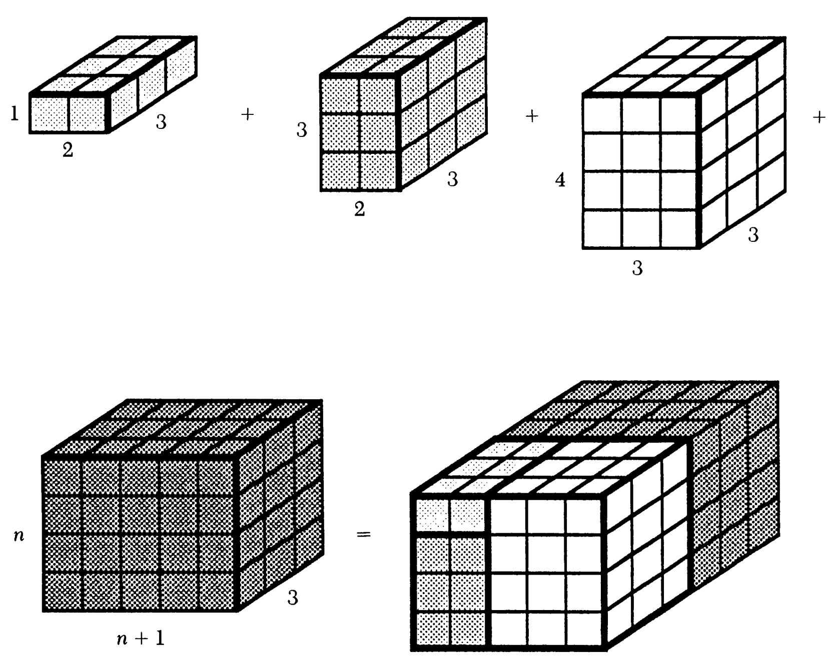
—Sidney H. Kung
Jacksonville University

REFERENCES

1. E. Carlson, private communication.
2. F. R. K. Chung and E. N. Gilbert, Steiner trees for the regular simplex, *Bull. Inst. Math. Acad. Sinica* 4 (1976), 313–325.
3. F. R. K. Chung and R. L. Graham, Steiner trees for ladders, *Annals of Discrete Math.* 2 (1978), 173–200.
4. ______, A new bound for Euclidean Steiner minimal trees, *Annals of the N.Y. Acad. Sci.* 440 (1985), 328–346.
5. E. J. Cockayne and D. E. Hewgill, Exact computation of Steiner minimal trees in the plane, *Inform. Proc. Letters* 22 (1986), 151–156.
6. M. Gardner, Mathematical games, *Scientific American* (June 1986), 16–22.
7. M. R. Garey, R. L. Graham and D. S. Johnson, The complexity of computing Steiner minimal trees, *SIAM J. Appl. Math.* 32 (1977), 835–859.
8. M. R. Garey and D. S. Johnson, *Computers and Intractability, a Guide to the Theory of NP-Completeness*, W. H. Freeman and Co., San Francisco, CA, 1979.
9. E. N. Gilbert and H. O. Pollak, Steiner minimal trees, *SIAM J. Appl. Math.* 16 (1968), 1–29.
10. R. Kallman, On a conjecture of Gilbert and Pollak on minimal trees, *Stud. Appl. Math.* 52 (1973), 141–151.
11. J. B. Kruskal, On the shortest spanning subtree of a graph and the traveling salesman problem, *Proc. Amer. Math. Soc.* 7 (1956), 48–50.
12. H. O. Pollak, Some remarks on the Steiner problem, *J. Combin. Theory Ser. A* 24(3) (1978), 278–295.
13. P. Winter, An algorithm for the Steiner problem in the Euclidean plane, *Networks* 15 (1985), 323–345.

Proof without Words: Sum of Special Products

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



—Sidney H. Kung
Jacksonville University

NOTES

A Diophantine Equation from Calculus

GEORGE P. GRAHAM
CHARLES E. ROBERTS
Indiana State University
Terre Haute, IN 47809

1. Introduction

In constructing exercises for homework and tests one often tries to choose constants in a problem so that the answer has a simple numerical form. For one frequently used calculus problem this effort leads to a Diophantine equation, equation (1) below. General solutions of this type of equation are well-known [**2**], [**4**], [**6**], [**7**], and especially see [**3**, pages 405–406]. In this investigation we will be interested in certain kinds of solutions. Also we will make use of a symmetry of the equation and the solutions which seems not to have been noticed previously. To this end we will develop the solution in a way that facilitates the use of the symmetry. Furthermore we include a table of certain solutions suitable for new variations of the original calculus problem.

2. The Problem

The problem is the familiar one from calculus: Construct an open box from a rectangular piece of tin of dimensions a and b by cutting equal squares of edge x from the corners and bending up the resulting tabs. Find x which maximizes the volume of the box.

The dimensions a and b in use are (a, a), any number; integer multiples of $(a, b) = (5, 8)$; integer multiples of $(a, b) = (8, 15)$; and integer multiples of $(a, b) = (16, 21)$. No texts [**1**], [**5**], [**8**], [**9**], [**10**], [**11**], [**12**], [**13**] that we have found use any other dimensions that will lead to a rational value for x, yet there are many such pairs.

The volume V of the resulting box has formula $V = x(a-2x)(b-2x)$, and has maximum value when $x = [(a+b) - \sqrt{a^2+b^2-ab}\,]/6$. Thus x is rational when a, b, and c are integers with

$$a^2 + b^2 - ab = c^2. \tag{1}$$

In particular we examine the solutions in which a, b, and c are positive, relatively prime integers with $a < b$. We call these the *primitive* solutions.

3. Square Solutions

When $a = b$, (1) becomes $a^2 = c^2$ and every pair of integers (a, a) provides the rational solution for x, $x = a/6$. This is not very interesting.

4. Interesting Solutions

Now we consider solutions with $a < b$. There is a symmetry in the set of solutions, namely if (a, b, c) is a solution then so is $(b-a, b, c)$. This is easily verified by substitution in (1). It is also easy to verify that when (a, b, c) is primitive then $(b-a, b, c)$ is primitive, and different from the given solution.

Equation (1) is transformed by

$$\begin{aligned} a &= n - m \\ b &= n + m \end{aligned} \tag{2}$$

to obtain

$$n^2 + 3m^2 = c^2 \tag{3}$$

(Such a transformation follows standard procedure for simplifying quadratic forms.)

The inverse transformation to (2) is

$$\begin{aligned} m &= (b-a)/2 \\ n &= (b+a)/2, \end{aligned} \tag{4}$$

from which it follows that $0 < m < n$. The primitive solutions of (1) are related in a simple way to the primitive solutions of (3).

In discussing this relationship it is convenient to distinguish between solutions (1) (or (3)) in which a and b (m and n) are both odd, and those in which a and b (m and n) have opposite parity. The first we call *pure*, the second *mixed*.

CASE 1: The mixed primitive solutions of (3) correspond to the pure primitive solutions of (1).

Let (m, n) be a mixed primitive solution of (3) corresponding to the solution $a = n - m$, $b = n + m$. Clearly $0 < a < b$ and a and b are both odd. A common divisor of a and b must be odd, hence it will be a common divisor of $m = (b-a)/2$ and $n = (b+a)/2$. The only such divisor is 1, hence (a, b, c) is a pure primitive solution of (1).

When (a, b) is a pure primitive solution of (1) a similar argument will show that $m = (b-a)/2$, $n = (b+a)/2$ is a mixed primitive solution of (3), establishing the correspondence in this case.

CASE 2: The pure primitive solutions of (3) correspond to two times the mixed primitive solutions of (1).

Let (m, n, c^*) be a pure primitive solution of (3). Then $(n-m, n+m, c^*)$ is an integral solution of (1) in which $0 < n - m < n + m$. However, it is clear that each number is even since (n, m) is pure. Let $a = (n-m)/2$, $b = (n+m)/2$, and $c = c^*/2$. This is a positive integral solution of (1) which is mixed, otherwise, as in case (1), it would correspond by (4) to the mixed integral solution of (3), $(m/2, n/2, c^*/2)$. This would contradict the primitivity of (m, n, c^*).

This solution of (1) is, in fact, primitive, for any divisor of a, b, c will also divide $m = b - a$, $n = b + a$, and $c^* = 2c$, a sequence whose only common divisor is 1. Thus the pure primitive solution (m, u, c^*) of (3) corresponds to the mixed primitive solution $(a = (n-m)/2, b = (n+m)/2, c = c^*/2)$ of (1).

Starting with a mixed primitive solution of (1) one sees, through a similar argument, that such a solution corresponds through (4) to one-half a pure primitive solution of (3).

We will use the method given in [7] and the symmetry mentioned above for investigating the primitive solutions of (3), and hence of (1). First factor $3m^2 = c^2 - n^2$, obtaining $3m^2 = (c-n)(c+n)$. Let $d = \gcd(c-n, c+n)$. Then $3m^2 = rsd^2$ where $c - n = rd$, $c + n = sd$, $r < s$, and r and s are relatively prime. Furthermore, one of r and s is a square and the other is three times a square.

The factor d must be either 1 or 2. To see this suppose that $d \neq 1$. Now $d^2 | 3m^2$, so if $3|d$ it follows that $3|m$. But $d | [(c+n)-(c-n)] = 2n$, and if $3|d$, then $3|2n$ whence $3|n$. This is impossible since m and n are relatively prime. Therefore $d^2|m^2$ and $d|m$. From above we also have $d|2n$. Since $\gcd(m,n) = 1$ we conclude that $d = 2$.

Next we readily see that from $3|3m^2 = rsd^2$ and $d = 1$ or 2, one can conclude that $3|r$ and $3 \nmid s$ or $3 \nmid r$ and $3|s$. First consider the case in which $3|r$.

In this kind of solution

$$m^2 = (r/3)sd^2, \tag{5}$$

with $r/3$ and s relatively prime and both squares of integers, say $r/3 = u^2$ and $s = v^2$ where $\gcd(3u, v) = 1$. Substituting in (5) results in $m^2 = u^2v^2d^2$ whence

$$m = uvd, \quad 3u < v, \quad \gcd(3u, v) = 1. \tag{6}$$

Recalling that $2n = (c+n) - (c-n) = v^2d - 3u^2d$, we have

$$n = (v^2 - 3u^2)d/2. \tag{7}$$

When $d = 1$, v and u must both be odd. But when $d = 2$ then v and u are of opposite parity, for otherwise m and n would both be even.

Equations (6) and (7) express m and n in terms of u, v, and d. Substitution in (2) yields

$$\begin{aligned} a &= (v^2 - 2uv - 3v^2)d/2 \\ b &= (v^2 + 2uv - 3v^2)d/2. \end{aligned} \tag{8}$$

The conditions on u and v are $0 < 3u < v$ and $\gcd(3u, v) = 1$. When $d = 1$, u and v are both odd giving a pure primitive solution (m, n) for (3). Then (8) is twice a mixed primitive solution of (1). When $d = 2$, u and v have opposite parity, giving a mixed primitive solution of (3). The solution of (1) given by (8) is then a pure primitive solution.

All such choices of u, v, and d yield half of the primitive solutions of (3) and thus of (1). The other half, in which $3 \nmid r$ and $3|s$, can be found by a similar analysis. However the symmetry of (1) observed earlier interchanges these two kinds of primitive solutions. For $(a, b, c) \to (b-a, b, c)$ induces $(m, n, c) \to ((n-m)/2, (n+3m)/2, c)$ for solutions of (3). Then

$$\begin{aligned} c - (n+3m)/2 &= [c + (c-n) - 3m]/2 \\ c + (n+3m)/2 &= [3c - (c-n) + 3m]/2. \end{aligned} \tag{9}$$

If $3|c-n$ (a solution of the first kind), primitivity of (m, n, c) implies that $3 \nmid c$, and from (9) it follows that $3 \nmid c - (n+3m)/2$, and 3 does divide $c + (n+3m)/2$. The new solution is therefore of the second kind.

TABLE 1 lists some solutions of the first kind, obtained by taking values of u, v, and d. TABLE 2 lists transformed solutions. Each table provides new integer pairs a and b for dimensions for the rectangular box problem.

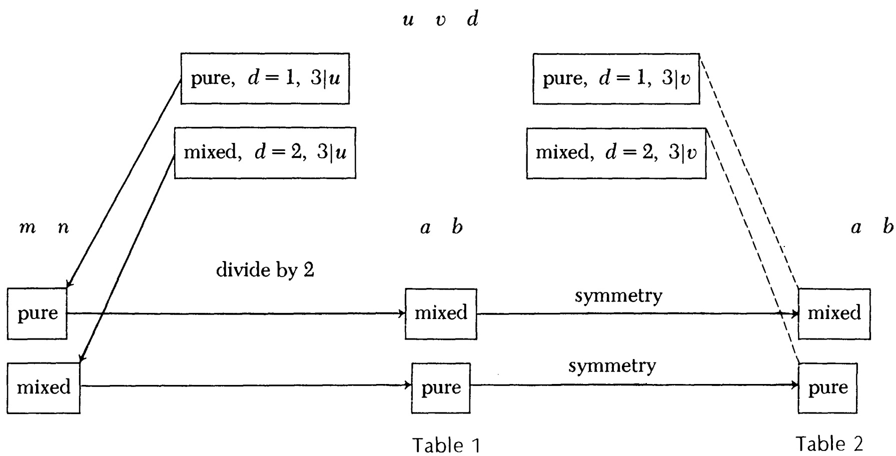
TABLE 1

u	v	d	m	n	a	b	c
1	5	1	5	11	3	8	7
1	7	1	7	23	8	15	13
3	11	1	33	47	7	40	37
5	17	1	85	107	11	96	91
7	23	1	191	161	15	176	169
1	4	2	8	13	5	21	19
2	7	2	28	37	9	65	61
3	10	2	60	73	13	133	127
4	13	2	104	121	17	225	217

TABLE 2

a	b	c
5	8	7
7	15	13
33	40	37
85	96	91
161	176	169
16	21	19
56	65	61
120	133	127
208	225	217

A diagram relating the values of u, v, d with (m, n) and in turn with (a, b) is given below.



5. A Final Comment

The symmetry observed in the solutions of (1) is a special case of a symmetry for equations of the form

$$a^2 + b^2 - qab = c^2. \tag{10}$$

The symmetry is $(a, b, c) \to (qb - a, b, c)$, a symmetry of order 2.

By the substitution

$$\begin{aligned} a &= (q^2 - 2)m + qn \\ b &= qm + n \end{aligned} \tag{11}$$

equation (9) changes to

$$(4 - q^2)m^2 + n^2 = c^2 \tag{12}$$

and the corresponding symmetry for solutions of (12) is

$$(m, n, c) \to ([(q^2 - 2)m + qn]/2, [(q^3 - 4q)m + (q^2 - 2)n]/2, c).$$

REFERENCES

1. Dennis Berkey, *Calculus*, Saunders, 1984, page 194, square solution.
2. H. Davenport, *The Higher Arithmetic*, 2nd ed., Hutchison University Library, London, 1964.
3. Leonard E. Dickson, *History of the Theory of Numbers*, Vol. 2, Chelsea Publishing Company, 1952.
4. ______, *Introduction to the Theory of Numbers*, Dover, 1957.
5. John Fraleigh, *Calculus with Analytic Geometry*, 2nd ed., Addison-Wesley, 1985, page 186, arbitrary dimensions a and b.
6. Karl Friedrich Gauss, *Disquisitiones Arithmeticae*, Yale University Press, 1966.
7. A. O. Gelfond, *Solving Equations in Integers*, Little Mathematics Library, Mir Publishers, Moscow, 1981.
8. Larson and Hostetler, *Calculus*, 2nd ed., D. C. Heath and Co., 1982, page 203, square solution.
9. Munem and Foulis, *Calculus*, 2nd ed., Worth, 1984, page 212, $a = 8$, $b = 15$ and page 218, square solution.
10. Richard Silverman, *Calculus with Analytic Geometry*, Prentice-Hall, 1985, page 204, square solution and page 212, $a = 8$, $b = 15$.
11. E. W. Swokowski, *Calculus with Analytic Geometry*, Prindle, Weber, and Schmitt, 1983, page 183, $a = 16$, $b = 21$.
12. G. B. Thomas, *Calculus with Analytic Geometry*, 2nd ed., Addison-Wesley, 1956, page 96, square solution and page 103, $a = 8$, $b = 15$.
13. G. B. Thomas and R. L. Finney, *Calculus and Analytic Geometry*, 6th ed., Addison-Wesley, 1984, page 205, square solution and page 212, $a = 8$, $b = 15$.

How Small Is a Unit Ball?

DAVID J. SMITH
MAVINA K. VAMANAMURTHY
University of Auckland
Auckland, New Zealand

The volume of the cube of edge d in $\mathbb{R}^n$ is d^n so that, as the dimension n increases, this volume increases, stays constant, or decreases to zero according as $d > 1$, $d = 1$, or $d < 1$. The situation for the ball of radius r in $\mathbb{R}^n$ is quite different.

For $n = 0, 1, 2, \ldots$, and $r > 0$, let $V_n(r)$ denote the n-dimensional volume of the n-dimensional ball of radius r in $\mathbb{R}^n$. Then $V_0(r) = 1$, $V_1(r) = 2r$, $V_2(r) = \pi r^2$, $V_3(r) = (4/3)\pi r^3$, and in general

$$V_n(r) = \frac{r^n \pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)}.$$

The derivation of this formula for $V_n(r)$ is a useful pedagogical device, and studying the formula reveals some interesting properties. For example, $V_n(1)$ increases for $0 \leqslant n \leqslant 5$ and decreases for $5 \leqslant n < \infty$. Its maximum value is $V_5(1) = 8\pi^2/15$, and $\lim_{n \to \infty} V_n(1) = 0$. In fact $\sum_{n=0}^{\infty} V_n(r)$ converges for all $r > 0$, and $\sum_{n=0}^{\infty} V_{2n}(r) = e^{\pi r^2}$. Thus, for fixed r, $V_n(r)$ tends to zero as n tends to infinity.

We derive the volume formula in three ways. The first and second methods use cross-sections and Fubini's theorem, and the third uses a polar-coordinate transformation and some simple properties of determinants.

The asymptotic properties already mentioned and some others are then developed.

REFERENCES

1. Dennis Berkey, *Calculus*, Saunders, 1984, page 194, square solution.
2. H. Davenport, *The Higher Arithmetic*, 2nd ed., Hutchison University Library, London, 1964.
3. Leonard E. Dickson, *History of the Theory of Numbers*, Vol. 2, Chelsea Publishing Company, 1952.
4. ______, *Introduction to the Theory of Numbers*, Dover, 1957.
5. John Fraleigh, *Calculus with Analytic Geometry*, 2nd ed., Addison-Wesley, 1985, page 186, arbitrary dimensions a and b.
6. Karl Friedrich Gauss, *Disquisitiones Arithmeticae*, Yale University Press, 1966.
7. A. O. Gelfond, *Solving Equations in Integers*, Little Mathematics Library, Mir Publishers, Moscow, 1981.
8. Larson and Hostetler, *Calculus*, 2nd ed., D. C. Heath and Co., 1982, page 203, square solution.
9. Munem and Foulis, *Calculus*, 2nd ed., Worth, 1984, page 212, $a=8$, $b=15$ and page 218, square solution.
10. Richard Silverman, *Calculus with Analytic Geometry*, Prentice-Hall, 1985, page 204, square solution and page 212, $a=8$, $b=15$.
11. E. W. Swokowski, *Calculus with Analytic Geometry*, Prindle, Weber, and Schmitt, 1983, page 183, $a=16$, $b=21$.
12. G. B. Thomas, *Calculus with Analytic Geometry*, 2nd ed., Addison-Wesley, 1956, page 96, square solution and page 103, $a=8$, $b=15$.
13. G. B. Thomas and R. L. Finney, *Calculus and Analytic Geometry*, 6th ed., Addison-Wesley, 1984, page 205, square solution and page 212, $a=8$, $b=15$.

How Small Is a Unit Ball?

DAVID J. SMITH
MAVINA K. VAMANAMURTHY
University of Auckland
Auckland, New Zealand

The volume of the cube of edge d in $\mathbb{R}^n$ is d^n so that, as the dimension n increases, this volume increases, stays constant, or decreases to zero according as $d>1$, $d=1$, or $d<1$. The situation for the ball of radius r in $\mathbb{R}^n$ is quite different.

For $n=0,1,2,\dots,$ and $r>0$, let $V_n(r)$ denote the n-dimensional volume of the n-dimensional ball of radius r in $\mathbb{R}^n$. Then $V_0(r)=1$, $V_1(r)=2r$, $V_2(r)=\pi r^2$, $V_3(r)=(4/3)\pi r^3$, and in general

$$V_n(r)=\frac{r^n\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

The derivation of this formula for $V_n(r)$ is a useful pedagogical device, and studying the formula reveals some interesting properties. For example, $V_n(1)$ increases for $0\leqslant n\leqslant 5$ and decreases for $5\leqslant n<\infty$. Its maximum value is $V_5(1)=8\pi^2/15$, and $\lim_{n\to\infty}V_n(1)=0$. In fact $\sum_{n=0}^{\infty}V_n(r)$ converges for all $r>0$, and $\sum_{n=0}^{\infty}V_{2n}(r)=e^{\pi r^2}$. Thus, for fixed r, $V_n(r)$ tends to zero as n tends to infinity.

We derive the volume formula in three ways. The first and second methods use cross-sections and Fubini's theorem, and the third uses a polar-coordinate transformation and some simple properties of determinants.

The asymptotic properties already mentioned and some others are then developed.

The following basic properties of Gamma (Γ) and Beta (B) functions are used (cf. [4], p. 235 ff.):

(i) $x\Gamma(x) = \Gamma(x+1)$ for all $x > 0$.
(ii) $\Gamma(n+1) = n!$ for $n = 0, 1, 2, \dots$
(iii) $\Gamma(1/2) = \sqrt{\pi}$.
(iv) $\Gamma(x+1) \sim \sqrt{2\pi x}\,(x/e)^x$, $(x \to \infty)$.
(v) $B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx = \Gamma(p)\Gamma(q)/\Gamma(p+q)$, for all p, q positive.

1. First Method (cf. [1], p. 412)

Regard $\mathbb{R}^n$ as $\mathbb{R}^{n-2} \times \mathbb{R}^2$. Then $(x_1, \dots, x_n)$ is in the ball $B^n(r)$ if and only if $x_1^2 + x_2^2 + \cdots + x_{n-2}^2 + x_{n-1}^2 + x_n^2 \leqslant r^2$, that is, if and only if

$$x_1^2 + x_2^2 + \cdots + x_{n-2}^2 \leqslant r^2 - x_{n-1}^2 - x_n^2.$$

Hence,

$$V_n(r) = \int_{B^n(r)} dx_1\,dx_2 \cdots dx_n,$$
$$= \int_{B^2(r)} \left(\int_{B^{(n-2)}\left(\sqrt{r^2 - x_{n-1}^2 - x_n^2}\right)} dx_1 \cdots dx_{n-2} \right) dx_{n-1}\,dx_n.$$

By induction

$$V_n(r) = \frac{\pi^{(n-2)/2}}{\Gamma\left(\frac{n-2}{2}+1\right)} \int_{B^2(r)} \left(r^2 - x_{n-1}^2 - x_n^2\right)^{(n-2)/2} dx_{n-1}\,dx_n.$$

Using polar coordinates in $\mathbb{R}^2$ this expression becomes

$$\frac{\pi^{(n-2)/2}}{\Gamma(n/2)} \int_0^{2\pi} d\theta \int_0^r \left(r^2 - t^2\right)^{(n-2)/2} t\,dt = \frac{2\pi^{n/2}}{\Gamma(n/2)} \frac{r^n}{n} = \frac{\pi^{n/2} r^n}{\Gamma\left(\frac{n}{2}+1\right)}.$$

2. Second Method (cf. [1], p. 416)

Regard $\mathbb{R}^n$ as $\mathbb{R}^{n-1} \times \mathbb{R}$. Then

$$V_n(r) = \int_{B^1(r)} \left(\int_{B^{n-1}\left(\sqrt{r^2 - x_n^2}\right)} dx_1 \cdots dx_{n-1} \right) dx_n, \quad \text{and (by induction)},$$
$$= \frac{\pi^{(n-1)/2}}{\Gamma\left(\frac{n-1}{2}+1\right)} \int_{-r}^{r} \left(r^2 - x_n^2\right)^{(n-1)/2} dx_n,$$
$$= \frac{2\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)} \int_0^r \left(r^2 - x_n^2\right)^{(n-1)/2} dx_n, \quad \left(\text{set } x_n = r\sqrt{t}\right),$$
$$= \frac{2\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)} \frac{r^n}{2} \int_0^1 (1-t)^{(n-1)/2} t^{-(1/2)}\,dt,$$

$$= r^n \frac{\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)} B\left(\frac{n+1}{2}, \frac{1}{2}\right) = r^n \frac{\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)} \frac{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)},$$

$$= r^n \frac{\pi^{(n-1)/2}\pi^{1/2}}{\Gamma\left(\frac{n}{2}+1\right)} \qquad \left(\text{since } \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}\right),$$

$$= \frac{r^n \pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

3. Third Method

Lemma.

$$\int_0^\pi \sin\theta\, d\theta \int_0^\pi \sin^2\theta\, d\theta \cdots \int_0^\pi \sin^n\theta\, d\theta = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

Proof. For $n = 1$, the left-hand side is

$$\int_0^\pi \sin\theta\, d\theta = -\cos\theta\Big]_0^\pi = 2$$

and the right-hand side is

$$\frac{\pi^{1/2}}{\Gamma\left(\frac{3}{2}\right)} = \frac{\pi^{1/2}}{\frac{1}{2}\Gamma\left(\frac{1}{2}\right)} = 2.$$

For $n \geqslant 2$,

$$\int_0^\pi \sin^n\theta\, d\theta = 2\int_0^{\pi/2} \sin^n\theta\, d\theta, \qquad \left(\text{put } \sin\theta = \sqrt{t}\right),$$

$$= 2\int_0^1 \frac{t^{n/2}}{2\sqrt{t}\sqrt{1-t}}\, dt,$$

$$= \int_0^1 t^{(n-1)/2}(1-t)^{-1/2}\, dt = B\left(\frac{n+1}{2}, \frac{1}{2}\right),$$

$$= \frac{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)}.$$

Hence by induction,

$$\int_0^\pi \sin\theta\, d\theta \int_0^\pi \sin^2\theta\, d\theta \cdots \int_0^\pi \sin^n\theta\, d\theta = \frac{\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)} \frac{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)} = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

We introduce n-dimensional spherical coordinates as follows (c.f. [3], p. 211):

$$
\begin{aligned}
x_1 &= t \sin\theta_1 \sin\theta_2 \cdots \sin\theta_{n-1}\\
x_2 &= t \cos\theta_1 \sin\theta_2 \cdots \sin\theta_{n-1}\\
x_3 &= t \cos\theta_2 \sin\theta_3 \cdots \sin\theta_{n-1}\\
&\cdots\\
x_{n-1} &= t \cos\theta_{n-2} \sin\theta_{n-1}\\
x_n &= t \cos\theta_{n-1}
\end{aligned}
$$

where $0 \leqslant t \leqslant r$, $0 \leqslant \theta_i \leqslant \pi$ for $i = 2, \ldots, n-1$, and $0 \leqslant \theta_1 \leqslant 2\pi$. For brevity, denote $\sin\theta_i = s_i$, $\cos\theta_i = c_i$, $\cot\theta_i = k_i$, $i = 1, \ldots, n-1$. Then

$$
\frac{\partial(x_1, \ldots, x_n)}{\partial(t, \theta_1, \ldots, \theta_{n-1})} = \begin{vmatrix}
x_1/t & x_1 k_1 & x_1 k_2 & & x_1 k_{n-1}\\
x_2/t & -x_2/k_1 & x_2 k_2 & \cdots & x_2 k_{n-1}\\
x_3/t & 0 & -x_3/k_2 & & \cdots\\
 & & 0 & & \\
\cdots & & & & x_{n-1} k_{n-1}\\
x_n/t & 0 & 0 & \cdots & -x_n/k_{n-1}
\end{vmatrix}
$$

$$
= t^{-1} x_1 x_2 \cdots x_n k_1 k_2 \cdots k_{n-1} \begin{vmatrix}
1 & 1 & 1 & & \\
1 & -\frac{1}{k_1^2} & 1 & \cdots & \\
1 & 0 & -\frac{1}{k_2^2} & & \\
\cdots & & 0 & & 1\\
1 & 0 & 0 & \cdots & -\frac{1}{k_{n-1}^2}
\end{vmatrix}
$$

Using elementary column operations, we can reduce this to

$$
t^{-1} x_1 \cdots x_n k_1 \cdots k_{n-1} \begin{vmatrix}
1 & 0 & 0 & 0\\
1 & -\left(1 + \frac{1}{k_1^2}\right) & 0 & \\
1 & -1 & -\left(1 + \frac{1}{k_2^2}\right) & \cdots\\
 & \cdots & & 0\\
1 & -1 & \cdots & -\left(1 + \frac{1}{k_{n-1}^2}\right)
\end{vmatrix}
$$

$$
= t^{-1} x_1 \cdots x_n k_1 \cdots k_{n-1} (-1)^{n-1} \left(1 + \frac{1}{k_1^2}\right) \cdots \left(1 + \frac{1}{k_{n-1}^2}\right)
$$

$$
= t^{-1} x_1 \cdots x_n (-1)^{n-1} \frac{1}{s_1 \cdots s_{n-1}} \frac{1}{c_1 \cdots c_{n-1}} = (-1)^{n-1} t^{n-1} s_2 s_3^2 \cdots s_{n-1}^{n-2}.
$$

Hence,

$$V_n(r) = \int_{B^n(r)} dx_1\, dx_2 \cdots dx_n$$
$$= \int_0^r dt \int_0^{2\pi} d\theta_1 \int_0^{\pi} d\theta_2 \cdots \int_0^{\pi} d\theta_{n-1} \left| \frac{\partial(x_1, \dots, x_n)}{\partial(t, \theta_1, \dots, \theta_{n-1})} \right|$$
$$= \frac{r^n}{n} 2\pi \int_0^{\pi} \sin u\, du \cdots \int_0^{\pi} \sin^{n-2} u\, du.$$

By the above Lemma, this is equal to

$$\frac{r^n}{n} 2\pi \frac{\pi^{(n-2)/2}}{\Gamma\left(\frac{n-2}{2}+1\right)} = \frac{r^n \pi^{n/2}}{(n/2)\Gamma(n/2)} = \frac{r^n \pi^{n/2}}{\Gamma(n/2+1)}.$$

4. Asymptotic Properties

(i)

$$\sum_{n=0}^{\infty} V_n(r)$$

converges for all $r > 0$.

This is an immediate consequence of the following result and its proof.

THEOREM. *$\sum_{n=0}^{\infty} n^{(n+1)/2} V_n(r)$ converges if and only if $0 < r < \frac{1}{\sqrt{2\pi e}}$.*

Proof. Stirling's Formula states that

$$\Gamma\left(\frac{x}{2}+1\right) \sim \sqrt{\pi x}\left(\frac{x}{2e}\right)^{x/2}.$$

Hence,

$$V_n(r) = r^n \frac{\pi^{n/2}}{\Gamma(n/2+1)} \sim \frac{r^n \pi^{n/2}}{(\pi n)^{1/2} (n/2e)^{n/2}}$$
$$= \frac{r^n (2\pi e)^{n/2}}{\pi^{1/2} n^{(n+1)/2}}.$$

Thus,

$$\left[n^{(n+1)/2} V_n(r)\right]^{1/n} \sim \frac{r(2\pi e)^{1/2}}{\pi^{1/2n}},$$

so that

$$\lim_{n\to\infty} \left[n^{(n+1)/2} V_n(r)\right]^{1/n} = r\sqrt{2\pi e}.$$

By the Cauchy root test,

$$\sum_{n=0}^{\infty} n^{(n+1)/2} V_n(r)$$

converges for $r < 1/\sqrt{2\pi e}$ and diverges for $r > 1/\sqrt{2\pi e}$. If $r = 1/\sqrt{2\pi e}$, then

$\lim_{n\to\infty} n^{(n+1)/2}V_n(r) = 1/\sqrt{\pi} \neq 0$, and, hence, the series diverges.

(ii)

$$\sum_{n=0}^{\infty} V_n(r) = e^{\pi r^2}\left(1 + \frac{2}{\sqrt{\pi}}\int_0^{r\sqrt{\pi}} e^{-t^2}\,dt\right).$$

Proof. $V_{2n}(r) = \dfrac{(\pi r^2)^n}{n!}$, and

$$V_{2n+1}(r) = \frac{2r(2\pi r^2)^n}{1\cdot 3\cdot 5\cdots(2n+1)}.$$

Hence,

$$\sum_{n=0}^{\infty} V_{2n}(r) = e^{\pi r^2}.$$

To calculate $\Sigma V_{2n+1}(r)$, we note that

$$\sum_{n=0}^{\infty} \frac{n!}{(2n+1)!}x^n = \frac{2}{\sqrt{x}}e^{x/4}\int_0^{\sqrt{x}/2} e^{-t^2}\,dt$$

(cf. [2], no. 5.21.13).
Hence, putting $x = 4\pi r^2$,

$$\begin{aligned}\sum_{n=0}^{\infty} V_{2n+1}(r) &= 2r\sum_{n=0}^{\infty}\frac{(2\pi r^2)^n}{1\cdot 3\cdot 5\cdots(2n+1)}\\ &= 2r\frac{1}{r\sqrt{\pi}}e^{\pi r^2}\int_0^{r\sqrt{\pi}} e^{-t^2}\,dt\\ &= \frac{2}{\sqrt{\pi}}e^{\pi r^2}\int_0^{r\sqrt{\pi}} e^{-t^2}\,dt.\end{aligned}$$

(iii) For any $r > 0$ there is an $N > 0$ such that $V_n(r)$ is monotone decreasing for $n > N$.

Proof.

$$V_{2n}(r) = \frac{r^{2n}\pi^n}{n!}$$

$$V_{2n+1}(r) = \frac{2r(2\pi r^2)^n}{1\cdot 3\cdot 5\cdots(2n+1)}.$$

Hence,

$$\frac{V_{2n}(r)}{V_{2n+1}(r)} = \frac{1}{2r}\,\frac{3\cdot 5\cdots(2n+1)}{2\cdot 4\cdots(2n)},$$

and

$$\frac{V_{2n-1}(r)}{V_{2n}(r)} = \frac{1}{\pi r}\,\frac{2\cdot 4\cdots(2n)}{3\cdot 5\cdots(2n-1)}.$$

$V_n(r)$ monotone decreasing means

$$\frac{V_{2n}(r)}{V_{2n+1}(r)} > 1 \quad \text{and} \quad \frac{V_{2n-1}(r)}{V_{2n}(r)} > 1.$$

That is, $r < \min(a_n/2, b_n/\pi)$, where

$$a_n = \frac{3\cdot 5\cdots(2n+1)}{2\cdot 4\cdots(2n)} \quad \text{and} \quad b_n = \frac{2\cdot 4\cdots(2n)}{1\cdot 3\cdot 5\cdots(2n-1)}.$$

(Note that in this case, $r^2 < a_n b_n/2\pi = (2n+1)/2\pi$, hence $n + (1/2) > \pi r^2$.) Now $a_n = \prod_{i=1}^n (1 + 1/2i)$ and $b_n = \prod_{i=1}^n (1 + 1/(2i-1))$, hence both are strictly increasing and unbounded (cf [4], p. 32), so for each $r > 0$ there is an N such that $V_n(r)$ decreases for $n > N$. Direct calculation shows, for example, that $V_n(1)$ increases for $n \leqslant 5$ and decreases for $n > 5$, and $V_5(1) = 8\pi^2/15$.

In conclusion, we remark that analogous results are true for the n-area $\sigma_n(r)$ of the n-dimensional sphere $S^n(r)$, and similar arguments may be used since

$$\sigma_{n-1}(r) = \frac{n}{r} V_n(r).$$

REFERENCES

1. T. M. Apostol, *Calculus*, Volume 2, second edition, Wiley, 1969.
2. E. R. Hansen, *A Table of Series and Products*, Prentice-Hall, 1975.
3. K. Rogers, *Advanced Calculus*, Merrill, 1976.
4. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, Fourth Edition, Cambridge, 1927.

More on Incircles

HÜSEYIN DEMIR
CEM TEZER
Middle East Technical University
Ankara, Turkey

The contents of this note came into being during the authors' search for a "synthetic" proof of the following result by H. Demir (FIGURE 1):

"Consider a triangle ABC and points P, Q on the line segment BC. If the incircles of the subtriangles ABP and AQC are congruent then the incircles of the subtriangles ABQ and APC are congruent."

(Notice that the requirements of the "Five Circle Theorem" ([2]) are partly redundant.)

Singularly enough, this question turned out to be less accessible than a more general result which was conjectured at the very outset of our investigations and later proved by means of the methods which will constitute the body of the present work:

$V_n(r)$ monotone decreasing means

$$\frac{V_{2n}(r)}{V_{2n+1}(r)} > 1 \quad \text{and} \quad \frac{V_{2n-1}(r)}{V_{2n}(r)} > 1.$$

That is, $r < \min(a_n/2, b_n/\pi)$, where

$$a_n = \frac{3 \cdot 5 \cdots (2n+1)}{2 \cdot 4 \cdots (2n)} \quad \text{and} \quad b_n = \frac{2 \cdot 4 \cdots (2n)}{1 \cdot 3 \cdot 5 \cdots (2n-1)}.$$

(Note that in this case, $r^2 < a_n b_n / 2\pi = (2n+1)/2\pi$, hence $n + (1/2) > \pi r^2$.) Now $a_n = \prod_{i=1}^{n}(1 + 1/2i)$ and $b_n = \prod_{i=1}^{n}(1 + 1/(2i-1))$, hence both are strictly increasing and unbounded (cf [4], p. 32), so for each $r > 0$ there is an N such that $V_n(r)$ decreases for $n > N$. Direct calculation shows, for example, that $V_n(1)$ increases for $n \leqslant 5$ and decreases for $n > 5$, and $V_5(1) = 8\pi^2/15$.

In conclusion, we remark that analogous results are true for the n-area $\sigma_n(r)$ of the n-dimensional sphere $S^n(r)$, and similar arguments may be used since

$$\sigma_{n-1}(r) = \frac{n}{r} V_n(r).$$

REFERENCES

1. T. M. Apostol, *Calculus*, Volume 2, second edition, Wiley, 1969.
2. E. R. Hansen, *A Table of Series and Products*, Prentice-Hall, 1975.
3. K. Rogers, *Advanced Calculus*, Merrill, 1976.
4. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, Fourth Edition, Cambridge, 1927.

More on Incircles

HÜSEYIN DEMIR
CEM TEZER
Middle East Technical University
Ankara, Turkey

The contents of this note came into being during the authors' search for a "synthetic" proof of the following result by H. Demir (FIGURE 1):

"Consider a triangle ABC and points P, Q on the line segment BC. If the incircles of the subtriangles ABP and AQC are congruent then the incircles of the subtriangles ABQ and APC are congruent."

(Notice that the requirements of the "Five Circle Theorem" ([2]) are partly redundant.)

Singularly enough, this question turned out to be less accessible than a more general result which was conjectured at the very outset of our investigations and later proved by means of the methods which will constitute the body of the present work:

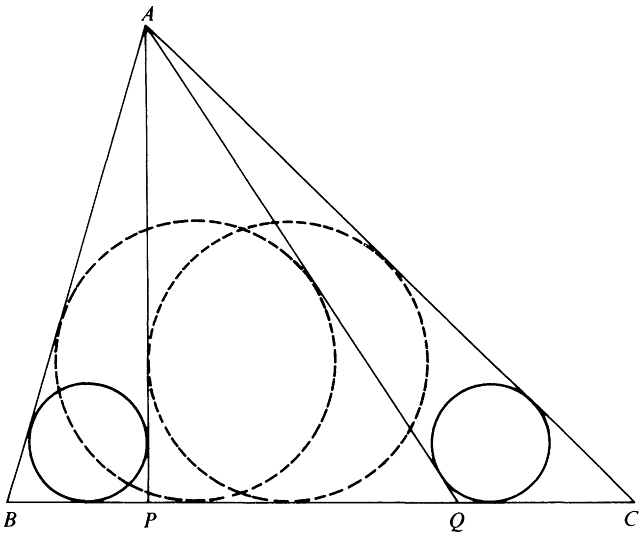


FIGURE 1

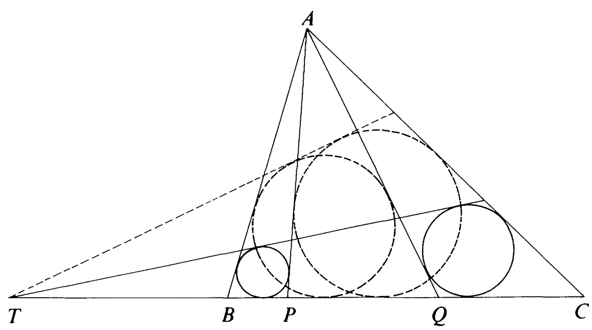


FIGURE 2

PROPOSITION 1 (FIGURE 2). *Consider a triangle ABC and points P, Q on the line segment BC. If T is the external homothety center of the incircles of the subtriangles ABP and AQC, then T is also the external homothety center of the incircles of the subtriangles ABQ and APC.*

It is not difficult to see that Proposition 1 implies the above mentioned result by H. Demir.

The authors found that Proposition 1 and several other interesting results could be obtained merely by a careful elucidation of necessary and sufficient conditions for a convex quadrangle to be circumscriptible.

1. Circumscriptible Quadrangles

In the following we consider a convex quadrangle $ABCD$, that is, a quadrangle which encloses a convex region that has as its boundary, the union of line segments AB, BC, CD, DA. Such a quadrangle will be said to be *circumscriptible* if there exists a circle lying in the convex region enclosed by the quadrangle, touching each side

AB, BC, CD, DA. We shall further exclude the triangle degeneracy by forbidding any three points from among A, B, C, D to be collinear.

The following simple result is quite standard (see p. 135 in [1] for a similar situation) and forms the basis of all our subsequent observations. The proof, which will be left to the reader as a mild challenge can be effected by repeated applications of the congruence of line segments that have one endpoint in common and are tangent to a circle at the other.

LEMMA (FIGURE 3). *Given a triangle AEF and points B, D on the line segments AE, AF, respectively, let ED, FB intersect in C. The following statements are equivalent*:

(i) *The convex quadrangle $ABCD$ is circumscriptible.*
(ii) $|AE| - |AF| = |CE| - |CF|$
(iii) $|BE| + |BF| = |DE| + |DF|$.

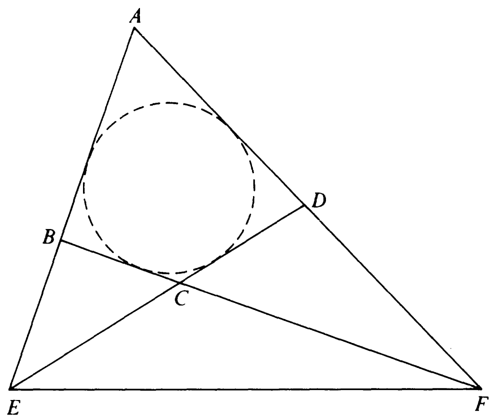


FIGURE 3

2. On the Third Incircle

Let AEF be a triangle, K, B and M, D be points on the line segments AE and AF, respectively, K, M lying nearer A than B, D (FIGURES 4, 5). Let EM and FK, EM and FB, ED and FB, ED and FK intersect in L, P, C, Q, respectively.

PROPOSITION 2 (FIGURE 4). *If any two from among the quadrangles $AKLM$, $ABCD$, $LPCQ$ are circumscriptible, then so is the third.*

Proof. Assume without loss of generality that $AKLM$ and $LPCQ$ are circumscriptible. By the Lemma

$$|AE| - |AF| = |LE| - |LF|$$

as $AKLM$ is circumscriptible and

$$|LE| - |LF| = |CE| - |CF|$$

as $LPCQ$ is circumscriptible. Therefore,

$$|AE| - |AF| = |CE| - |CF|.$$

Consequently, $ABCD$ is circumscriptible.

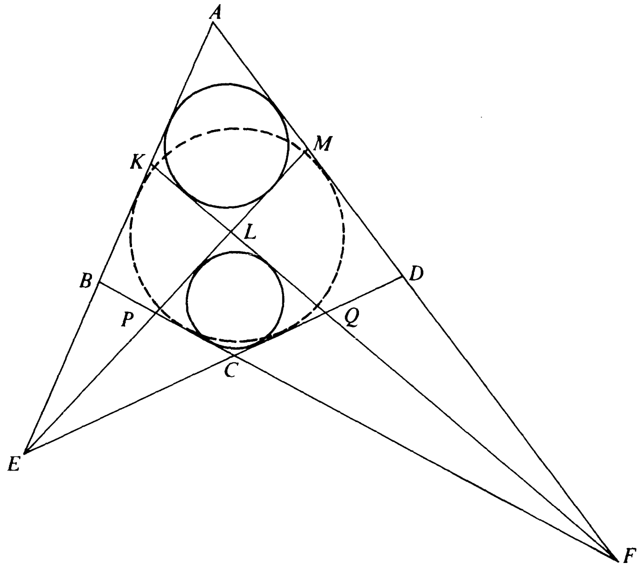


FIGURE 4

PROPOSITION 3 (FIGURE 5). *If any two from among the quadrangles KBPL, ABCD, MLQD are circumscriptible, then so is the third.*

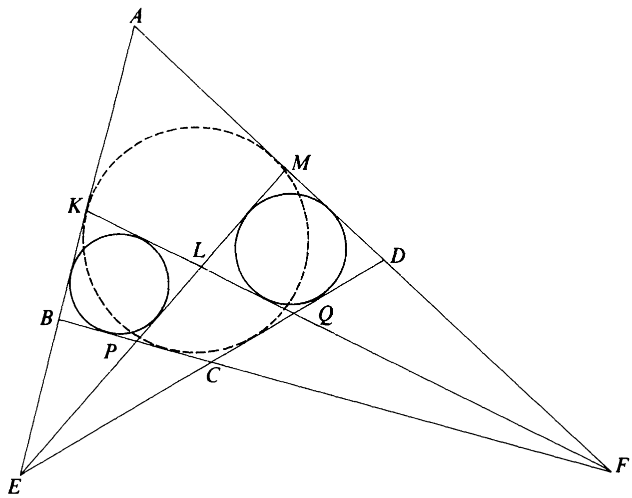


FIGURE 5

Proof. Assume without loss of generality, that $KBPL$ and $MLQD$ are circumscriptible. By the Lemma

$$|BE| + |BF| = |LE| + |LF|$$

as $KBPL$ is circumscriptible. Similarly

$$|LE| + |LF| = |DE| + |DF|$$

as $MLQD$ is circumscriptible. Combining these equations we obtain

$$|BE| + |BF| = |DE| + |DF|.$$

Therefore, $ABCD$ is circumscriptible.

Proof of Proposition 1 (FIGURE 6). Let Γ_1, Γ_2 be the incircles of the subtriangles ABP, AQC, respectively, with external homothety center T. Let the second common tangent of Γ_1, Γ_2 through T intersect AB, AP, AQ, AC in B', P', Q', C', respectively. Let Γ be the incircle of the triangle $AP'Q'$ and the second tangent to Γ through T intersect AB, AP, AQ, AC in B'', P'', Q'', C'', respectively. As $BPP'B'$ and $P'Q'Q''P''$ are circumscriptible so is $BQQ''B''$. Similarly as $QCC'Q'$ and $P'Q'Q''P''$ are circumscriptible so is $PCC''P''$. The incircles of $BQQ''B''$ and $PCC''P''$ are incircles of the subtriangles ABQ and APC respectively. TC'' is obviously their second common tangent through T.

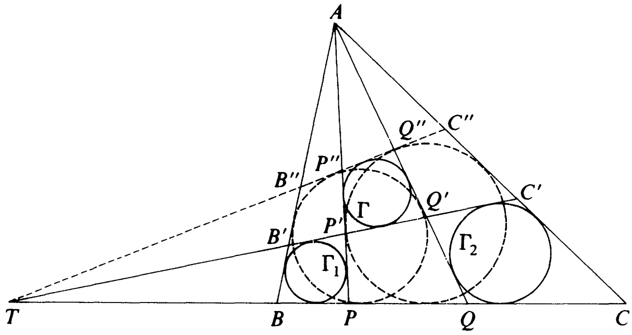


FIGURE 6

3. On the Fourth Incircle

By Propositions (2), (3) it is not difficult to see that if any three of the quadrangles $AKLM$, $KBPL$, $LPCQ$, $MLQD$ are circumscriptible, then so is the fourth (FIGURES 4, 5). A method which is traditionally ascribed to Fedorov but was possibly known to other and earlier mathematicians provides a straightforward proof of a much more general result and a simple relation between the inradii of the quadrangles in question. For a source on Fedorov's method we refer the reader to §78 in [3].

The method of Fedorov concerns the tangency of cycles (circles with orientation) and directed lines. In this connection tangency is expected to respect orientation. For instance, in FIGURE 7 the tangent directed line a goes "with" the cycle, that is, respects its orientation, whereas b does not. The essential idea is to assign to each cycle in the plane a point in space and to each pair of directed lines in the plane (except for well-isolated cases) a line in space in a one-to-one manner such that two

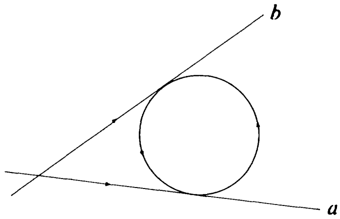


FIGURE 7

pairs of directed lines in the plane have a common tangent cycle if and only if the corresponding pair of lines in space intersect.

This "assignment" or "correspondence" is possibly best described by identifying space with $\mathbb{R}^3$ and the plane with $\mathbb{R}^2$. We assign the point $(a, b, \pm r)$ to the circle $(x-a)^2+(y-b)^2=r^2$, taking the plus sign if the circle is counterclockwise oriented, taking the minus sign if the circle is clockwise oriented. Consider, on the other hand, a pair of directed lines which is not a pair of parallel directed lines with the same direction. It can be routinely checked that the set of points in space corresponding to the cycles tangent to the given pair of directed lines in the plane is a line in space. To each pair of directed lines in the plane which is not a pair of parallel directed lines with the same direction, we assign the line in space specified as above. A simple inspection corroborates that these assignments fulfill the requirements put forth at the beginning of this section.

PROPOSITION 4 (FIGURE 8). *Consider the triangle* A_1EF *and points* B_1, A_2, B_2 *on the line segment* A_1E, *points* D_1, A_4, D_4 *on the line segment* A_1F *in order of increasing distance from* A_1. *Let* ED_1 *intersect* FB_1, FA_2, FB_2 *in* C_1, D_2, C_2; *let* EA_4 *intersect* FB_1, FA_2, FB_2 *in* B_4, A_3, B_3; *and let* ED_4 *intersect* FB_1, FA_2, FB_2 *in* C_4, D_3, C_3, *respectively. The quadrangles* $A_iB_iC_iD_i$, $i=1,2,3,4$ *are circumscriptible if any three of them are. If* $A_iB_iC_iD_i$ *are circumscriptible with inradii* r_i, $i=1,2,3,4$, *then*

$$\frac{1}{r_1}+\frac{1}{r_3}=\frac{1}{r_2}+\frac{1}{r_4}.$$

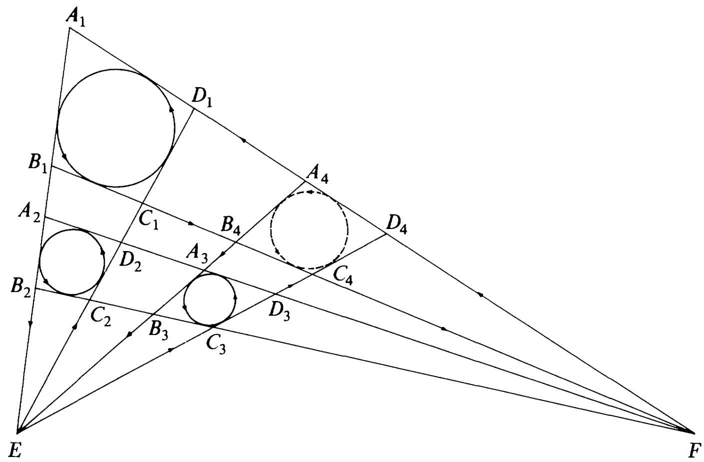


FIGURE 8

Proof. Let $A_iB_iC_iD_i$ be circumscriptible for $i=1,2,3$. Choosing suitable orientations for the circles and lines in question we obtain points K_i in space corresponding to the incircles of $A_iB_iC_iD_i$, $i=1,2,3$, and the points E, F in space corresponding to the point-circles E, F in the plane (FIGURE 9). Then K_1, K_2, E are collinear and K_2 lies between K_1 and E. Similarly K_2, K_3, F are collinear and K_3 lies between K_2 and F. Therefore K_1, K_2, K_3, E, F are coplanar and EK_3 intersects K_1F in a point K_4 between K_1 and F. Hence, $A_4B_4C_4D_4$ is circumscriptible.

Let $A_iB_iC_iD_i$ be circumscriptible with inradii r_i, $i=1,2,3,4$. Let K_1K_3, K_2K_4 intersect EF in L, M, respectively. Choose S, T on EF such that K_2S and K_4T are

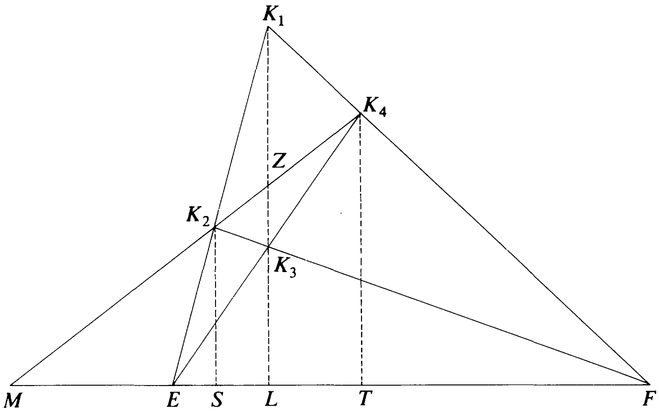


FIGURE 9

parallel to K_1K_3. And r_1, r_2, r_3, r_4 are proportional to $|K_1L|, |K_2S|, |K_3L|, |K_4T|$, respectively. Let K_1K_3 and K_2K_4 intersect in Z. Then K_1, Z, K_3, L form a harmonic division (Ch. 4 in [5]).

Hence,

$$\frac{2}{|LZ|} = \frac{1}{|K_1L|} + \frac{1}{|K_3L|}. \tag{1}$$

Similarly K_4, Z, K_2, M form a harmonic division. Hence,

$$\frac{2}{|MZ|} = \frac{1}{|K_2M|} + \frac{1}{|K_4M|}$$

unless M is at infinity, from which we obtain

$$\frac{2}{|LZ|} = \frac{1}{|K_2S|} + \frac{1}{|K_4T|}. \tag{2}$$

Combining (1) and (2) we obtain

$$\frac{1}{|K_1L|} + \frac{1}{|K_3L|} = \frac{1}{|K_2S|} + \frac{1}{|K_4T|}.$$

Consequently,

$$\frac{1}{r_1} + \frac{1}{r_3} = \frac{1}{r_2} + \frac{1}{r_4}.$$

The above described and employed method of assigning points of $\mathbb{R}^3$ to cycles in $\mathbb{R}^2$ is by no means the only one. We draw the attention of the reader to the "Six Circle Theorem" treated in §94.2 of [4] and in [6].

4. Conclusion

Our further investigations led us to a not unamusing but rather disconnected collection of lesser results. Instead of offering a list of them at the risk of incurring the impatience of our readers, we present a diagram which we like to call "the pseudolattice" (FIGURE 10). In the pseudolattice each quadrangle in which the sides are made up of the same number of segments—the pseudosquare, so to speak—is circumscriptible.

Through each lattice point there exist an ellipse and a hyperbola orthogonal to each other with common foci ∞_x, ∞_y which have the property that if they enter a pseudosquare by one vertex, they leave the same by the opposite vertex.

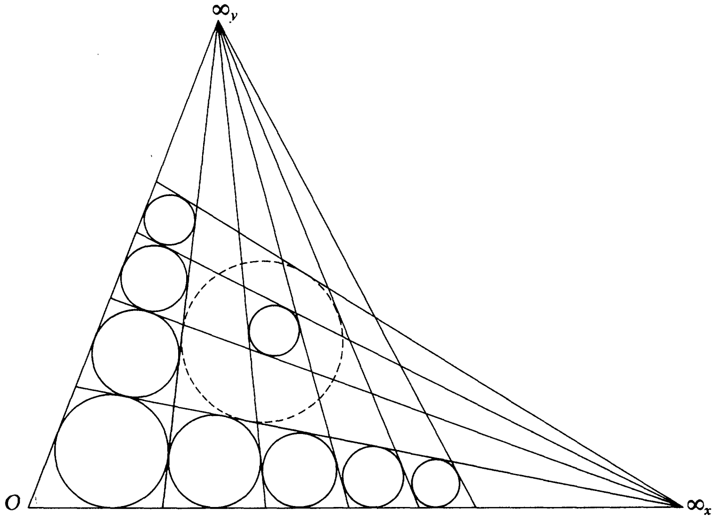


FIGURE 10

REFERENCES

1. N. Altshiller-Court, *College Geometry*, Barnes and Noble, 1952.
2. H. Demir, Incircles Within, this MAGAZINE 59 (1986), 77–83.
3. B. V. Kutuzov, *Geometry*, School Mathematics Study Group, Studies in Mathematics, No. 4, Chicago, 1960.
4. D. Pedoe, *A Course of Geometry for Colleges and Universities*, Cambridge University Press, 1970.
5. L. S. Shively, *An Introduction to Modern Geometry*, John Wiley and Sons Inc., New York, 1939.
6. J. F. Rigby, The geometry of the cycles and generalised Laguerre inversion, in *The Geometric Vein: The Coxeter Festschrift*, C. Davis, et al., editors, Springer, 1981, pp. 355–378.

I hold every man a debtor to his profession, from the which as men of course do seek to receive countenance and profit, so ought they of duty to endeavour themselves by way of amends to be a help and an ornament thereunto.

—Francis Bacon

Partial Fractions in Euclidean Domains

ROBERT W. PACKARD
STEPHEN E. WILSON
Department of Mathematics
Northern Arizona University
Flagstaff, AZ 86011

When we teach the technique of using partial fractions to integrate a rational function, we usually first tell the students that a rational function which is proper (i.e., in which the degree of the top is less than the degree of the bottom) can be broken up into partial fractions, and we show the students the algebraic tricks involved in solving for the constants. We then mention that to integrate any rational function whatever, first do the division to get a polynomial plus a proper rational function, and then apply the previous theory. Thus, we produce the two statements below. In each of these, $\mathbf{F}[x]$ stands for the set of polynomials over the reals.

Statement 1. If $f(x)$ and $g(x)$ are polynomials in $\mathbf{F}[x]$ with $\deg(f) < \deg(g)$, then $f(x)/g(x)$ may be written in the form

$$\frac{f(x)}{g(x)} = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \frac{s_{ij}(x)}{p_i(x)^j},$$

where $p_i(x)$ are the irreducible factors of $g(x)$, and $\deg(s_{ij}) < \deg(p_i)$ for all i, j.

Statement 2. If $f(x)$ and $g(x)$ are *any* polynomials in $\mathbf{F}[x]$, then $f(x)/g(x)$ may be written in the form

$$\frac{f(x)}{g(x)} = Q(x) + \sum_{i=1}^{n} \sum_{j=1}^{m_i} \frac{s_{ij}(x)}{p_i(x)^j},$$

where $Q(x)$ is a polynomial, $p_i(x)$ are the irreducible factors of $g(x)$, and $\deg(s_{ij}) < \deg(p_i)$ for all i, j.

Because $\mathbf{F}[x]$ is a Euclidean domain, and because the division algorithm figures prominently in the proof of the partial fraction result, a natural place to expect a generalization of this idea is in the theory of Euclidean domains. It is, perhaps, surprising that the natural generalization of Statement 2 holds in any Euclidean domain, while that of Statement 1 does not. And while the partial fraction expansion of a rational function is unique, this can not be expected to hold generally in all Euclidean domains.

Our most general setting will be a Euclidean domain R with valuation ϕ and field of quotients W. With this in mind, let us review definitions briefly. A **valuation** on an integral domain R is a function ϕ from $R\setminus\{0\}$ into any well-ordered set with the property that $\phi(ab) \le \phi(a)$ and $\phi(ab) = \phi(a)$ exactly when b is a unit. R is a **Euclidean domain** under ϕ provided that the Division Algorithm holds; i.e., for any a and b in R with $b \neq 0$, there are q and r in R such that $a = bq + r$ and either $r = 0$ or $\phi(r) < \phi(b)$. The following are examples of Euclidean domains:

(1) $\mathbf{F}[x]$, the polynomials with coefficients in some field $\mathbf{F}$, with $\phi(f) =$ degree of f.

(2) $\mathbf{Z}$, the integers, with $\phi(a) = |a|$

(3) $\mathbf{G}$, the Gaussian integers $= \{x + iy | x$ and y are integers$\}$, with $\phi(a) = \phi(x + iy) = x^2 + y^2$.

We will refer to these domains in the results and examples that follow. For other examples of interesting integral domains, see [2, Theorems 246, 247].

Note that the q and r guaranteed by the division algorithm are seldom unique. For example, in $\mathbf{Z}$ with $a=5$, $b=3$, we can use either $q=1$, $r=2$ (the natural choice), or $q=2$, $r=-1$. In $\mathbf{F}[x]$, q and r *are* unique for each a and b. In $\mathbf{G}$, for some a, b, one has up to four choices for q and r. If R is a Euclidean domain in which, for each a and b, q and r are unique, then R must, in fact, be a field or $\mathbf{F}[x]$ for some field $\mathbf{F}$. (This follows from [3, problem 4, p. 124].)

Our first lemma is basic for breaking up fractions into partial fractions.

Lemma 1. *If a, b, c are any elements of R such that $bc \neq 0$ and the greatest common divisor of b and c is 1, then there are elements d, e in R such that a/bc in W can be written $d/b + e/c$. Furthermore,*

$$\frac{d}{b}+\frac{e}{c}=\frac{D}{b}+\frac{E}{c}$$

implies that for some k in R, $D=d-kb$, $E=e+kc$.

Proof. Since $(b, c)=1$, and since R is a principal ideal domain, there are x, y in R such that $xc+yb=1$. See [1, p. 303]. Hence,

$$\frac{a}{bc}=\frac{axc+ayb}{bc}=\frac{ax}{b}+\frac{ay}{c}=\frac{d}{b}+\frac{e}{c}.$$

If

$$\frac{d}{b}+\frac{e}{c}=\frac{D}{b}+\frac{E}{c},$$

then $(d-D)c=(E-e)b$. Thus, $c|(E-e)$ and so there is an element k in R such that $E=e+kc$. Then $D=d-kb$.

For convenience, define $\operatorname{Rem}(b)$ by

$$\operatorname{Rem}(b)=\{r \mid r=0 \text{ or } \phi(r)<\phi(b)\}$$

for each $b \neq 0$ in a Euclidean domain R. Further, if x is in $\operatorname{Rem}(b)$, we will call the fraction x/b a **proper** fraction.

Referring to Lemma 1, if $d=pb+r$ where r is in $\operatorname{Rem}(b)$ and $e=qc+s$ where s is in $\operatorname{Rem}(c)$, we can rewrite the fractions d/b and e/c as $p+(r/b)$ and $q+(s/c)$, respectively. That is, each can be written as a quotient plus a proper fraction. Combining the quotients, we can rewrite a/bc as $Q+(r/b)+(s/c)$, where $Q=p+q$, and r/b and s/c are proper fractions. Because of nonuniqueness at several stages, there may be more than one candidate for Q. The following technical corollary implies that the range of possibilities for Q does not depend on the choices of d and e in Lemma 1. Thus, if we have exhausted the possibilities for Q with one choice of d and e, we have done so for all.

Corollary 1. *With b, c, d, e, D, E as in Lemma* 1,

$$\left\{p+q \,\middle|\, \begin{array}{l} D=pb+r,\ r\in \operatorname{Rem}(b) \\ E=qc+s,\ s\in \operatorname{Rem}(c) \end{array}\right\}=\left\{\hat{p}+\hat{q} \,\middle|\, \begin{array}{l} d=\hat{p}b+\hat{r},\ \hat{r}\in \operatorname{Rem}(b) \\ e=\hat{q}c+\hat{s},\ \hat{s}\in \operatorname{Rem}(c) \end{array}\right\}$$

Proof. For X, x in R, let

$$Q_x^X = \{ q | X = qx + r,\ r \in \mathrm{Rem}(x) \}.$$

Then the first set above is $Q_b^D + Q_c^E$ and the second is $Q_b^d + Q_c^e$. But since $D = d - kb$, $D - pb = d - (p + k)b$. It follows that $p \in Q_b^D \Leftrightarrow r = D - pb \in \mathrm{Rem}(b) \Leftrightarrow r = d - (p + k)b \in \mathrm{Rem}(b) \Leftrightarrow p + k \in Q_b^d$. Thus, $Q_b^d = Q_b^D + k$, and similarly, $Q_c^e = Q_c^E - k$. The conclusion follows directly.

We are now ready to formulate a generalization of Statement 2:

THEOREM 1. *If $a, b \in R$, $b = \prod_{i=1}^n p_i^{m_i}$, p_i irreducible, $n \geqslant 1$, then a/b in W can be written in the form*

$$Q + \sum_{i=1}^{n} \sum_{j=1}^{m_i} \frac{s_{ij}}{p_i^j},$$

where Q is in R and $s_{ij} \in \mathrm{Rem}(p_i)$. Call such an expression in proper fractions a **proper expansion** *of a/b.*

Proof. If $b = \prod_{i=1}^n p_i^{m_i}$, then using Lemma 1 inductively, we obtain

$$\frac{a}{b} = \sum_{i=1}^{n} \frac{d_i}{p_i^{m_i}}.$$

Successive division m_i times by p_i gives

$$\frac{d_i}{p_i^{m_i}} = q_{im_i} + \sum_{j=1}^{m_i} \frac{s_{ij}}{p_i^j},$$

where $s_{ij} \in \mathrm{Rem}(p_i)$ for each i, j. Letting $Q = \sum_{i=1}^n q_{im_i}$, we are done.

For particular domains, it is sometimes possible to obtain stronger results, including results corresponding to Statement 1. For example, the theorem which motivated this investigation, stated in full generality, is:

THEOREM 2. *If $f(x)$ and $g(x)$ are in $\mathbf{F}[x]$ (and here $\mathbf{F}$ is any field) then the fraction $f(x)/g(x)$ has a unique proper expansion. If, moreover, $\deg(f) < \deg(g)$ then the $Q(x)$ in that expression must be 0.*

Also, it is not hard to prove:

THEOREM 3. *If $|a| < |b|$ in $\mathbf{Z}$, then a/b has a proper expansion in which $Q = 0$.*

To illustrate Theorem 3, consider the following examples. In each example, the proper expansion is not unique. In the first example, nonuniqueness due to expansions of numerators is demonstrated. In the second example, nonuniqueness due to quotients is demonstrated.

Example 1. In $\mathbf{Z}$,

$$\frac{131}{3^2 2^4} = \frac{131[4 \cdot 16 - 7 \cdot 9]}{3^2 2^4} = \frac{524}{3^2} - \frac{917}{2^4}$$

$$= 58 + \frac{2}{3^2} - 58 + \frac{11}{2^4} = \frac{2}{3^2} + \frac{2^3 + 2 + 1}{2^4} = \frac{2}{3^2} + \frac{1}{2} + \frac{1}{2^3} + \frac{2}{2^4}$$

$$= \frac{3 - 1}{3^2} + \frac{2^3 + 2 + 1}{2^4} = \frac{1}{3} - \frac{1}{3^2} + \frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^4}.$$

Example 2. In **Z**,

$$\frac{11}{3^2 2^4} = \frac{44}{3^2} - \frac{77}{2^4}$$

$$= 4 + \frac{8}{3^2} - 4 - \frac{13}{2^4} = \frac{2\cdot 3+2}{3^2} - \frac{2^3+2+1}{2^4} = \frac{2}{3} + \frac{2}{3^2} - \frac{1}{2} - \frac{1}{2^3} - \frac{1}{2^4}$$

$$= 5 - \frac{1}{3^2} - 5 + \frac{3}{2^4} = -\frac{1}{3^2} + \frac{2+1}{2^4} = -\frac{1}{3^2} + \frac{1}{2^3} + \frac{1}{2^4}.$$

In each of these examples, our first step has been to apply the technique of proof of Lemma 1 to break the fraction into two fractions, with each denominator a power of a prime.

Theorem 1 generalizes Statement 2 to all Euclidean domains. To illustrate that Statement 1 will not so generalize, consider the fraction $(4+206i)/209$ in **G**, the Gaussian integers. The denominator $209 = 11\cdot 19$, and both 11 and 19 are prime in **G**.

Example 3. In **G**,

$$\frac{4+206i}{(19)(11)} = \frac{4+206i}{(19)(11)}[7\cdot 11 - 4\cdot 19]$$

$$= \frac{28+1442i}{19} - \frac{16+824i}{11} = \begin{pmatrix} 1+76i+\dfrac{9-2i}{19} \\ \text{or} \\ 2+76i-\dfrac{10+2i}{19} \end{pmatrix} - \begin{pmatrix} 1+75i+\dfrac{5-i}{11} \\ \text{or} \\ 2+75i-\dfrac{6+i}{11} \end{pmatrix}$$

$$= i + \frac{9-2i}{19} - \frac{5-i}{11} = i - \frac{10+2i}{19} + \frac{6+i}{11}$$

$$= 1 + i - \frac{10+2i}{19} - \frac{5-i}{11} = -1 + i + \frac{9-2i}{19} + \frac{6+i}{11}.$$

Remark. In Example 3, $Q=0$ is not attainable because all possible quotients have been listed, and by Corollary 1 all possible *sums* of quotients have been listed.

Summary. We have shown that in any Euclidean domain, any fraction has a proper expansion; that in **Z** and in **F**$[x]$, if the fraction is itself proper, the expansion may be chosen with $Q=0$. We have shown by examples in **Z** and **G** that neither this nor uniqueness can be expected to hold in general.

Two questions seem to us to merit further study: (1) In cases like Example 3 where $Q=0$ is not attainable for a proper fraction, can we limit Q a bit more precisely? For example, in **G**, we can show that if a/b is proper, it has a proper expansion in which $\phi(Q) < 9$. (2) What condition or conditions on the domain or on the valuation will guarantee that a proper fraction has a proper expansion in which Q may be taken to be 0?

REFERENCES

1. John B. Fraleigh, *A First Course in Abstract Algebra*, Addison-Wesley, Reading, MA 1982.
2. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press, 1983.
3. Nathan Jacobson, *Lectures in Abstract Algebra*, Van Nostrand, Princeton, NJ, 1966.

The Gunfight at the OK Corral

JAMES T. SANDEFUR
Georgetown University
Washington, DC 20057

The Movie. The final scene from the movie "The Good, the Bad, and the Ugly," is a three-way gunfight. Each gunfighter has exactly one bullet. They draw and fire, simultaneously, with each gunfighter firing at the better shot of his opponents.

In this article, we wish to consider a variation of this movie, in which the survivors of one round of the gunfight repeat the entire process (assuming there are at least two survivors). Thus, the gunfight continues, one round at a time, until there are less than two survivors. The questions we wish to answer are: for each gunfighter, what is the probability of surviving the entire gunfight? and what is the probability that no one survives?

This gunfight is an example of an absorbing Markov chain, and is discussed (as a three-way tank battle) in some finite mathematics books [1]. In this paper, we develop well-known results of absorbing Markov chains using the theory of difference equations, and then use these results to answer our questions. This is a natural approach in the sense that dynamical methods are used to analyze a dynamical model.

The Model. In this section, we will construct a mathematical model of the above gunfight. We need to have some information concerning each of the gunfighters. Let's designate our gunfighters as G, B and U (for Good, Bad, and Ugly). Suppose from previous observations we have determined (approximately) that: gunfighter G is a good shot and hits his target 60% of the time; gunfighter U is not as good a shot and hits his target 50% of the time; and that gunfighter B is a bad shot and hits his target only 30% of the time.

Thus, on the first round of the gunfight, G shoots at U, while U and B shoot at G. (Those who have seen the movie know that this is not the way it happened.)

We call each possible outcome of a round of the gunfight a **state**. There are 7 possible outcomes or states. The first is that no one survives the round, designated by N. We list each of the other states with the letters of the survivors; that is, G, U, B, GB, UB, and GUB.

Notice that GU is not a possible state. For GU to be a state, B must be shot first. But if all three gunfighters are alive, no one is shooting at B. Thus, B cannot be shot first. (We are excluding the possibility that B is hit by a ricocheting bullet.)

A state is called an **absorbing state** if, once that state is the result of one round, it will be the result of all following rounds. In other words, once that state is reached the game is over. The states, N, G, U, and B are absorbing states.

A state that is not an absorbing state is called a **nonabsorbing state**. These states are GB, UB, and GUB.

Let $p_1(n)$ be the probability that state N (nobody survives) occurs after n rounds of shots have been fired. Likewise, let $p_2(n)$, $p_3(n)$, and $p_4(n)$ be the probabilities that states G, U, and B have been reached after n rounds of shots have been fired, respectively. Let $q_1(n)$, $q_2(n)$, and $q_3(n)$ be the probabilities that the nonabsorbing states GB, UB, and GUB, respectively, are the result of the nth round of the gunfight.

Let's compute $p_1(n+1)$. There are three cases (or ways) in which nobody survives round $n+1$. The first case is that nobody survives round n (with probability $p_1(n)$). The second case is that only G and B survive round n (with probability $q_1(n)$) and then G and B shoot each other. The third case is that only U and B survive round n (with probability $q_2(n)$) and then U and B shoot each other.

The second case is a three-stage process. The first stage is that only G and B survive round n ($q_1(n)$). The second stage is that G shoots B (with probability .6). The third stage is that B shoots G (with probability .3). Observe that stages 2 and 3 occur simultaneously. Thus, the probability of the second case occurring is (using the multiplication principle for independent events) $.18q_1(n)$. Likewise, using three stages, we get that the probability of the third case is $(.5)(.3)q_2(n) = .15q_2(n)$. Adding the probabilities of the three cases, we get that

$$p_1(n+1) = p_1(n) + .18q_1(n) + .15q_2(n).$$

In a similar manner,

$$p_2(n+1) = p_2(n) + .42q_1(n),$$

and

$$p_3(n+1) = p_3(n) + .35q_2(n).$$

To get $p_4(n+1)$, the probability that only B survives the nth round, we must consider four cases. Three of them are easy. It is more difficult to compute the fourth case: that G, U, and B survive round n; that G shoots U; and that U and/or B shoots G.

Case 4 is a three-stage process. The first stage is that all three survive round n, which occurs with probability $q_3(n)$. The second stage is that G shoots U, which occurs with probability .6. The third stage is that U and/or B shoots G. The easiest way to compute this is to use the complement rule; that is, the probability we seek is 1 minus the probability that neither U nor B shoots G. The probability that U misses and B also misses is .5(.7). Thus the probability of the third stage is $1 - .35 = .65$. The product of the three stages, $.6(.65)q_3(n) = .39q_3(n)$, is the probability of the fourth case. Thus,

$$p_4(n+1) = p_4(n) + .12q_1(n) + .15q_2(n) + .39q_3(n).$$

In a similar manner,

$$q_1(n+1) = .28q_1(n) + .21q_3(n),$$

$$q_2(n+1) = .35q_2(n) + .26q_3(n),$$

and

$$q_3(n+1) = .14q_3(n).$$

We now notice that we can write the last three equations in the matrix form

$$Q(n+1) = RQ(n), \tag{1}$$

where

$$Q(n)=\begin{pmatrix} q_1(n) \\ q_2(n) \\ q_3(n) \end{pmatrix}, \quad \text{and } R=\begin{pmatrix} .28 & 0 & .21 \\ 0 & .35 & .26 \\ 0 & 0 & .14 \end{pmatrix}.$$

Note that $q_1(0)=0$, $q_2(0)=0$, and $q_3(0)=1$, since the probability is 1 that *GUB* are alive before the first round of the gunfight. Thus, we are given $Q(0)$. Equation (1) is a **first-order linear system of difference equations,** and $Q(0)$ is the **initial value** of the system of difference equations. Observe that we could now compute $Q(1)$, $Q(2)$, and so forth.

In a similar fashion, we can rewrite the first four equations as

$$P(n+1)=P(n)+SQ(n), \tag{2}$$

where

$$P(n)=\begin{pmatrix} p_1(n) \\ p_2(n) \\ p_3(n) \\ p_4(n) \end{pmatrix}, \quad \text{and } S=\begin{pmatrix} .18 & .15 & 0 \\ .42 & 0 & 0 \\ 0 & .35 & 0 \\ .12 & .15 & .39 \end{pmatrix}.$$

Note that $p_j(0)=0$ for $j=1$, 2, 3, and 4, so we are given $P(0)$. Since we (theoretically) know $Q(0)$, $Q(1)$, and so forth, we can compute $P(1)$, $P(2)$, and so forth.

Equation (2) is an example of an **absorbing Markov chain,** although in most undergraduate texts it is written in a different form. We will give a precise definition later.

We observe that the vector, P, given by

$$P=\lim_{n\to\infty} P(n),$$

gives the probabilities of each of the 4 absorbing states being the (eventual) outcome of the gunfight (if this limit exists). We could approximate P by computing $P(1)$, $P(2)$, and so forth.

The Method. We will find a closed form solution to equation (2), and use this to compute the vector, P. We will find the solution using the **method of undetermined coefficients.**

Suppose we are given a first-order linear difference equation

$$q(n+1)=rq(n),$$

as well as the initial value, $q(0)$. Induction shows that $q(1)=rq(0)$, $q(2)=rq(1)=r^2q(0)$, and

$$q(k)=r^kq(0).$$

Consider the first-order nonhomogeneous difference equation,

$$p(n+1)=sp(n)+ar^n,$$

where $p(0)$ is given and $s\neq r$. Using the method of undetermined coefficients, we try to find a **particular solution** to the equation of the form $p(n)=br^n$. We substitute this "guess" into the difference equation and solve for the constant b. In this case, $b = a/(r-s)$. The **general solution** is then found by adding the solution to the linear equation, $p(n+1) = sp(n)$, to this particular solution, giving

$$p(n) = cs^n + br^n.$$

We then use $p(0)$ to compute $c = p(0) - b$, to get the particular solution to the nonhomogeneous difference equation.

Suppose we have a Markov chain in which the nonabsorbing states satisfy the system of equations

$$Q(n+1) = RQ(n). \tag{1}$$

We note that the solution to this system of difference equations is

$$Q(n) = R^nQ(0).$$

Suppose that the absorbing states satisfy the system of equations

$$P(n+1) = P(n) + SQ(n). \tag{2}$$

Substituting $R^nQ(0)$ for $Q(n)$ gives the nonhomogeneous system of difference equations

$$P(n+1) = P(n) + SR^nQ(0). \tag{3}$$

Assume that the matrix, R, is such that R^n goes to zero; that is, each component of R^n goes to zero. We then call (3) an **absorbing Markov chain.** (R^n going to zero implies that $Q(n)$, the probabilities of being in each of the nonabsorbing states after n rounds, goes to zero.)

Alternatively, an absorbing Markov chain could be defined as a Markov chain in which there is at least one absorbing state, and that you can reach some absorbing state from each of the nonabsorbing states. These definitions are equivalent, although the proof of this fact is beyond the scope of this paper. The advantage of our definition is that if R^n goes to zero, then it is clear that $(I - R)$ is invertible.

The general solution of the linear part of (3) is $P(n) = A$, where A is a constant vector. By the method of undetermined coefficients, we look for a solution to the nonhomogeneous difference equation of the form

$$P(n) = S(R^n)C.$$

Since

$$P(n+1) = S(R^{n+1})C = S(R^n)RC,$$

then substitution into (3) gives

$$S(R^n)RC = S(R^n)C + S(R^n)Q(0).$$

Bringing all the terms to the left gives

$$S(R^n)RC - S(R^n)C - S(R^n)Q(0) = 0,$$

(where 0 is the zero vector). Factoring out $S(R^n)$ gives

$$S(R^n)(RC - C - Q(0)) = 0.$$

Thus, the equation is satisfied if

$$RC - C - Q(0) = 0; \quad \text{that is,} \quad \text{if } (R - I)C = Q(0),$$

or, after multiplying both sides by $(R - I)^{-1}$ on the left,

$$C = (R - I)^{-1}Q(0).$$

Thus, adding this particular solution to the nonhomogeneous difference equation to the general solution to the linear difference equation, implies that the general solution to (3) is

$$P(n) = A + S(R^n)(R - I)^{-1}Q(0).$$

Since $P(0)$ is the zero vector and $R^0 = I$,

$$P(0) = A + S(R^0)(R - I)^{-1}Q(0),$$

or

$$A = S(I - R)^{-1}Q(0).$$

Hence, the particular solution to the system (3) is

$$P(n) = S(I - R^n)(I - R)^{-1}Q(0).$$

Since R^n goes to 0 as n gets large,

$$\lim_{n \to \infty} P(n) = S(I - R)^{-1}Q(0).$$

In summary, we have derived the following well-known result. Suppose we have an absorbing Markov chain, given by the systems of difference equations (1) and (2). Suppose we *start* in one of the nonabsorbing states, given by the vector $Q(0)$, which will have a 1 in the position corresponding to the starting nonabsorbing state and 0's in all other positions. Then the probabilities of *eventually ending* in each of the absorbing states are given by the vector

$$P = S(I - R)^{-1}Q(0).$$

Remark. Suppose, instead of starting in one nonabsorbing state, you randomly choose your nonabsorbing state by flipping a coin, rolling a die, drawing a marble, or by some other random method. Then in the above computation, you only need to compute the correct $Q(0)$, where each component, $q_j(0)$, is the probability of picking the jth nonabsorbing state as the initial state.

To solve our gunfighter problem we compute

$$(I - R)^{-1} = \begin{pmatrix} 25/18 & 0 & 175/516 \\ 0 & 20/13 & 20/43 \\ 0 & 0 & 50/43 \end{pmatrix}$$

and

$$S(I - R)^{-1} = \begin{pmatrix} 1/4 & 3/13 & 45/344 \\ 7/12 & 0 & 49/344 \\ 0 & 7/13 & 7/43 \\ 1/6 & 3/13 & 97/172 \end{pmatrix}$$

$S(I-R)^{-1}Q(0)$ is the last column of the matrix $S(I-R)$.$^{-1}$ Therefore, the probability that everyone dies is $45/344 = 0.13$; that G wins is 0.14; that U wins is 0.16; and that the worst gunfighter, B, wins is 0.56. (There is .01 roundoff error.)

Remark. Let the vector, $A(n)$, be the expected number of times that we are in each nonabsorbing state after n rounds. The system of difference equations that gives $A(n)$ is

$$A(n+1) = A(n) + RQ(n).$$

Using the techniques given above, we get that

$$A(n) = (I-R)^{-1}(I-R^{n+1})Q(0).$$

Since R^n goes to zero, we have derived the well-known result that

$$A = (I-R)^{-1}Q(0)$$

gives the expected number of times the process will be in each of the nonabsorbing states. The sum of the components of A, 1.96 in this example, gives the expected number of rounds the gunfight will last.

The Motive. The standard method for finding P, the probabilities of absorption into each of the absorbing states, is by constructing and solving a system of linear equations [1]. This is a static method, in which the probabilities are assumed to exist and then computed. Students can understand the mathematics involved in this approach, while still not being intuitively sure why the answer is right.

The method used in this paper is dynamic. Loosely speaking, in a dynamic process the result after a finite time evolves toward the answer. For example, $P(n)$ converges to P in the discussion above.

Discrete dynamics, in the form of difference equations or recurrence relations, is a natural (but underused) approach to many mathematical models. This approach is easily accessible to those with only a good algebra background. For a more complete survey of elementary discrete dynamical methods, see [2].

Personally, I would like to see introductory mathematics courses (both for science and nonscience majors) use dynamical methods more extensively. This would help students to think in terms of "cause and effect", as well as to see the value of mathematics. We might even find that our students enjoy mathematics.

REFERENCES

1. A. W. Goodman and J. S. Ratti, *Finite Mathematics with Applications*, Macmillan, New York, 1975.
2. J. T. Sandefur, *Discrete Mathematics with Finite Difference Equations*, Lecture Notes, Mathematics Department, Georgetown University, 1983.

Some Program Anomalies and the Parameter Theorem

DAVID POKRASS
Clemson University
Clemson, SC 29634

1. Introduction

Recursive function theory studies functions that can be evaluated by an algorithm based on some computing model such as a Turing machine or computer program. These functions are called *partial recursive* (*p.r.*) and are defined on subsets of **N**, where $\mathbf{N} = \{0, 1, \dots\}$, with values in **N**. (A good introduction to the theory may be found in [**2**].) This paper explains the *Parameter Theorem*, illustrates some unusual computer programs, and explains their connection to recursive function theory. Basing our computing model on a modern high level language, we can *construct* programs whose existence is known through more formal means. Our functions will operate on *strings* rather than numbers, however. There is no loss in generality since strings can be viewed as members of **N**, and vice-versa. From a practical aspect, this approach is somewhat more desirable since conventional programming languages often allow making use of long strings, but place more restrictive limits on the size of integers. We first pose a problem: *Does there exist a program* (*written in some conventional programming language*) *which halts if and only if its input is identical to the program's text* (*i.e., source code*)?

For the remainder of this note X is the set of ASCII characters 'a,' 'b,' '+,' space, line feed, and so forth, needed to write programs in most programming languages. We denote by X^* the set of all *strings* (finite sequences) over X. Let P_1 be the set of all one variable *partial recursive functions* on X^*. A function Ψ belongs to P_1 if and only if there is a subset of S of X^* and some *algorithm* which computes Ψ by associating with each element of S a well-defined element of X^*. Similarly, Ψ belongs to P_2 if and only if there is a set $S \subseteq X^* \times X^*$, and an algorithm for Ψ which associates with each element of S a well-defined element of X^*. A partial recursive function Ψ in P_1 (or in P_2) is *recursive* if the algorithm is defined on all of X^* (or $X^* \times X^*$). This is equivalent to saying that the algorithm which evaluates Ψ always halts.

Our computing model will be based on Ada[†]; its high-level string-handling properties simplify later examples. For readability, however, we will express algorithms in a more familiar Pascal-like style. Readers familiar with either programming language should find forthcoming algorithms sufficiently clear. A formal presentation in Ada is given in the Appendix.

A string of ASCII characters, such as in FIGURE 1, represents a function in our language which takes strings to strings. Of course, strings of ASCII characters such as "The quick brown fox" and "%*!!" represent nothing at all in our programming language.

Let A_1 be the set of all strings in X^* representing functions in our language of one input STRING which outputs a STRING. Similarly let A_2 be the set of all strings in

[†]Ada is a trademark of the U.S. Department of Defense.

X^* representing functions of two input STRINGS which output a STRING. For reasons which will simplify our construction in the next section, names of main functions in A_1 and A_2 are always F. Let $A = A_1 \cup A_2$. An element of A is called a *program*.

```
function F(X,Y: STRING)
begin
  return X;
end;
```

FIGURE 1

Each member of A_1 (or A_2) defines or *implements* a p.r. function of one (or two) variables. FIGURE 1 shows a string belonging to A_2. This string implements the usual projection function $\Psi(x, y) = x$. According to our definition, this function is recursive since its algorithm always halts for any ordered pair of strings.

If σ is a program in A, then the set of inputs for which σ returns a string is the domain of the p.r. function implemented by σ. The domain of the function implemented by the program in FIGURE 1 is $X^* \times X^*$. On the other hand, the only string in the domain of the function implemented by the program in FIGURE 2 is "hello." Indeed, the program first checks the equality of the input string with "hello." If the two strings are equal the program returns "goodbye." Otherwise, the program drops into the infinite loop. By our definition, the function implemented by this program is partial recursive but not recursive.

```
function F(X: STRING)
begin
  if X = "hello" then
    return "goodbye";
  else
    loop forever;
end;
```

FIGURE 2

Programs such as the one in FIGURE 2 execute indefinitely and fail to return for certain inputs. Another way a program can fail to return is by encountering a run-time error such as an out-of-range array index. In such cases, we agree that the partial recursive function is undefined for x.

Replacing X in line 3 of FIGURE 2 with "jello" creates a program which never halts for any input. The function implemented by such a program is the *empty function* whose domain is the empty set.

For each string σ in X^* we define Ψ_σ^1 to be the function implemented by σ if σ is a member of A_1, and the empty function otherwise. Similarly, for each string σ in X^* we define Ψ_σ^2 to be the function implemented by σ if σ is a member of A_2, and the empty function otherwise. We have thus defined an indexing so that for each σ in X^*, Ψ_σ^1 and Ψ_σ^2 each define one member of P_1 and P_2, respectively. Such an indexing is called a *numbering* of the partial recursive functions. Note that for each p.r. function Ψ in P_1 there are infinitely many σ for which $\Psi = \Psi_\sigma^1$ and, similarly, for Ψ in P_2. For if σ implements Ψ, then by correctly inserting any finite number of spaces the new program will still implement Ψ.

2. Parameter Theorem

In this section we explain the Parameter Theorem which is central to many important results in recursive function theory. First we describe the important S function which maps an ordered pair of strings into a string.

Given two strings x and y, $S(x, y)$ is the string formed by taking the template of FIGURE 3 and inserting y and a *representation* of x in the places indicated.

```
function F(X: STRING)

y

begin
  return F(X, representation for x);
end;
```

FIGURE 3

For a concrete example, let x be the string "hi", and let y be the string shown in FIGURE 1. In this case the string that S produces appears in FIGURE 4.

```
function F(X: STRING)

function F(X, Y: STRING)
begin
  return X;
end;

begin
  return F(X, "hi");
end;
```

FIGURE 4

While the string y always appears as a substring of $S(x, y)$, we carefully said that a *representation* of x appears. In most programming languages a string x and the string $r(x)$ that represents x in that language can be different. For example, if x contains nonprintable characters such as the linefeed, most programming languages represent the nonprintable character by some special sequence of printable characters. Another case might arise if the string x contains a quote character, for then it is usually required to represent the quote character by two consecutive quote characters.

Note that FIGURE 4 is a legal member of A_1, and y has become a subprogram of the main program. However if y had been the string "The quick brown fox," then the string produced by S would be FIGURE 5 which is not a program since "The quick brown fox" is not a statement or declaration in the language. When, in general, will $S(x, y)$ be in A_1?

PROPOSITION 1. *If y is a member of A_2, then the string $\sigma = S(x, y)$ is a member of A_1.*

```
function F(X: STRING)

The quick brown fox

begin
  return F(X, "hi");
end;
```

FIGURE 5

Proof. Recall σ is a member of A_1 if it is a legal function named F of one input STRING. By inspection of FIGURE 3, it is clear that σ will belong to A_1 as long as it is a legal program. Observe that the next-to-last line of σ always contains a *call* to a string valued function F of *two* string arguments. For example, in FIGURE 4 this call is given by

$$\text{F(X, "hi")}.$$

The string σ will be a legal program if it contains a local subprogram F of two parameters to match its call. If y is in A_2, σ will contain such a subprogram.

We are ready to state the Parameter Theorem. For functions f and g, $f(x) = g(x)$ means *either* f and g are both not defined for x *or* they are defined for x and have the same value.

THEOREM 1 (The Parameter Theorem). *For any $x \in X^*$ and $y \in A_2$*

$$\Psi_y^2(a, x) = \Psi_{S(x, y)}^1(a) \tag{1}$$

as functions of a.

Before considering the proof, note the theorem, loosely speaking, says: *Take a program $y \in A_2$ (i.e., which has two inputs) and consider the function (of one variable* a) *obtained by fixing the program's second input parameter to x. Then this function is identical to the function of one variable implemented by $\sigma = S(x, y)$.*

Proof (Informal). Informally consider the special case when x is "hi" and y is the program shown in FIGURE 1. Recall FIGURE 4 shows the string $\sigma = S(x, y)$ for this special case. Our theorem asserts that for any string a, FIGURE 1's program, when given input $(a, \text{"hi"})$, outputs the same value as does FIGURE 4's program with input a. Recall, from the proof of Proposition 1, the call to F in the next to last line of FIGURE 4 is a call to a local subprogram, appearing above it as string y. The program in FIGURE 4 always makes a call to its local subprogram, identical to the program in FIGURE 1, and supplies for the two input parameters its (FIGURE 4's) one input parameter along with "hi." FIGURE 4's program, hence, returns whatever value was returned by the FIGURE 1 subprogram. Hence, the two functions are equal. Since the above argument did not depend on any special properties of x and y, our proof is actually general.

3. Implementing the S Function

Consider FIGURE 6, a program in A_2 which *implements* S. We have expressed the algorithm, shown in italics, in English.

```
function F(X, Y:STRING)
   A:STRING := "function F(X:STRING)";
   B:STRING := "begin" & LF & "return F(X,";
   C:STRING:= "); "& LF & "end;";
begin
   Start with string A.
   Append string Y.
   Append string B.
   Append a representation of X.
   Append string C.
   Return this value.
end;
```

FIGURE 6

The program builds its output string in five pieces. Three of the pieces, A, B, and C are always the same, no matter what the program's input is. These strings contain sections of program text necessary to make the output appear as in FIGURE 3. The other two pieces needed are the second input parameter Y, and a string which *represents* the first input parameter, X, as a string. (Recall that a string and its representation in a language may be different.) String A contains the initial line of program text with which members of A_2 begin. String B contains the program text which will appear after the string Y but before the representation of X. This also includes a needed line feed character. Finally, string C is the substring containing some punctuation, a line feed, and the necessary ending text for a member of A_2.

4. Program Anomalies

Given Theorem 1, we may answer the question raised at the beginning of this paper. Our next theorem asserts the existence of a function Ψ_σ^1 with domain $\{\sigma\}$. Hence σ is a program in A_1 that halts exactly when its input parameter is σ (i.e., identical to its source code).

THEOREM 2. *There is a function* $\Psi_\sigma^1 \in P_1$ *whose domain is* $\{\sigma\}$.

Proof. Let b be any fixed element of X^* (say, "match"). Define the function h as follows: *For any* u, v *in* X^*, *calculate* $S(v, v)$. *If* $S(v, v)$ *equals* u *define* $h(u, v)$ *to be* b. *Otherwise let* $h(u, v)$ *be undefined*. Since we have defined h with an algorithm, h is in P_2. Hence, we may construct a program $z \in A_2$ which implements h. Let σ be the string $S(z, z)$.

We first claim that $\Psi_\sigma^1(a)$ and $h(a, z)$ are the same functions of a. For let a be any string in X^*. Since $\sigma = S(z, z)$, we may rewrite $\Psi_\sigma^1(a)$ as $\Psi_{S(z,z)}^1(a)$. Next, since z is in A_2, Theorem 1 may be used to rewrite $\Psi_{S(z,z)}^1(a)$ as $\Psi_z^2(a, z)$. Because z is the program which implements h, by definition of Ψ_z^2, we may rewrite $\Psi_z^2(a, z)$ as $h(a, z)$.

Now recall $h(a, z)$ is defined if and only if $a = S(z, z)$. That is, if and only if, $a = \sigma$. Hence, the domain of $\Psi_\sigma^1(a)$ contains only σ.

The proof of Theorem 2 shows us how to construct the program σ. FIGURE 7 shows a program which implements the function h in the proof and can therefore serve as the program z. The program has two input parameters U and V, and gets into an infinite loop unless $U = S(V, V)$. If it does return, the program returns the (arbitrarily chosen) string "match." Note our program of FIGURE 6 serves as a local subprogram to the program in FIGURE 7. Finally, to obtain σ we *apply* the function S to (z, z). That is, the string shown in FIGURE 7 is supplied to both parameters in the program shown in FIGURE 6.

```
function F(U,V:STRING)
  function S(X,Y:STRING)
     as in Figure 6
  end;
begin
  if U = S(V,V) then
    return "match"
  else
    loop forever
end;
```

FIGURE 7

It is interesting to note that in [4] Thompson gives an example of a program which always halts and which has the property that its output is identical to its source code. This program, like our program in Theorem 2, has a "self-referential" quality. However Thompson's program halts for all input strings, while our program halts only for a single input. The existence of a program similar to Thompson's can actually be shown by modifying the proof of Theorem 2 so that $h(u, v)$ is defined *always* to be $S(v, v)$, regardless of the value of u. The construction of the program and the argument that it has the required property are very similar to our proof of Theorem 2, however we omit the details.

Finally, we wish to thank the editor and referee for valuable suggestions which greatly enhanced the style and content of this paper.

Appendix: Programming Notes

We include this optional section for readers who wish to construct programs described in the present paper. For readability, literal strings in this paper were written with double quotes, e.g., as "hi." Two complications result from this notation. First, because of the problem that a quote might appear within a string, the S function must be sophisticated enough to represent the quote as two consecutive quotes. A more serious problem, from a practical view, is that the output of S must contain the substring

$$F(X, \textit{representation of } Y)$$

within a single line of text. However, the string Y (and hence its representation) may be quite long, especially when it is the program z described above. Since many compilers place limits on the length of input lines, the resulting string may not compile. An alternative is to denote strings as arrays of individual characters. In Ada we would write ('*h*', '*i*'). This notation has the advantage that the string expression can be spread over several lines. With this approach $S(x, y)$ might appear as in FIGURE 8 rather than in FIGURE 4. Using this method, we generated the program of Theorem 2 as 1330 lines, depicted in FIGURE 9. It was then tested and appeared to behave correctly.

```
function F(X:STRING) return STRING is

function F(X,Y:STRING) return STRING is
begin
  return X;
end;

begin
return F(X,(
'h',
'i'));
end;
```

FIGURE 8

```
function F(X:STRING) return STRING is

function F(U,V:STRING) return STRING is
  function S(X,Y:STRING) return STRING is
    A: constant STRING := ('f','u','n','c','t','i','o','n','',
    Many lines not shown.
'e',
'n',
'd',
';',
LF));
end;
```

FIGURE 9

REFERENCES

1. G. Bray and D. Pokrass, *Understanding Ada—A Software Engineering Approach*, John Wiley and Sons, New York, 1985.
2. M. Davis and E. Weyuker, *Computability, Complexity, and Languages*, Academic Press, New York, 1983.
3. R. Soare, *Recursively Enumerable Sets and Degrees: The Study of Computable Functions and Computably Generated Sets*, Springer-Verlag, New York, 1987.
4. K. Thompson, Reflections on trusting trust, *Communications of the ACM* 27, 8 (1984) 761-763.

Random Jottings on G. H. Hardy

...

His spelling was not immaculate, odd for anyone of his time (we could all spell all but perfectly from the age of 12). And he once queried whether when someone who wrote 'depreciate' 'didn't really mean "deprecate"—there is a word deprecate isn't there?'

Littlewood's Miscellany
(Béla Bollobas, Ed.)
Cambridge University Press, 1986

Which Real Matrices Have Real Logarithms?

JEFFREY NUNEMACHER
Ohio Wesleyan University
Delaware, OH 43015

The exponential function $\exp(x) = \sum_{n=0}^{\infty} x^n/n!$ can be defined in any domain in which the infinite series converges; thus, in particular, it makes sense in any Banach algebra. This note is concerned with characterizing the range of exp in $M_n(\mathbf{R})$, the set of all $n \times n$ real matrices, which is the simplest nontrivial real Banach algebra. Let us agree to write $\log B = A$ whenever $\exp A = B$. Then we wish to determine which real matrices possess a real logarithm.

It is elementary that for the domain $\mathbf{R}$ the range of exp is $\mathbf{R}^+ = \{y > 0\}$ and for the domain $\mathbf{C}$ the range is $\mathbf{C}\setminus\{0\}$. For the domain $M_n(\mathbf{C})$, consisting of all $n \times n$ complex matrices, it is not too difficult to show that the range is $GL_n(\mathbf{C})$, the set of all nonsingular complex matrices. That it can be no larger than $GL_n(\mathbf{C})$ follows from the Jacobi identity $\det(\exp A) = \exp(\operatorname{tr} A)$, a classical result which is also a consequence of remarks below. Standard infinite series expansions for the logarithm can be used to construct $\log B$ for certain large open sets in $GL_n(\mathbf{C})$. For example, if $\|B - I\| < 1$, where $\| \ \|$ denotes any operator norm, then a logarithm can be defined by the series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(B-I)^n,$$

or if $\left\|(B-I)(B+I)^{-1}\right\| < 1$ then by the series

$$\sum_{n=1}^{\infty} \frac{2}{2n-1}\left[(B-I)(B+I)^{-1}\right]^{2n-1}.$$

Using the Jordan form, it is possible to construct a complex logarithm for any B in $GL_n(\mathbf{C})$ in terms of a finite series. For this approach see Pullman [**6**, §3.7], and for other proofs of this result see Bellman [**1**, Chapter 11, §20–21]; Gantmacher [**2**, Chapter 8, §8]; and Henrici [**3**, Theorem 2.6h]. A good general reference for exp in a Banach algebra setting is Henrici [**3**, Chapter 2].

In analogy with the one-dimensional case, the range of exp when the domain is $M_n(\mathbf{R})$ might be conjectured to be all real matrices with all eigenvalues positive. But the example

$$\exp\begin{pmatrix} 0 & -\pi \\ \pi & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

shows this guess to be incorrect. Or, pursuing a different idea of positivity, one might conjecture the range to be all elements of $GL_n(\mathbf{R})$ with positive determinant. The Jacobi identity again shows that it can be no larger. But this guess is not correct either: it is rather easy by direct calculation to show that no real matrix A can satisfy

$$\exp A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}.$$

We prove the following characterization.

THEOREM. *An $n \times n$ real matrix B has a real logarithm if and only if B is nonsingular and for each negative eigenvalue λ and each k in $\mathbf{N}$ the number of $k \times k$ Jordan λ-blocks occurring in the Jordan form of B is even.*

The result is surely known to experts in matrix analysis or Lie theory, but we have been unable to find it anywhere in the common literature. There is a statement as an exercise in Hirsch and Smale [4, p. 132], but the condition is not formulated there with complete precision.

In the study of Lie groups the exponential mapping from the Lie algebra of a Lie group into the group is a fundamental tool (see, for example, Howe [5]). The case when the Lie group is $GL_n(\mathbf{R})$, its Lie algebra is $M_n(\mathbf{R})$, and the exponential mapping is the one defined above gives the simplest nontrivial real example. In addition, every continuous 1-parameter subgroup of $GL_n(\mathbf{R})$ can be written as $\exp(tA)$ for some A in $M_n(\mathbf{R})$ (see Howe [5], Theorem 10), so the range of exp consists of all those matrices which can belong to some continuous 1-parameter subgroup. In different language, the mapping $\exp(tA)$ defines a flow on $\mathbf{R}^n$ which gives the solution of the constant coefficient linear differential equation $x'(t) = Ax(t)$. Thus by finding the range of exp, we are identifying all matrices which can belong to such a flow.

To appreciate better the condition given in the theorem, consider the case of a real nonsingular 2×2 matrix B with eigenvalues λ_1 and λ_2. Then B will have a real logarithm in any of the following cases:

a) if λ_1 and λ_2 $(=\bar{\lambda}_1)$ are both nonreal;

b) if λ_1 and λ_2 are both positive;

c) if λ_1 and λ_2 are equal and negative and B has Jordan form $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}$.

B will *not* have a real logarithm in the excluded cases, namely:

d) if λ_1 is positive and λ_2 is negative;

e) if λ_1 and λ_2 are both negative but not equal;

f) if λ_1 and λ_2 are equal and negative and B has Jordan form $\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$.

A consequence of the theorem is the following perhaps surprising result concerning the topological nature of the range of exp.

COROLLARY. *If $n > 1$ the range of exp is neither open nor closed in $GL_n(\mathbf{R})$.*

Proof of the corollary. We construct appropriate convergent sequences for the 2×2 case which can then be embedded into $GL_n(\mathbf{R})$ by adding 1's down the diagonal. The sequence $\begin{pmatrix} -1 & 1/n \\ 0 & -1 \end{pmatrix}$ is not in the range by f) above, yet its limit $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ is by c). Thus the complement of the range is not closed, so the range is not open. The sequence of companion matrices $\begin{pmatrix} 0 & 1 \\ -1-1/n^2 & 2 \end{pmatrix}$ is in the range by a) above, yet its limit $\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$ is not by f). Thus, the range is also not closed in $GL_n(\mathbf{R})$.

Before giving the proof of the theorem, we recall a few facts from matrix theory. Two matrices M and N in $M_n(\mathbf{C})$ are called similar if there exists a matrix P in $GL_n(\mathbf{C})$ so that $M = PNP^{-1}$. If M and N are real and similar, then P can be chosen to be real also.

It is a standard result from linear algebra that any M in $M_n(\mathbf{C})$ is similar to some matrix J in Jordan form. J is said to be in Jordan form if it is a block diagonal matrix with certain blocks $J_k(\lambda)$, known as $k \times k$ Jordan λ-blocks, on the diagonal, where

$$J_k(\lambda)=\begin{pmatrix}\lambda & 1 & & & & 0\\ & \lambda & 1 & & & \\ & & \lambda & 1 & & \\ & & & \ddots & & \\ & & & & \lambda & 1\\ 0 & & & & & \lambda\end{pmatrix}.$$

A matrix J in Jordan form to which M is similar is unique up to the ordering of the blocks $J_k(\lambda)$, so it is referred to as "the" Jordan form of M. The λ's which occur in J are the eigenvalues of M.

If M is in $M_n(\mathbf{R})$, then its Jordan form J is special in the sense that for every nonreal eigenvalue λ each $J_k(\lambda)$ in J has a twin $J_k(\bar{\lambda})$ which is also present in J. Conversely, if a matrix in Jordan form has such a pairing of Jordan λ-blocks for all nonreal λ, then there exists a real matrix to which it is similar. This characterization of real matrices in terms of their Jordan form follows from the less familiar real Jordan form for real matrices, for which a good reference is Shilov [7, §6.6].

Let J denote the Jordan form of a matrix M in $M_n(\mathbf{C})$, so that $M=PJP^{-1}$ for some P in $GL_n(\mathbf{C})$. It follows easily from explicit calculation that $\exp M=P(\exp J)P^{-1}$ and $\exp J$ is block diagonal with the blocks on the diagonal equal to $\exp J_k(\lambda)$ for those $J_k(\lambda)$ which occur in J. Further explicit calculation shows that

$$\exp J_k(\lambda)=e^{\lambda}\begin{pmatrix}1 & 1 & 1/2! & 1/3! & \cdots & 1/(k-1)!\\ & 1 & 1 & 1/2! & \cdots & 1/(k-2)!\\ & & 1 & 1 & \cdots & 1/(k-3)!\\ & & & \ddots & \ddots & \\ & & & & & \cdots\\ 0 & & & & 1 & 1\\ & & & & & 1\end{pmatrix}.$$

It follows that the Jordan form of $\exp J_k(\lambda)$ is $J_k(\exp\lambda)$. The reason is that the number of Jordan λ-blocks in the Jordan form of any matrix M is equal to $\dim\ker(M-\lambda I)$, and it is clear in our case that $\dim\ker(\exp J_k(\lambda)-(\exp\lambda)I)=1$ since the columns, except for the first, are obviously linearly independent. Thus there can be only one Jordan $\exp(\lambda)$-block in the Jordan form of $\exp J_k(\lambda)$, and this forces the Jordan form to be $J_k(\exp\lambda)$.

Putting these observations together, we conclude that the Jordan form of $\exp M$ has the same Jordan structure as that of M with each $k\times k$ Jordan λ-block in M corresponding to a $k\times k$ Jordan $(\exp\lambda)$-block in $\exp M$. The Jacobi identity mentioned above is an easy consequence of this statement.

Proof of the theorem. The main idea is to match up the blocks in the Jordan forms of A and B and to make use of the special nature of the Jordan form of a real matrix. As noted above, the Jacobi identity implies that for any A in $M_n(\mathbf{R})$ the matrix $\exp A$ must be nonsingular.

Let us consider a nonsingular real matrix B. Suppose that λ is a negative real eigenvalue of B and that $J_k(\lambda)$ is some Jordan λ-block in the Jordan form of B. If $B=\exp A$ for some A in $M_n(\mathbf{R})$, then $J_k(\lambda)$ must be the Jordan form of $\exp J_k(\log\lambda)$ for some block $J_k(\log\lambda)$ in the Jordan form of A, where $\log\lambda$ is necessarily nonreal. Since A is a real matrix, the Jordan $(\log\lambda)$-block $J_k(\log\lambda)$ has a twin $J_k(\overline{\log\lambda})$ in the Jordan form of A. But $\exp J_k(\overline{\log\lambda})$ yields another copy of the block $J_k(\lambda)$ in the Jordan form of B. Thus we conclude that a necessary condition for B to have a real logarithm is that each Jordan λ-block $J_k(\lambda)$ for λ negative in the Jordan form of B must have a twin; i.e., for each k in $\mathbf{N}$ and each negative eigenvalue λ of B the number of $k \times k$ Jordan λ-blocks must be even.

This necessary condition (together with nonsingularity) is also sufficient. Let us suppose that it is satisfied. We construct a block diagonal matrix A' from the Jordan form of B as follows. For each nonreal eigenvalue λ of B; we choose any complex value for $\log \lambda$ and put into the diagonal of A' Jordan blocks $J_k(\log \lambda)$ and $J_k(\overline{\log \lambda})$ to correspond in size with any Jordan blocks $J_k(\lambda)$ and $J_k(\bar{\lambda})$ which occur in the Jordan form of B. Since we are assuming that the $k \times k$ Jordan λ-blocks come in pairs for negative eigenvalues λ, we can do the same for each such negative λ. Finally, for each positive real eigenvalue λ we put into A' a $k \times k$ Jordan ($\log \lambda$)-block using the real choice of $\log \lambda$ to correspond to each $J_k(\lambda)$ present in the Jordan form of B.

The A' so constructed is the Jordan form of some real matrix A'' since it has the required pairing of $k \times k$ Jordan λ-blocks, and $B' = \exp A'$ has the same Jordan form as B by construction. Thus $\exp A''$ is similar to B, and since they are both real matrices, there is a nonsingular real matrix P so that $B = P(\exp A'')P^{-1}$. Then $A = PA''P^{-1}$ is the required real logarithm for B. This completes the proof.

Acknowledgement. The author wishes to thank the referees for their perceptive and helpful comments.

REFERENCES

1. Richard Bellman, *An Introduction to Matrix Analysis*, 2nd ed., McGraw Hill, New York, 1970.
2. F. R. Gantmacher, *Applications of the Theory of Matrices*, Vol. 1, Chelsea, New York, 1959.
3. Peter Henrici, *Applied and Computational Complex Analysis*, Vol. 1, Wiley-Interscience, 1974.
4. Morris Hirsch and Steven Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, New York, 1974.
5. Roger Howe, Very basic Lie theory, *Amer. Math. Monthly* 90 (1983), 600–623.
6. N. J. Pullman, *Matrix Theory and its Applications*, Marcel Dekker, New York, 1976.
7. Georgi Shilov, *Linear Algebra*, Dover, New York, 1977.

On a hot summer day in the 1970's I met one of the teaching assistants; he was barefoot and carrying his shoes. He asked if I thought it would be all right if he taught his class barefoot. I said that I couldn't see any objection. Later I recalled that during the 1940's a young instructor (now a well-known mathematician) had been rebuked by his department chairman for going to class with his shoes not shined.

—Ralph P. Boas

PROBLEMS

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

BRUCE HANSON, *associate editor*
St. Olaf College

Proposals

To be considered for publication, solutions should be received by September 1, 1989.

1317. *Proposed by Allen J. Schwenk, Western Michigan University, Kalamazoo.*

Let $X_n = \{1, 2, \dots, n\}$ and let S be any nonempty collection of subsets of X_n. Define S' to be the collection of all subsets of X_n that are subsets of an odd number of elements of S. For example, when $X_3 = \{1, 2, 3\}$ and $S = \{\{3\}, \{1,2\}, \{1,3\}\}$, we find $S' = \{\varnothing, \{2\}, \{1,2\}, \{1,3\}\}$. In Problem **1267**, Ronald Graham asked us to prove that $(S')' = S$.

Observe that occasionally $S' = S$, for example when $S = \{\varnothing, \{2\}, \{1,2\}, \{1,3\}, \{2,3\}\}$. Characterize and count the collections of subsets for which $S' = S$.

1318. *Proposed by David Doster, Choate Rosemary Hall, Wallingford, Connecticut.*

Define a sequence (p_n) recursively by the formula

$$p_n = \left(\frac{6n-1}{4n}\right)p_{n-1} - \left(\frac{2n-1}{4n}\right)p_{n-2}, \qquad p_0 = 1, \qquad p_1 = 5/4.$$

Evaluate $\lim_{n\to\infty} p_n$.

1319. *Proposed by Jeffrey Shallit, Dartmouth College, Hanover, New Hampshire.*

Let $\sigma(N)$ denote the sum of the divisors of N. Show that $\sigma(N)$ is a power of 2 if and only if N is the product of distinct Mersenne primes. (A prime p is Mersenne if $p = 2^a - 1$ where a is prime.)

 We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals should be accompanied by solutions, if at all possible, and by any other information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution. An asterisk (*) next to a problem number indicates that neither the proposer nor the editors supplied a solution.

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

Solutions and new proposals should be mailed in duplicate to Loren C. Larson, Department of Mathematics, St. Olaf College, Northfield, MN 55057.

1320. *Proposed by Václav Konečný, Ferris State University, Michigan.*

Let $C(I)$ be a circle with center I, the incenter of triangle ABC. Let D, E, F be the points of intersection of $C(I)$ with the lines from I that are perpendicular to sides BC, CA, AB respectively. Show that AD, BE, CF are concurrent.

1321. *Proposed by Miha'ly Bencze, Braşov, Romania.*

Let $x, y, z \geqslant \sqrt{3}$. Prove that

$$\Gamma(x+y+z+1) \geqslant (x+1)(y+1)(z+1)\Gamma(x+1)\Gamma(y+1)\Gamma(z+1).$$

Quickies

Answers to the Quickies are on page 143.

Q745. *Proposed by L. Van Hamme, Vrije Universiteit Brussel, Belgium.*

Prove that

$$1+\sum_{n=1}^{\infty}\frac{1}{n+1}\left(\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots(2n)}\right)^2=\frac{4}{\pi}.$$

Q746. *Proposed by Murray S. Klamkin, University of Alberta, Canada.*

If V_i denotes the $(n-1)$-dimensional volume of the face F_i opposite the vertex A_i of the n-dimensional simplex S_n: $A_0A_1 \cdots A_n$, show that

$$V_0+V_1+\cdots+V_n > 2V_i$$

for all i.

Q747. *Proposed by Rodica Simion, George Washington University, Washington, D.C., and Frank Schmidt, Bryn Mawr College, Pennsylvania.*

Given points A and B, and $n-1$ points $P_1, P_2, \ldots, P_{n-1}$, in the plane, can a further point P_n always be found so that

a. A will lie on the least squares line, L, "through" $P_1, P_2, \ldots, P_n$?

b. Both A and B will lie on L?

Solutions

A Maximum Problem **April 1988**

1292. *Proposed by Murray S. Klamkin, University of Alberta, Canada.*

Determine the maximum value of

$$x_1^2x_2+x_2^2x_3+\cdots+x_n^2x_1$$

given that $x_1+x_2+\cdots+x_n=1$, $x_1, x_2, \ldots, x_n \geqslant 0$ and $n \geqslant 3$.

I. *Solution by the 1988 USA Olympiad Math Team.*

The maximum is $4/27$, achieved when the x_k comprise a cyclic permutation of $(2/3, 1/3, 0, \ldots, 0)$.

Suppose $n>3$. Choose j so that $x_{j+1} \geqslant x_j$ (where $x_{n+1}=x_1$). The modified sequence $(\ldots, x_{j-2}, x_{j-1}+x_j, x_{j+1}, \ldots)$, with only $n-1$ terms, gives at least as high a value for the objective function as does the original sequence, since

$$x_{j-2}^2(x_{j-1}+x_j)+(x_{j-1}+x_j)^2 x_{j+1} \geqslant x_{j-2}^2 x_{j-1}+(x_{j-1}^2+x_j^2)x_{j+1}$$
$$\geqslant x_{j-2}^2 x_{j-1}+x_{j-1}^2 x_j+x_j^2 x_{j+1}.$$

Thus it suffices to prove the bound of $4/27$ only for $n=3$.

So, suppose $n=3$. Without loss of generality, cycle the three indices so that x_2 takes on the intermediate value. Then $(x_2-x_1)(x_2-x_3) \leqslant 0 \leqslant x_1x_2$, so

$$x_1^2x_2+(x_2-x_1)(x_2-x_3)x_3 \leqslant x_1^2x_2+x_1x_2x_3.$$

These terms can be arranged to give

$$x_1^2x_2+x_2^2x_3+x_3^2x_1 \leqslant (x_1+x_3)^2x_2=(1-x_2)^2x_2$$
$$\leqslant 4\left(\frac{\frac{1}{2}(1-x_2)+\frac{1}{2}(1-x_2)+x_2}{3}\right)^3=4/27,$$

the last by the AM-GM inequality whenever $0 \leqslant x_2 \leqslant 1$.

II. *Solution by Eugene Lee, Boeing Commercial Airplane Company, Seattle.*

We are to find the maximum of $f_n(x)=x_1^2x_2+x_2^2x_3+\cdots+x_n^2x_1$, $x \in R^n$, $n \geqslant 3$, over the $(n-1)$-simplex σ_{n-1}: $x_1+\cdots+x_n=1$, $x_i \geqslant 0$. Note that at a boundary point $(2/3, 1/3, 0, \ldots, 0)$, $f_n=4/27$. We will prove that this is indeed the maximum.

First we show that the maximum cannot occur in the interior of σ_{n-1}. Indeed, by Lagrange multipliers, critical points of $f_n(x)$ restricted to the hyperplane $x_1+\cdots+x_n=1$ must satisfy

$$2x_1x_2+x_n^2=\lambda$$
$$\vdots$$
$$2x_nx_1+x_{n-1}^2=\lambda.$$

Multiplying the ith equation by x_i and summing, we have $3f_n=\lambda$. Also, summing all the equations, we obtain

$$n\lambda=x_1^2+\cdots+x_n^2+2(x_1x_2+\cdots+x_nx_1) \leqslant (x_1+\cdots+x_n)^2=1,$$

so that at any such critical point,

$$f_n=\frac{\lambda}{3} \leqslant \frac{1}{3n} < \frac{4}{3^3},$$

showing that the maximum of f_n over σ_{n-1} must occur on the boundary.

For $n=3$, a typical face of σ_2 is σ_1 obtained by setting $x_3=0$, on which $f_3(x)=f_3(x_1,x_2,0)=x_1^2(1-x_1)$, the maximum of which, by calculus, is $4/27$ occurring at $x_1=2/3$. Thus, by symmetry, $\max_{\sigma_2} f_3=4/27$. Generally, a typical face of σ_{n-1} is σ_{n-2} obtained by setting $x_n=0$, on which

$$f_n(x_1,\ldots,x_{n-1},0)=x_1^2x_2+\cdots+x_{n-2}^2x_{n-1}$$
$$\leqslant x_1^2x_2+\cdots+x_{n-2}^2x_{n-1}+x_{n-1}^2x_1$$
$$=f_{n-1}(x_1,\ldots,x_{n-1}).$$

Hence

$$\max_{\sigma_{n-1}} f_n \leqslant \max_{\sigma_{n-2}} f_{n-1} \leqslant \cdots \leqslant \max_{\sigma_2} f_3 = 4/27,$$

and the remark in the first paragraph shows that all inequalities are equalities.

Also solved by Mark S. Ashbaugh and Rafael D. Benguria (Chile), Duane M. Broline, Robert C. Carson, Chico Problem Group, François Dubeau and Jean Savoie (Canada), Thomas E. Elsner, Raoul-Fr. Gloden (Italy), H.-J. Seiffert (West Germany), and the proposer. There were five incorrect solutions, caused, in most cases, by not considering the boundary.

gcd-preserving Transformations **April 1988**

1293. *Proposed by William P. Wardlaw, U.S. Naval Academy, Annapolis.*

Let **A** be an $n \times n$ matrix with integer entries. Prove that gcd(**AX**) = gcd(**X**) for every integer column n-tuple **X** if and only if det(**A**) = ± 1. (gcd(**X**) denotes the greatest common divisor of the entries of **X**.)

Solution by David Callan, University of Bridgeport, Connecticut.

Clearly gcd(**X**)|gcd(**AX**). Suppose det(**A**) = ± 1. Then $\mathbf{A}^{-1} = \pm\,\mathrm{adj}(\mathbf{A})$ has integer entries (adj(**A**) is the adjoint of **A**). So gcd(**AX**)|gcd($\mathbf{A}^{-1}\mathbf{AX}$) = gcd(**X**), giving gcd(**AX**) = gcd(**X**).

Conversely, suppose gcd(**AX**) = gcd(**X**) for each integer n-tuple **X**. Then **A** is invertible, for if not, there is an integer n-tuple $\mathbf{X} \neq 0$ such that $\mathbf{AX} = \mathbf{0}$, and thus gcd(**AX**) $\neq$ gcd(**X**). Let $m = \det(\mathbf{A})$, so $\mathbf{A}\,\mathrm{adj}(\mathbf{A}) = \mathrm{diag}(m, m, \dots, m)$. Applying the hypothesis to the columns of adj(**A**) we find that m divides each entry of adj(**A**). It follows that $m^{n+1} | \det(\mathbf{A}\,\mathrm{adj}(\mathbf{A})) = m^n$, and thus $m = \pm 1$.

Also solved by Duane M. Broline, Chico Problem Group, Jim Delany, Fred Dodd, Hugh M. Edgar, Nathan S. Feldman and Daniel Scully (students), Richard A. Gibbs, Hans Kappus (Switzerland), David W. Koster, David E. Manes, Ricardo Perez Marco (Spain), Uri Peled, M. Riazi-Kermani, Harvey Schmidt, Jr., Daniel B. Shapiro, and the proposer.

Shapiro generalized the result in the following way. Let us say that an $m \times n$ matrix with integer entries **A** has the "gcd property" if gcd(**AX**) = gcd(**X**) for every integer column n-tuple **X**.

THEOREM. *Let* **A** *be an* $m \times n$ *matrix over* **Z**. *The following are equivalent.*

(i) **A** *has the* gcd *property.*
(ii) *There exists an* $n \times m$ *matrix* **P** *over* **Z** *with* $\mathbf{PA} = \mathbf{I}_n$.
(iii) *The* gcd *of all the* $n \times n$ *minors of* **A** *is* 1.
(iv) **A** *can be enlarged to an* $m \times m$ invertible *matrix over* **Z**.

An Antiderivative Problem **April 1988**

1294. *Proposed by William Dunham, Hanover College, Indiana.*

Express $\int \sqrt{\tan x}\, dx$ in closed form.

Solution by Leon Gerber, St. John's University, New York.

Let $u = \sqrt{\tan x}$. Then

$$\int \sqrt{\tan x}\, dx = \int u \frac{2u\,du}{u^4+1} = \int \frac{u^2-1}{u^4+1}\,du + \int \frac{u^2+1}{u^4+1}\,du$$

$$= \int \frac{1-u^{-2}}{(u+u^{-1})^2-2}\,du + \int \frac{1+u^{-2}}{(u-u^{-1})^2+2}\,du$$

$$= \frac{1}{2\sqrt{2}} \log \frac{u + u^{-1} - \sqrt{2}}{u + u^{-1} + \sqrt{2}} + \frac{1}{\sqrt{2}} \arctan \frac{u - u^{-1}}{\sqrt{2}} + C$$

$$= \frac{1}{2\sqrt{2}} \log \frac{\tan x + \sqrt{2 \tan x} - 1}{\tan x + \sqrt{2 \tan x} + 1} + \frac{1}{\sqrt{2}} \arctan \frac{\tan x - 1}{\sqrt{2 \tan x}} + C.$$

Also solved by Charles Ashbacher, Mark S. Ashbaugh, Seung-Jin Bang (Korea), John T. Baskin, Frank P. Battles, J. C. Binz (Switzerland), Jerry E. Bolick, George Boros, Lawrence S. Braden, Duane M. Broline, Barry Brunson, David Callan, Robert C. Carson, Timothy Chow (student), Keith Conrad (student), Robert Dahlin, Patrick M. Donnelly, François Dubeau (Canada), David Earnshaw, G. A. Edgar, Kurt Eisemann, Ervin Eltze, Russell Euler, Michael J. Fitzgerald (student), Arthur H. Foss, Raoul-Fr. Gloden (Italy), M. R. Gopal, Ralph P. Grimaldi, Jerrold W. Grossman, William H. Gustafson, C. T. Haskell, Jonathan Higa (student), Rebecca E. Hill, Brian Hogan, Joe Howard, Padmini T. Joshi, Geoffrey A. Kandall, Hans Kappus (Switzerland), M. Riazi-Kermani, Murray S. Klamkin (Canada), Benjamin G. Klein, Václav Konečný, Joe Konhauser, Y. H. Harris Kwong, Kee-Wai Lau (Hong Kong), S. A. Lilge, Kim McInturff, David E. Manes, Charles W. Mitchell, Jr., Steve Monson, Y. Mythili (India), Roger B. Nelsen, William A. Newcomb, Stephen Noltie, Cornel G. Ormsby, David E. Penney, Bob Prielipp, Kennard Reed, Jr., Adam Riese, P. K. Sahoo, Robert P. Schilleman, Volkhard Schindler (East Germany), Harry Sedinger, H.-J. Seiffert (West Germany), Bruce Shawyer (Canada), Florentin Smarandache (Romania), Adam Stinchcomke, Randall J. Swift, Jan Söderkvist (Sweden), Richard Tebbs, Claude C. Thompson, Nora S. Thornber, Jack V. Wales, Jr., Gary L. Walls, Edward T. H. Wang (Canada), Leon Weintrob (Israel), Joseph Wiener, A. Yanusuica, I. J. Zucker (England), and the proposer.

Readers found this problem in Leithold's *The Calculus with Analytic Geometry* (5th edition), Harper and Row, 1986, Exercise 59, p. 705, and in Edwards and Penny, *Calculus and Analytic Geometry* (2nd edition), Englewood Cliffs, 1986, Problem 119, p. 494. It also appears in G. N. Berman's *A Problem Book in Mathematical Analysis*, Mir Publishers, Moscow, 1977, Exercise 2131, p. 143. The answer, in the latter, is given in the form

$$\frac{1}{\sqrt{2}}\left(\log|\sin x + \cos x - \sqrt{\sin 2x}| + \arcsin(\sin x - \cos x)\right) + C.$$

One technique used for this integral (substitution followed by partial fractions) works as well for any rational power of $\tan x$. The integral was carried out successfully by the software program MACSYMA but not by MAPLE.

Spiraling Squares **April 1988**

1295. *Proposed by Edward Kitchen, Santa Monica, California.*

Let R be a given rectangle. Construct a square outwards on the length of R; construct another square outwards on the length of the resulting rectangle. Continue this process anticlockwise indefinitely.

a. Prove that the centers of the spiraling squares lie on two perpendicular lines.

b. As the process continues, show that the ratio of the sides of the rectangles approaches the golden mean.

Solution by Duane M. Broline, Eastern Illinois University, Charleston.

Let R be a $t_0 \times t_1$ rectangle oriented so that its sides are parallel to the coordinate axes and with the "length", t_1, in the direction of the positive x-axis. The points of the plane will be identified with complex numbers in the usual manner so that the lower left corner P_0 is 0 and the upper right-hand corner is $P_1 = t_1 + t_0 i$. A $t_1 \times t_1$ rectangle is added to the top of R so that its upper left corner is $P_2 = P_1 - t_1 + t_1 i$. Let $t_2 = t_0 + t_1$. A $t_2 \times t_2$ rectangle is added to the right of this rectangle so that its lower left corner is $P_3 = P_2 - t_2 - t_2 i$. Let $t_3 = t_2 + t_1$. A $t_3 \times t_3$ rectangle is added to the bottom so that its lower right corner is $P_4 = P_3 + t_3 - t_3 i$. Now let $t_4 = t_3 + t_2$ and add a $t_4 \times t_4$ rectangle to the right so that its upper left right corner is $P_5 = P_4 + t_4 + t_4 i$. This completes one cycle of the construction. A simple computation shows that

$$P_{n+1} = P_n + i^n(1+i)t_n, \quad \text{when } n = 1,2,3,4.$$

Thus, by induction, the opposite corners of the spiraling squares are given by the recurrence relations

$$P_{n+1} = P_n + i^n(1+i)t_n,$$

$$t_{n+1} = t_n + t_{n-1}, \; n > 0,$$

where

$$P_1 = t_1 + t_0 i.$$

For $n \geqslant 1$, let M_n be the midpoint of the square whose opposite vertices are P_n and P_{n+1}. Using the recurrence relations, we find that

$$M_{n+2} - M_n = \frac{i^{n+1}}{2}(1+3i)t_{n+1}.$$

If n is odd, $\frac{i^{n+1}}{2}t_{n+1}$ is a real number and the line joining M_{n+2} and M_n is parallel to the line joining 0 to $1+3i$. Hence the points M_n, n odd, are collinear and the slope of this line is 3. A similar argument shows that the points M_n, n even, are all on a line of slope $-1/3$. This verifies part a.

The nth rectangle in this process has dimensions t_n by t_{n-1}. If

$$r_n = \frac{t_n}{t_{n-1}}, \qquad n \geqslant 1,$$

then

$$r_{n+1} = \frac{t_{n+1}}{t_n} = \frac{t_n + t_{n-1}}{t_n} = 1 + \frac{1}{r_n}, \qquad n \geqslant 1.$$

In particular, $1 < r_n < 2$, $n \geqslant 2$. If ϕ is the golden mean, then $\phi = 1 + 1/\phi$, and $1 < \phi < 2$. Furthermore,

$$|r_{n+1} - \phi| = \left|\frac{1}{r_n} - \frac{1}{\phi}\right| = \frac{|\phi - r_n|}{\phi r_n} < \frac{1}{\phi}|r_n - \phi|.$$

These facts imply that the sequence $\{r_n\}$ approaches ϕ.

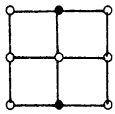
Also solved by Michael Bertrand, J. C. Binz (Switzerland), Chico Problem Group, Jim Delany, François Dubeau (Canada), Ragnar Dybvik (Norway), P. Teresa Farnum, Herta T. Freitag, John F. Goehl, Jr., Francis M. Henderson, J. Heuver (Canada), Hans Kappus (Switzerland), M. Riazi-Kermani, Václav Konečný, L. Kuipers (Switzerland), Eugene Lee, Middle Tennessee State University Problem Solving Group, Stephen Noltie, Volkhard Schindler (East Germany), Harry Sedinger, Seshadri Sivakumar (Canada), J. M. Stark, Jan Söderkvist (Sweden), and the proposer.

Other work related to spiral patterns in the plane is contained in the following paper: Herbert L. Holden, Fibonacci tiles, *The Fibonacci Quarterly*, vol. 13, no. 1, February 1975, pp. 45–49.

Lattice Packing **April 1988**

1296. *Proposed by Howard Cary Morris, Xerox Computer Services, Los Angeles.*

Consider an $n \times n$ square lattice with points colored either black or white. A *square path* is a closed path in the shape of a square with edges parallel to the edges of the lattice. Let $M(n)$ be the minimum number of black points needed for an $n \times n$ square lattice so that every square path has at least one black point on it. (For example, $M(3) = 2$; see FIGURE.)



a. Show that $\lim_{n \to \infty} M(n)/n^2$ exists.
b. Evaluate $\lim_{n \to \infty} M(n)/n^2$.
*c. Find a formula for $M(n)$.

Solution (to parts a and b) by Daniel E. Erickson, Jet Propulsion Laboratory, Pasadena, California.

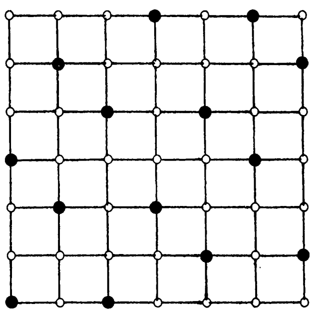
Let P be a black point in the lattice, and suppose S is a 2 by 2 square path that passes through P. Assign P a "credit" of $1/k$ if S passes through exactly k black squares. Let $T(P)$ be the sum of all the credits assigned to P as S varies over all 2 by 2 squares that pass through P.

Note that the sum of $T(P)$ as P varies over all black points in the square array is $(n-1)^2$ (each of the $(n-1)^2$ 2 by 2 arrays contributes 1 to the total).

It is clear that $T(P) \leqslant 1$ if P is a corner square, and $T(P) \leqslant 2$ if P is on the outer edge. Suppose that P is a square in the interior of the lattice. It lies on exactly four 2 by 2 square paths, and there must be at least one black square on the 3 by 3 square path surrounding P. Thus, for such a $P, T(P) \leqslant 7/2$.

Thus, in all cases, $(7/2)M(n) \geqslant (n-1)^2$, or equivalently, $M(n) \geqslant 2(n-1)^2/7$.

On the other hand, the pattern shown below for a 7 by 7 lattice (2/7 of the points are black) can be extended to an arbitrary n by n lattice (tile an m by m lattice, $m = 7\lceil n/7 \rceil$, with copies of the lattice below, and then remove $m - n$ rows and columns from the top and from the right, respectively).



These observations show that $\lim_{n \to \infty} M(n)/n^2 = 2/7$.

Parts a and b were also solved by Duane M. Broline and the proposer. No solutions were received for part c.

Erickson expanded the problem to higher dimensions (square paths replaced by surfaces of cubes), and showed that the general limit is $2/(2^{r+1} - 1)$, where r is the dimension of the space.

Answers

Solutions to the Quickies on p. 137.

A745. Put

$$s_n = 4n\left(\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots(2n)}\right)^2.$$

Since

$$s_{n+1} - s_n = \frac{1}{n+1}\left(\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots(2n)}\right)^2$$

the partial sum of the series is s_n. Now Wallis' formula says that $\lim_{n\to\infty} s_n = 4/\pi$.

A746. Orthogonally project the faces $F_0, F_1, \ldots, F_{i-1}, F_{i+1}, \ldots, F_n$ onto the space of F_i. If A'_i, the projection of A_i, lies in the convex hull H_i of $A_0, A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n$, then

$$V_0' + V_1' + \cdots + V_{i-1}' + V_{i+1}' + \cdots + V_n' = V_i,$$

where V_j' denotes the $(n-1)$-dimensional volume of the orthogonal projection of face F_j. If A'_i lies outside the convex hull H_i, then

$$V_0' + V_1' + \cdots + V_{i-1}' + V_{i+1}' + \cdots + V_n' > V_i.$$

Finally, since $V_j > V_j'$, we are done. Note that there is equality only in the case of a degenerate simplex in which case one vertex lies in the convex hull of the remaining vertices.

Note that for $n = 2$, we have $a_1 + a_2 + a_3 > 2a_i$ for the sides of a triangle, and for $n = 3$, we have $F_1 + F_2 + F_3 + F_4 > 2F_i$ for the areas of the faces of a tetrahedron.

A747. a. Yes. It is well known that L passes through the centroid $C = (\bar{x}, \bar{y})$ of $P_1, P_2, \ldots, P_n$. Hence, the coordinates (x_n, y_n) of P_n can be (uniquely) chosen so that $A = C$.

b. No. Suppose that $P_1, P_2, \ldots, P_{n-1}$ lie in a disk of radius r. Then L must pass within r of the disk, else the line L' through P_n and the center of the disk gives a better fit for $P_1, P_2, \ldots, P_n$.

Comment

Problem **Q720** proved that if **A** is a nonsingular $n \times n$ matrix then adj(adj(**A**)) = (det**A**)$^{n-2}$**A**. The Comment (February 1988, p. 59), proves this equality is also true when **A** is singular. But the printed proof tacitly assumes the matrices are real or complex. What if the matrix is over some other field (e.g., over a finite field)?

Daniel B. Shapiro, The Ohio State University, shows how this can be done. The standard method of proof is to use a "generic" matrix. Let F be any field and let $x_{i,j}$ be a system of n^2 independent indeterminates over F. Define $\mathbf{X} = (x_{i,j})$ an $n \times n$ matrix over the larger field $F(\{x_{i,j}\})$. One can check that **X** is non-singular (in fact, it is a nice exercise to prove that det(**X**) is an irreducible polynomial in $F[x_{1,1}, \ldots, x_{n,n}]$). Applying the result of Q720 we have adj(adj(**X**) = (det **X**)$^{n-2}$**X**. All the entries here are polynomials in $x_{1,1}, x_{1,2}, \ldots, x_{n,n}$ so we can evaluate the polynomial identity at any value $x_{i,j} = a_{i,j} \in F$. Consequently, adj(adj(**A**)) = (det**A**)$^{n-2}$**A** for every $n \times n$ matrix **A** over F.

This "generic" technique has a number of similar applications in linear algebra. Two related references, where "extension by continuity" is valid are (i) W. Wardlaw, A transfer device for matrix theorems, this MAGAZINE 59 (1986), pp. 30–33, and (ii) W. Watkins, Polynomial identities for matrices, *Amer. Math. Monthly* 82 (1975), pp. 364–368.

NEWS AND LETTERS

SEVENTEENTH ANNUAL U.S.A. MATHEMATICAL OLYMPIAD

1. By a *pure repeating decimal* (in base 10) we mean a decimal $0.\overline{a_1 \cdots a_k}$ which repeats in blocks of k digits beginning at the decimal point. An example is $.243243243\ldots = \frac{9}{37}$. By a *mixed repeating decimal* we mean a decimal $0.b_1 \cdots b_m \overline{a_1 \cdots a_k}$ which eventually repeats, but which cannot be reduced to a pure repeating decimal. An example is $.011363636\cdots = \frac{1}{88}$.

Prove that if a mixed repeating decimal is written as a fraction $\frac{p}{q}$ in lowest terms, then the denominator q is divisible by 2 or 5 or both.

2. The cubic equation $x^3 + ax^2 + bx + c = 0$ has three real roots. Show that $a^2 - 3b \geq 0$, and that $\sqrt{a^2 - 3b}$ is less than or equal to the difference between the largest and smallest roots.

3. A function $f(S)$ assigns to each nine-element subset S of the set $\{1, 2, 3, \ldots, 20\}$ a whole number from 1 to 20. Prove that, regardless of how the function f is chosen, there will be a ten-element subset $T \subset \{1, 2, 3, \ldots, 20\}$ such that $f(T - \{k\}) \neq k$ for all $k \in T$.

4. Let I be the incenter of triangle ABC, and let A', B', and C' be the circumcenters of triangles IBC, ICA, and IAB, respectively. Prove that the circumcircles of triangles ABC and $A'B'C'$ are concentric.

5. A polynomial product of the form

$$(1-z)^{b_1}(1-z^2)^{b_2}(1-z^3)^{b_3}(1-z^4)^{b_4}(1-z^5)^{b_5}\cdots(1-z^{32})^{b_{32}},$$

where the b_k are positive integers, has the surprising property that if we multiply it out and discard all terms involving z to a power larger than 32, what is left is just $1 - 2z$. Determine, with proof, b_{32}. (The answer can be written as the difference of two powers of 2.)

29th INTERNATIONAL MATHEMATICAL OLYMPIAD

1. Consider two coplanar circles of radii R and r $(R > r)$ with the same center. Let P be a fixed point on the smaller circle and B a variable point on the larger circle. The line BP meets the larger circle again at C. The perpendicular l to BP at P meets the smaller circle again at A. (If l is tangent to the circle at P then $A = P$.)

(i) Find the set of values of $BC^2 + CA^2 + AB^2$.

(ii) Find the locus of the midpoint of AB.

2. Let n be a positive integer and let $A_1, A_2, \ldots, A_{2n+1}$ be subsets of a set B. Suppose that

(a) each A_i has exactly $2n$ elements,
(b) each $A_i \cap A_j$ $(1 \leq i < j \leq 2n + 1)$ contains exactly one element, and
(c) every element of B belongs to at least two of the A_i.

For which values of n can one assign to every element of B one of the numbers 0 and 1 in such a way that each A_i has 0 assigned to exactly n of its elements?

3. A function f is defined on the positive integers by

$f(1) = 1,\ f(3) = 3,$
$f(2n) = f(n),$
$f(4n + 1) = 2f(2n + 1) - f(n),$
$f(4n + 3) = 3f(2n + 1) - 2f(n),$

for all positive integers n.

Determine the number of positive integers n, less than or equal to 1988, for which $f(n) = n$.

4. Show that the set of real numbers x which satisfy the inequality

$$\sum_{k=1}^{70} \frac{k}{x - k} \geq \frac{5}{4}$$

is a union of disjoint intervals, the sum of whose lengths is 1988.

5. ABC is a triangle right-angled at A, and D is the foot of the altitude from A. The straight line joining the incenters of the triangles ABD, ACD intersects the sides AB, AC at the points K, L respectively. S and T denote the areas of the triangles ABC and AKL respectively. Show that $S \geq 2T$.

6. Let a and b be positive integers such that $ab + 1$ divides $a^2 + b^2$. Show that $\frac{a^2 + b^2}{ab + 1}$ is the square of an integer.

49th PUTNAM COMPETITION: WINNERS AND SOLUTIONS

Teams from 270 schools competed in the 1988 William Lowell Putnam Mathematical Competition. The top five winning teams, in descending rank, are:

Harvard University
David J. Moews, Bjorn M. Poonen, Constantin S. Teleman

Princeton University
Daniel J. Bernstein, David J. Grabiner, Matthew D. Mullin

Rice University
Hubert L. Bray, Thomas M. Hyer, John W. McIntosh

University of Waterloo
Frank M. D'Ippolito, Colin M. Springer, Minh-Tue Vo

California Institute of Technology
William P. Cross, Robert G. Southworth, Glenn P. Tesler

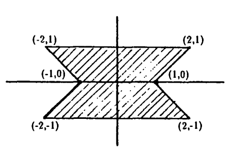
The five highest ranking individuals, named Putnam Fellows, are:

David J. Grabiner	Princeton University
Jeremy A. Kahn	Harvard University
David J. Moews	Harvard University
Bjorn M. Poonen	Harvard University
Ravi D. Vakil	University of Toronto

Solutions to the 1988 *Putnam problems were prepared for publication in this* Magazine *by Loren Larson, St. Olaf College.*

A-1. Let R be the region consisting of the points (x,y) of the cartesian plane satisfying both $|x| - |y| \leq 1$ and $|y| \leq 1$. Sketch the region R and find its area.

Sol. The part of R in the first quadrant is bounded by $x = 0$, $y = 0$, $x - y = 1$, and $y = 1$. This part is a trapezoid with vertices (0,0), (1,0), (2,1), and (0,1) and area 3/2. Since $(\pm x, \pm y)$ is in R when (x,y) is in R, the parts of R in the other quadrants are obtained using symmetry about both axes, and consequently, the area of R is 6.



A-2. A not uncommon calculus mistake is to believe that the product rule for derivatives says that $(fg)' = f'g'$. If $f(x) = e^{x^2}$, determine, with proof, whether there exists an open interval (a,b) and a non-zero function g defined on (a,b) such that this wrong product rule is true for x in (a,b).

Sol. The function defined by $g(x) = e^x\sqrt{2x - 1}$ has the property desired for $1/2 < a < x < b$ and $g(x) = e^x\sqrt{1 - 2x}$ has the property for $a < x < b < 1/2$.

To derive that result, set $y = g(x)$ and set up the differential equation $(e^{x^2}y)' = (e^{x^2})'y'$. This reduces to $2xy + y' = 2xy'$, which separates into the form $\frac{2x\,dx}{2x - 1} = \frac{dy}{y}$. It follows that

$$\log|y| = x + \tfrac{1}{2}\log|1 - 2x| + C,$$

where C is an arbitrary constant. If $1/2 < a < x < b$, this has the form

$$g(x) = Ae^x\sqrt{2x - 1}$$

where A is an arbitrary positive real number. If $a < x < b < 1/2$, it has the form $g(x) = Ae^x\sqrt{1 - 2x}$.

A-3. Determine, with proof, the set of real numbers x for which $\sum_{n=1}^{\infty}\left(\frac{1}{n}\csc\frac{1}{n} - 1\right)^x$ converges.

Sol. Let $a_n = \frac{1}{n}\csc\frac{1}{n} - 1$. Then

$$a_n = \frac{1}{n\sin\frac{1}{n}} - 1 = \frac{1}{n\left(\frac{1}{n} - \frac{1}{6n^3} + \frac{1}{120n^5} - \cdots\right)} - 1$$

$$= \frac{1}{1 - \frac{1}{6n^2} + \frac{1}{120n^4} + \cdots} - 1$$

$$= 1 + \frac{1}{6n^2} + \frac{1}{n^2}g(n) - 1 = \frac{1}{n^2}\left(\frac{1}{6} + g(n)\right),$$

where $g(n) \to 0$ as $n \to \infty$. Hence, there exist positive real numbers c, d, and N such that

$$c\,\frac{1}{n^2} \leq a_n \leq d\,\frac{1}{n^2}, \text{ for } n > N.$$

Using the comparison and the p-test, one finds that $\Sigma\, a_n^x$ converges for $x > 1/2$ and diverges for $0 < x \leq 1/2$. But it is easy to see that the series also diverges for $x \leq 0$. Hence the answer is $\{x\colon x > 1/2\}$.

A-4. (a) If every point of the plane is painted one of three colors, do there necessarily exist two points of the same color exactly one inch apart.

(b) What if "three" is replaced by "nine"? Justify your answers.

Sol. (a) The answer is yes. For the proof let A be an arbitrary point in the plane and let ABC be an artibrary equilateral triangle with side length 1 (where the units are inches, of course) that has A as one of its vertices. If any two of A, B, and C have the same color, the construction is finished. If not, let A' be the point obtained by reflecting A through the line BC. If A' has the same color as either B or C, the construction is finished. If not, then A and A' have the same color. Note that the distance between A and A' is $\sqrt{3}$, and that, in fact, any two points at distance $\sqrt{3}$ from one another can be obtained by making one of them a vertex of an equilateral triangle of side length 1 and then reflecting it through the side opposite it.

The result so obtained implies that, for any initial point A, either the reflected equilateral triangle argument finishes the desired construction for some B and C, or else that every point at distance $\sqrt{3}$ from A has the same color as A. The set of points such as A', at distance $\sqrt{3}$ from A, is a circle of radius $\sqrt{3}$; any chord of length 1, of that circle, yields a pair of points of the same color exactly one inch apart.

(b) The answer is no. For the proof, pave the plane with squares whose common side length is chosen so that the diagonals are nearly 1 but not equal to 1; the diagonal length 0.9 will do. If that length is used, then the side length of each square is $0.9/\sqrt{2}$, which is somewhat greater than 0.63. Color one square with color #1, color the eight squares adjacent to it with colors #2 - #9, and then repeat, throughout the plane, the coloring scheme of the large square (consisting of nine small squares) so obtained. (For present purposes it doesn't matter what consistent convention is followed for the boundaries of the squares; one possibility is to let the bottom and left boundaries of each square have the same color as the interior.)

The result is a nine-coloring of the plane in which no two points of the same color are exactly one inch apart. Indeed, for any point at all, the points of the same color are either within 0.9 inches from it or else farther than $2 \times .63 = 1.26$ inches.

A-5. Prove that there exists a <u>unique</u> function f from the set R^+ of positive real numbers to R^+ such that

$f(f(x)) = 6x - f(x)$ and $f(x) > 0$ for all $x > 0$.

Sol. For arbitrary $x > 0$, let a_0, a_1, a_2, ... be defined by $a_0 = x$ and $a_{n+1} = f(a_n)$. Then $a_{n+2} + a_{n+1} - 6a_n = 0$ for $n = 0, 1, 2, \ldots$. The characteristic roots of this difference equation are -3 and 2. Hence $a_n = (-3)^n c + 2^n k$ for some constants c and k. As $a_{n+1} = f(a_n) > 0$ for all n, we must have $c = 0$ and so $f(x) = 2x$. This unique f satisfies the conditions since it gives $f(f(x)) = f(2x) = 4x = 6x - f(x)$ and $2x > 0$ for $x > 0$.

A-6. If a linear transformation A on an n-dimensional vector space has $n + 1$ eigenvectors such that any n of them are linearly independent, does it follow that A is a scalar multiple of the identity? Prove your answer.

Sol. Yes, A must be a scalar multiple of the identity. Suppose that

$$\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n, \mathrm{x}_{n+1}$$

are eigenvectors of A such that any n of them are linearly independent, with corresponding eigenvalues

$$\lambda_1, \lambda_2, \ldots, \lambda_n, \lambda_{n+1}.$$

Since any n of the x's are linearly independent, it follows that any n of them span the whole space. Each x, therefore, is a linear combination of the others; in particular

$$\mathrm{x}_{n+1} = \sum_{i=1}^{n} \alpha_i x_i.$$

Note that none of the α's can be 0. Reason: if $\alpha_i = 0$, then the complement of x_i within $\{\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n, \mathrm{x}_{n+1}\}$ is linearly dependent.

Apply A to the equation to get

$$\mathrm{A}\,\mathrm{x}_{n+1} = \sum_{i=1}^{n} \alpha_i \lambda_i \mathrm{x}_i .$$

Since, however,

$$\mathrm{A}\,\mathrm{x}_{n+1} = \lambda_{n+1}\mathrm{x}_{n+1} = \sum_{i=1}^{n} \alpha_i \lambda_{n+1} \mathrm{x}_i,$$

it follows (from the linear independence of $\{\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n\}$) that

$$\alpha_i \lambda_i = \alpha_i \lambda_{n+1}$$

for each $i = 1, 2, \ldots, n$. Since $\alpha_i \neq 0$, this implies that $\lambda_{n+1} = \lambda_i$ for $i = 1, 2, \ldots, n$.

B-1. A <u>composite</u> (positive integer) is a product ab with a and b not necessarily distinct integers in $\{2, 3, 4 \ldots\}$. Show that every composite is expressible as $xy + xz + yz + 1$, with x, y, and z positive integers.

Sol. Letting $z = 1$, we have

$$xy + xz + yz + 1 = xy + x + y + 1 = (x+1)(y+1)$$

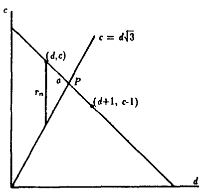
and this gives us all the composite positive integers when x and y range over all the positive integers.

B-2. Prove or disprove: If x and y are real numbers with $y \geq 0$ and $y(y + 1) \leq (x + 1)^2$, then $y(y - 1) \leq x^2$.

Sol. If $x \leq y - 1/2$ then $y^2 - y \leq x^2 + 2x + 1 - 2y \leq x^2$. If $x \geq y - 1/2 \geq 0$, then $x^2 \geq y^2 - y + 1/4 > y^2 - y$. If $0 \leq y < 1/2$, then $y(y - 1) \leq 0 \leq x^2$.

B-3. For every n in the set $Z^+ = \{1, 2, \dots\}$ of positive integers, let r_n be the minimum value of $|c - d\sqrt{3}|$ for all nonnegative integers c and d with $c + d = n$. Find, with proof, the smallest positive real number g with $r_n \leq g$ for all n in Z^+.

Sol. The number $|c - d\sqrt{3}|$ is the vertical distance from the point (d,c) to the line L_1: $c = d\sqrt{3}$. Let (d,c) and $(d + 1, c - 1)$ be lattice points on line L_2: $c + d = n$, lying on opposite sides of L_1. Let a be the minimum of the distances from (d,c) and $(d + 1, c - 1)$ to the point P where L_1 and L_2 intersect (see figure).



Then $r_n = a/\sqrt{2} + a\sqrt{3}/\sqrt{2}$, and $a \leq \sqrt{2}/2$, so $r_n \leq (1 + \sqrt{3})/2 = g$. However, because the numbers of the form $\frac{c - 1/2}{d + 1/2}$ are dense in the positive reals and so get arbitrarily close to $\sqrt{3}$, the midpoint $\left(d + \frac{1}{2}, c - \frac{1}{2}\right)$ can be arbitrarily close to P and hence a can be arbitrarily close to $\sqrt{2}/2$. Thus r_n can be arbitrarily close to g.

B-4. Prove that if $\sum_{n=1}^{\infty} a_n$ is a convergent series of positive real numbers, then so is

$$\sum_{n=1}^{\infty} (a_n)^{n/(n+1)}.$$

Sol. Let $S = \{n : a_n^{n/n+1} < 2a_n\}$. If $n \notin S$, $a_n^{n/n+1} \geq 2a_n$, or equivalently $1/2 \geq a_n^{1-(n/n+1)} = a_n^{1/n+1}$, which is the same as $1/2^n \geq a_n^{n/n+1}$. It follows that

$$\sum_{n=1}^{\infty} a_n^{n/n+1} \leq \sum_{n \in S} a_n^{n/n+1} + \sum_{n \notin S} 1/2^n < \infty.$$

B-5. For positive integers n, let $\mathbf{M}_n$ be the $2n + 1$ by $2n + 1$ skew-symmetric matrix for which each entry in the first n subdiagonals below the main diagonal is 1 and each of the remaining entries below the main diagonal is -1. Find, with proof, the rank of $\mathbf{M}_n$. (According to one definition the rank of a matrix is the largest k such that there is a $k \times k$ submatrix with non-zero determinant.) One may note that

$$\mathbf{M}_1 = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix} \text{ and}$$

$$\mathbf{M}_2 = \begin{pmatrix} 0 & -1 & -1 & 1 & 1 \\ 1 & 0 & -1 & -1 & 1 \\ 1 & 1 & 0 & -1 & -1 \\ -1 & 1 & 1 & 0 & -1 \\ -1 & -1 & 1 & 1 & 0 \end{pmatrix}.$$

Sol. In $\mathbf{M}_n$, the sum of the entries of any row (or any column) is 0. Hence, $\det \mathbf{M}_n = 0$ and the rank of $\mathbf{M}_n$ is less than $2n + 1$. Let S be the matrix obtained from $\mathbf{M}_n$ by deleting the i-th row and the i-th column. The main diagonal of S consists of 0's and every other entry is ± 1. Hence, det S is a sum of terms, each equal to ± 1, with a term for each of the d_{2n} derangements on $2n$ objects. Known formulas for d_m include

$$d_m = m!\left(\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^m \frac{1}{m!}\right)$$

(1) $d_m = (-1)^m(1 - m + m(m-1) - m(m-1)(m-2) + \cdots)$

(2) $d_m = md_{m-1} + (-1)^m$

(3) $d_m = (m - 1)(d_{m-1} + d_{m-2})$.

Using any one of (1), (2), or (3), one sees easily that d_m is odd when m is even; i.e., d_{2n} is odd. This tells us that det S is an odd integer and hence is nonzero. Thus the rank of $\mathbf{M}_n$ is $2n$.

B-6. Prove that there exist an infinite number of ordered pairs (a, b) of integers such that for every positive integer t the number $at + b$ is a triangular number if and only if t is a triangular number.

Sol. The infinite set $((2m + 1)^2, t_m)$ for $m = 1, 2, \dots$ has the desired properties.

On the one hand, if $n(n + 1)/2$ is a triangular number, then

$$(2m + 1)^2\frac{n(n + 1)}{2} + \frac{m(m + 1)}{2} = \frac{(2mn + m + n)(2mn + m + n + 1)}{2}.$$

On the other hand, suppose that x is an integer such that

$$(2m+1)^2x + \frac{m(m+1)}{2} = \frac{n(n+1)}{2}$$

for some positive integer n. Then

$$x = \frac{(n^2 - m^2) + (n - m)}{2(2m+1)^2} = \frac{(n-m)(n+m+1)}{2(2m+1)^2}.$$

The factors in the numerator, $n - m$ and $n + m + 1$, differ by $2m + 1$, and therefore, since x is an integer, $2m + 1$ divides both $n - m$ and $n + m + 1$. After this division is carried out, the two factors, $(n - m)/(2m + 1)$ and $(n + m + 1)/(2m + 1)$, differ by 1, and it follows that x is a triangular number.

LETTERS TO THE EDITOR

Editor:

I like the "Proof without Words: The Characteristic Polynomials of AB and BA Are Equal", that Sidney Kung provides on page 294 of the December 1988 issue of Mathematics Magazine. I merely wish to point out that the exact same factorizations given there provide a proof of the more general result that if A and B are $m \times n$ and $n \times m$ matrices, respectively, then, for $m \leq n$, $(-\lambda)^{n-m}\det(AB - \lambda I_m) = \det(BA - \lambda I_n)$, from which it is immediate that $\lambda^{n-m}\det(\lambda I_m - AB) = \det(\lambda I_n - BA)$. This proof has the advantage of depending only upon the permutation definition and the multiplicativity of the determinant.

R. Bruce Richter
Carleton University
Ottawa, Ontario
Canada K1S 5B6

Editor:

I read the interesting article entitled "The Ubiquitous π" by D. Castellanos (this MAGAZINE 61(1988) 67-98, 148-163). I found that the two formulae (4.14) at the bottom of p. 75 and (4.15) at the top of p. 76 are not correct and must be written as

$$\pi = \lim_{n\to\infty} \frac{1}{n}\left[\frac{2\cdot4\cdot6\cdots(2n)}{1\cdot3\cdot5\cdots(2n-1)}\right]^2$$

and

$$\pi = \lim_{n\to\infty} \frac{1}{n}\left[2^{2n}\Big/\binom{2n}{n}\right]^2.$$

W.A. Bassali
Kuwait University
P.O. Box 5969
13060 Safat, Kuwait

Editor:

I want to commend the two-part article "The Ubiquitous Pi" by Dario Castellanos (this MAGAZINE 61 (1988) 67-98, 148-163). His concise and rich treatise of this most fascinating of numbers provided much reading pleasure. I would like to share a few formulae of my own discovery.

$$(31)^{1/3} + \frac{5^4}{(12^3 - 11)^2} = 3.141592654132181...$$

$$\sqrt{\log_{10}\left[(21^2\cdot29)[(111^2)(47) + 2^5)]\right]} =$$

$$3.141592653585938....$$

John M. Coker
16300 West 65th Street
Shawnee Mission, KS 66217

[The author sent along a number of additional formulae, including some involving continued fractions. Those interested should address the author of the letter. Ed.]

COMMENT

H.K. Krishnapriyan, Department of Mathematics and Computer Science, Drake University, points out that the method presented in "On the Sum of Consecutive Kth Powers", by Jeffrey Nunemacher and Robert M. Young, this MAGAZINE 60 (1987) 237-238, appears in essentially this form in Hardy and Wright, *An Introduction to the Theory of Numbers*, fifth ed., p. 90.

SUMMER STUDY IN EUROPE

Two courses: (1) "On the Shoulders of Giants: A History of British Mathematics," at Oxford University, August 1-20, 1989. (2) "Mathematicians of Göttingen," at Salzburg, Austria, June 27 to July 25, 1989. (May include travel to Göttingen and Vienna.)

For course descriptions, fees, and other information, contact Paul R. Wolfson, Department of Mathematical Sciences, West Chester University, West Chester, PA 19383. Phone: (215) 436-2452.

THIRD BOSTON WORKSHOP

Third Boston Workshop at Wellesley College, Wellesley, MA, June 15-18, 1989.

Purpose: To support undergraduate teaching: emphasis on renewal of calculus, applied linear algebra and differential equations, algorithms.

Contact: Gilbert Strang, Room 2-240, M.I.T., Cambridge, MA 02139; (617) 253-4383.

Reporting History in Mathematics...

MATHEMATICAL VISIONS

The Pursuit of Geometry in Victorian England

Joan L. Richards

Mathematical Visions is a scholarly book on the fascinating subject of the foundations of geometry and its philosophical and sociological implications. In the late nineteenth century, particularly in England, these seemingly abstract questions were debated in the framework of education, philosophy, theology, and the natural sciences. These discussions laid the foundations for the astonishing developments in modern physics and the rigid interpretation of mathematical truth. Joan Richards puts these ideas into the full historical context of an era that intrigues and influences our thinking to this day.

1988, 280 pages, $34.95
ISBN: 0-12-587445-6

RAMANUJAN REVISITED

Proceedings of the Centenary Conference

George K. Andrews, Richard A. Askey, Bruce C. Berndt, K.G. Ramanathan, and Robert A. Rankin

Srinivasa Ramanujan was India's most famous mathematician. On the hundredth anniversary of his birth, his astounding mathematical discoveries are still the source of inspiration to many mathematicians working in number theory, analysis, and combinatorics as well as some physicists. Twenty-eight eminent speakers presented expository talks at the Ramanujan Centenary Conference on a wide range of topics. The volume as a whole is “mathematics-in-the-making” and provides an excellent introduction to numerous areas of current research.

1988, 632 pages, $49.50
ISBN: 0-12-058560-X

Making History in Mathematics...

FRACTALS EVERYWHERE

Michael Barnsley

This text focuses on how fractal geometry can be used to model real objects in the physical world. Rather than considering fractal images that have been generated randomly, the approach here is to start with a natural object and find a specific fractal to fit it. The widespread applications of fractal geometry extend to biological and physiological modeling, geography, coastlines, turbulence, images, feathers, and ocean spray. The applications to computer graphics and, in particular, image compression for data transmission and reconstruction are exciting new developments.

Providing the foundation for future development, **Fractals Everywhere** is ideal as a textbook on fractal geometry and its applications. Beautifully illustrated, including 32 full-color plates, it is also well suited as a reference work.

1988, 424 pages, $39.95
ISBN: 0-12-079062-9

ACADEMIC PRESS
Harcourt Brace Jovanovich, Publishers
Book Marketing Department #04049, 1250 Sixth Avenue, San Diego, CA 92101
CALL TOLL FREE 1-800-321-5068

CONTENTS

ARTICLES

NOTES

PROBLEMS

NEWS AND LETTERS

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