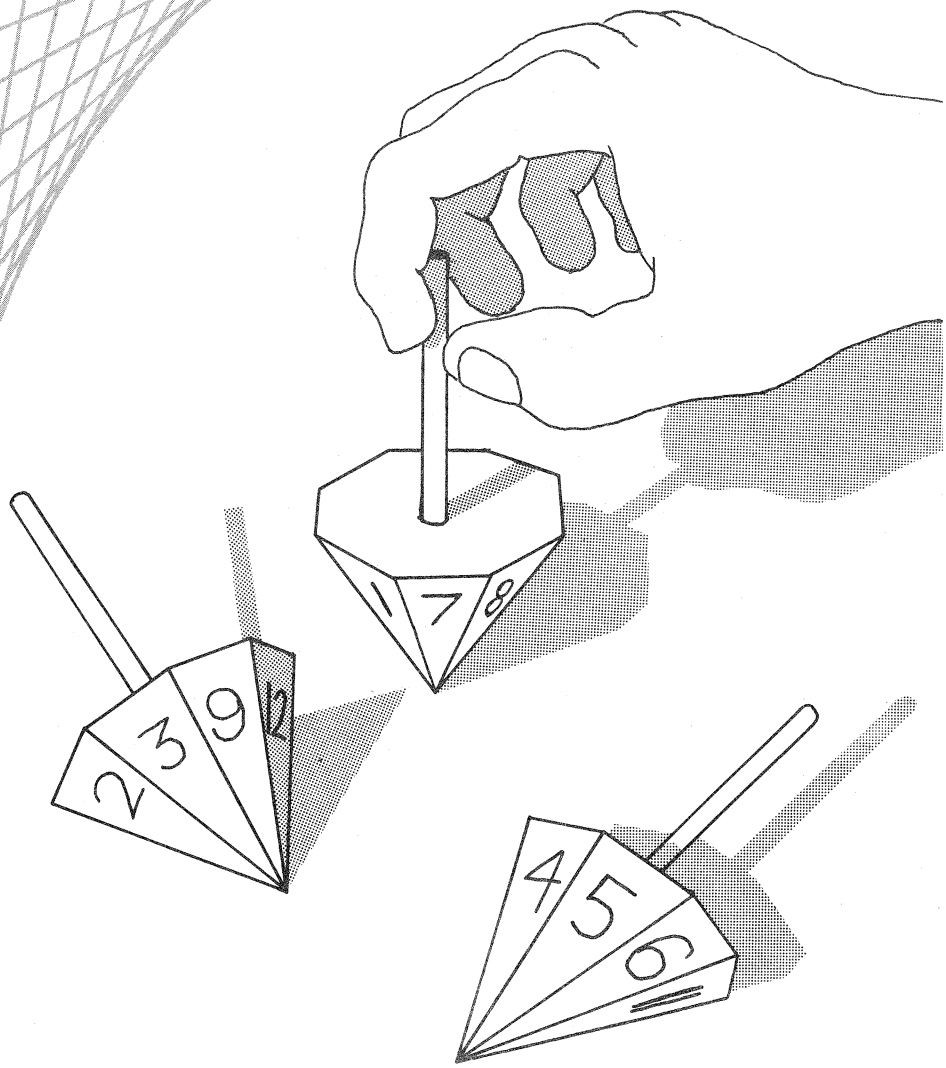
MATHEMATICS MAGAZINE



Vol. 49, No. 3
May, 1976
CODEN: MAMGA8

NONTRANSITIVITY • BIRTHDAY PROBLEM
POLYOMINOES • ULTRAMETRIC GEOMETRY

THE CARUS MATHEMATICAL MONOGRAPHS

The Monographs are a series of expository books intended to make topics in pure and applied mathematics accessible to teachers and students of mathematics and also to non-specialists and scientific workers in other fields.

These numbers are currently available:

1. *Calculus of Variations,* by G. A. Bliss.
2. *Analytic Functions of a Complex Variable,* by D. R. Curtiss.
3. *Mathematical Statistics,* by H. L. Rietz.
4. *Projective Geometry,* by J. W. Young.
6. *Fourier Series and Orthogonal Polynomials,* by Dunham Jackson.
7. *Vectors and Matrices,* by C. C. MacDuffee.
8. *Rings and Ideals,* by N. H. McCoy.
9. *The Theory of Algebraic Numbers* (Second edition), by Harry Pollard and Harold G. Diamond.
10. *The Arithmetic Theory of Quadratic Forms,* by B. W. Jones.
11. *Irrational Numbers,* by Ivan Niven.
12. *Statistical Independence in Probability, Analysis and Number Theory,* by Mark Kac.
13. *A Primer of Real Functions* (Second edition), by Ralph P. Boas, Jr.
14. *Combinatorial Mathematics,* by H. J. Ryser.
15. *Noncommutative Rings,* by I. N. Herstein.
16. *Dedekind Sums,* by Hans Rademacher and Emil Grosswald.
17. *The Schwarz Function and its Applications,* by Philip J. Davis.

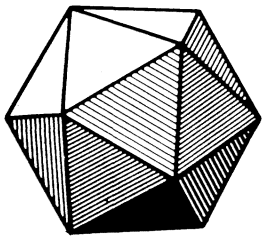
One copy of each Carus Monograph may be purchased by individual members of the Association for $5.00 each; additional copies and copies for nonmembers are priced at $10.00 each. (Orders for under $10.00 must be accompanied by payment. Prepaid orders will be delivered postage and handling free.)

Orders should be sent to:

MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, D.C. 20036

MATHEMATICS MAGAZINE

Volume 49, Number 3, May 1976



COVER: Dreidels, an old child's toy, generalize dice to permit equally probable outcomes for any number of sides. Appropriately numbered sets of dreidels can exhibit nontransitivity (see page 115) in which each dreidel is strictly dominated by some other dreidel in the set.

ARTICLES

NOTES

PROBLEMS

NEWS AND LETTERS

EDITORIAL POLICY

Mathematics Magazine is a journal of collegiate mathematics designed to enrich undergraduate study of the mathematical sciences. The *Magazine* should be an inviting, informal journal emphasizing good mathematical exposition of interest to undergraduate students. Manuscripts accepted for publication in the *Magazine* should be written in a clear and lively expository style. The *Magazine* is not a research journal, so papers written in the terse "theorem-proof-corollary-remark" style will ordinarily be unsuitable for publication. Articles printed in the *Magazine* should be of a quality and level that makes it realistic for teachers to use them to supplement their regular courses. The editors especially invite manuscripts that provide insight into applications and history of mathematics. We welcome other informal contributions, for example, brief notes, mathematical games, graphics and humor.

Editorial correspondence should be sent to: Mathematics Magazine. Department of Mathematics, St. Olaf College, Northfield, Minnesota 55057. Manuscripts should be prepared in a style consistent with the format of Mathematics Magazine. They should be typewritten and double spaced on $8\frac{1}{2}$ by 11 paper. Authors should submit the original and one copy and keep one copy as protection against possible loss. Illustrations should be carefully prepared on separate sheets of paper in black ink, the original without lettering and two copies with lettering added; the printers will insert printed letters on the illustration in the appropriate locations.

Authors planning to submit manuscripts may find it helpful to obtain the more detailed statement of guidelines available from the editorial office.

ABOUT OUR AUTHORS

Richard L. Tenney and Caxton C. Foster ("Non-transitive Dominance") are professors of Computer and Information Science at the University of Massachusetts at Amherst. Their interest in non-transitivity was sparked by Martin Gardner's 1970 article on non-transitive dice in Scientific American. Tenney earned a bachelors in mathematics at UCLA before turning to computer science in which he earned his masters and Ph.D. at Cornell. Foster began as a physicist at MIT and now holds a Ph.D. in electrical engineering from the University of Michigan. Today their joint research interests include parallel computation and microprocessors.

Roger Chalkley ("Matrices Derived from Finite Abelian Groups") who graduated from college a chemical engineer turned his hobby of mathematics into his profession. His dissertation on algebraic differential equations has led to continuing research and related interest in algebra and number theory. In particular this has resulted in several papers, of which this is the most recent, in which insights into matrices have been gained by relating matrices to unusual situations, in this case identifying certain matrices with tables of group characters.

Following the completion of his Ph.D. in 1958 Professor Chalkley worked as a mathematician for the Reactor Experimental Engineering Division of the Oak Ridge National Laboratory and then undertook postdoctoral study at the University of Zürich and the Eidgenössische Technische Hochschule where he worked with B. L. van der Waerden. He then taught two years at Knox College before coming to the University of Cincinnati where he has been since 1962.

BUSINESS INFORMATION. Mathematics Magazine is published by the Mathematical Association of America at Washington, D. C., five times a year in January, March, May, September, and November. Ordinary subscriptions are $10 per year. Members of the Mathematical Association of America or of Mu Alpha Theta may subscribe at special reduced rates. College and university mathematics departments may purchase bulk subscriptions (5 or more copies to a single address) for distribution to undergraduate students. Back issues may be purchased, when in print, for $2.00.

Subscription correspondence and notice of change of address should be sent to A. B. Willcox, Executive Director, Mathematical Association of America, Suite 310, 1225 Connecticut Avenue, N.W., Washington, D.C. 20036.

Advertising correspondence should be addressed to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, New York 14214.

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

Non-Transitive Dominance

The design of generalized dice in which A beats B, B beats C, and C beats A.

RICHARD L. TENNEY
CAXTON C. FOSTER
University of Massachusetts

If Al is taller than Bill and Bill is taller than Charlie, we may conclude that Al is taller than Charlie. This fact is abstracted mathematically by the statement that the relation "is taller than" is a transitive relation. Many other relations are also transitive: e.g., "greater than", "less than", "is isomorphic to", and "equals". Certainly, if all relations were transitive, transitivity would not be an interesting property to study. The relation "does not divide" (written $\nmid$) is not transitive, for from the facts $3 \nmid 5$ and $5 \nmid 12$, it does not follow that $3 \nmid 12$.

Intuitively one feels that relations having to do with dominance like "is better than" or "wins at chess from" or "is wiser than" should be transitive. But not all of them are, and this is surprising. Stories abound of chess masters who can beat everybody but a certain nemesis. This nemesis may be a rather second-rate player and be beaten regularly by many of the players that the chess master beats. Thus we have a case where A (the master) beats B and B beats C (the nemesis), but C beats A.

This example is to a certain extent unsatisfying because the reasons for it are unclear, or at least the problem may not be mathematical. Perhaps the fault lies in some imprecision in the definitions. So let us consider a better defined situation, one involving objects and events that can be described in detail, and one which involves some interesting mathematics.

We will consider a game between two players involving three dice, colored red, white, and blue for purposes of identification. Each player chooses a die and rolls it, and the one who rolls the higher number wins. The dice have been specially made for the game, and each face has an integer between one and nine on it; opposite faces of each die are identical; and the dice are fair in the sense that each side is equally likely. Surprisingly, this game, which sounds perfectly fair, can be rigged in such a way that the player who chooses the first die will lose an average of five out of nine games.

NONTRANSITIVE DICE: The three specially numbered dice

blue:	**2, 6, 7, 2, 6, 7**
white:	**3, 4, 8, 3, 4, 8**
red:	**1, 5, 9, 1, 5, 9**

have the property that white wins when matched against blue, red wins against white and blue wins against red. The various (equally probable) outcomes of these match-ups are enumerated in the accompanying table; asterisks denote those rolls for which the dominant die of the pair actually wins. Non-transitivity makes possible a hustler's honeymoon: by letting his opponent choose the first die, he can insure (long-term) victory by appropriate selection of the second die.

blue	white	white	red	red	blue
2	3*	3	1	1	2*
2	4*	3	5*	1	6*
2	8*	3	9*	1	7*
6	3	4	1	5	2
6	4	4	5*	5	6*
6	8*	4	9*	5	7*
7	3	8	1	9	2
7	4	8	5	9	6
7	8*	8	9*	9	7
5 wins for white		5 wins for red		5 wins for blue	

TABLE 1

Let us see how this is possible. Suppose the three dice have the following distinct numbers on their faces (each repeated twice, recall): red: 1, 5, 9; white: 3, 4, 8; blue: 2, 6, 7. If the first player chooses blue, the second chooses white; if the first chooses white, the second chooses red; and if the first chooses red, the second chooses blue. TABLE 1 gives the possible outcomes of rolling two dice in each case; asterisks mark those for which the second player wins. The "dominant" die wins five out of the nine possible rolls, and thus whichever die the first player chooses, there remains a "better" die for the second player to choose.

In two recent columns [**1, 2**] Martin Gardner has presented a number of other non-transitive games and situations. In particular, he discusses what is known as the voters' paradox (or cyclic majority problem) in which the voters when presented with pair-wise choices prefer A to B, B to C, and C to A. This problem has been studied by several people [**3, 4**], and Usiskin shows that as the number of candidates increases without bound there is a limit to the proportion of voters that prefer A to B, B to C, etc. (assuming that all such preferences are equal). This limit is 3/4 of the voters preferring the stronger candidate and 1/4 the weaker. For small numbers N of candidates, Usiskin derives values for P, the maximum preference proportion: for $N = 3$, $P = .618$; for $N = 4$, $P = 2/3$; for $N = 10$, $P = .732$.

In his column [**1**] Gardner notes that Efron was the first to point out the connection between the cyclic majority paradox and the non-transitive dominance relation among dice, with dice playing the role of candidates and the odds in favor of the dominant die being equivalent to the preference for the stronger candidate of each pair. In this paper we will do two things: first we will show that as the number of sides per die increases there is also a bound on the odds in favor of the dominant die and that this bound tends toward 3/4 as the number of sides goes to infinity. Second, we will present an algorithm for finding sets of dice with d dice of s sides each.

The odds

Standard dice are made of cubes with 6 sides. Other regular solids have 4, 8, 12 and 20 sides. For other integers n, one might be able to construct various irregular solids with n sides each equally likely to "come up". Rather simpler is the resort to an old child's toy called a dreidel. Historically a dreidel consists of a top with four flat faces each bearing one of four letters of the Hebrew alphabet. This top is spun and when it settles, one of the four faces is uppermost. It is easy to conceive of an n-sided dreidel. This requires that for n odd one either reads the face touching the floor, or better, replaces the n-gon by a $2n$-gon and duplicates each face, retaining thus n different faces, each appearing exactly twice.

To shorten the discussion which follows we present a few definitions and conventions: We will adopt a model of d dreidels, each with s sides. The dreidels are numbered $1, 2, \ldots, d$, and when

considered in pairs dreidel i dominates dreidel $i+1$ where arithmetic is done cyclically so that $d-1$ dominates d and d dominates 1, etc. Each of the total collection of $d \cdot s$ faces will contain a unique integer drawn from the set $\{1, \ldots, d \cdot s\}$. The number on the jth face of the ith dreidel will be symbolized by $A_{i,j}$. We agree to order the faces on each dreidel in monotonically increasing order so that $A_{i,j} < A_{i,j+1}$ for all i and all j. Sometimes we will refer to the ith dreidel as A_i.

We denote by x_i the number of different rolls of dreidels i and $i+1$ such that i wins from $i+1$ (i.e., x_i is the number of pairs $\langle j, k \rangle$ such that $A_{i,j} > A_{i+1,k}$). Since there are s^2 possible rolls of two s-sided dreidels, the probability that i wins is x_i/s^2. We denote by $\alpha_{i,j}$ the number of faces of dreidel $i+1$ that the jth face of dreidel i is larger than. It is the number of ways that i can still win from $i+1$ given that the ith dreidel came up $A_{i,j}$. Thus $x_i = \alpha_{i,1} + \alpha_{i,2} + \cdots + \alpha_{i,s}$. Further we know that if side j of dreidel i is larger than $\alpha_{i,j}$ faces of dreidel $i+1$ then side $j+1$ of dreidel i must be larger than at least $\alpha_{i,j}$ faces of dreidel $i+1$ ($\alpha_{i,j+1} \geqq \alpha_{i,j}$) since $A_{i,j+1} > A_{i,j}$.

Now we are ready to find the bounds on x_i as a function of the number of sides, s. Consider any column j in the array $A_{i,j}$. For some adjacent pair $\langle i, i+1 \rangle$ it must be the case that $A_{i,j} < A_{i+1,j}$ (because $<$ is transitive and each of $\{A_{1,j}, \ldots, A_{d,j}\}$ is distinct). Let us refer to dreidels A_i and A_{i+1} as B and C to simplify subscripts. For this pair, the maximum possible advantage of B over C would be achieved if all the sides $B_{j+1}, \ldots, B_s$ were greater than all the sides of C. This would allow B to win over C in $s(s-j)$ of the possible spins. Furthermore, to maximize the chance of B winning over C, $B_1, \ldots, B_j$ should be chosen large enough to win against $C_1, \ldots, C_{j-1}$ even though they must lose to $C_j, \ldots, C_s$; in short, they can account for $j(j-1)$ winning spins. Adding these together, we see that the advantage x of B over C is bounded by

$$x \leqq s(s-j) + j(j-1) = s^2 - sj + j^2 - j$$

and this bound must hold for each column j. We now ask for which column j does this analysis give the smallest x? Simple calculus and some integer arithmetic shows that the value $j = \lceil s/2 \rceil$ is the most constraining (where $\lceil z \rceil$ is the smallest integer $\geqq z$). This gives a bound on x of $x \leqq \lfloor (3s^2 - 2s)/4 \rfloor$ (where $\lfloor z \rfloor$ is the largest integer $\leqq z$). (The reader is warned that there are two cases to consider: one for even s and another for odd s; the formula holds for both cases.)

s	2	3	4	5	6	10	∞
p	.500	.555	.625	.640	.666	.700	.750

Maximum probability p of winning the dreidel game for various numbers of sides s. As shown in the accompanying article, p is derived from s by the formula $p = \lfloor (3s^2 - 2s)/4 \rfloor / s^2$, where $\lfloor z \rfloor$ is the largest integer less than or equal to z.

TABLE 2

TABLE 2 displays the probability that dreidel i will win from dreidel $i+1$. This probability, p, is found by dividing x by s^2; it approaches 3/4 as s approaches infinity, which is the same limit found by Usiskin as d approaches infinity. We will show by construction that this bound on x can be approached when d is large enough so this is a tight bound. For the general set of d dreidels with s sides the smaller of the two bounds (ours and Usiskin's) applies.

Construction of Dreidels

We now show how to construct all sets of d dreidels with s sides for which dreidel i will win x out of the s^2 possible rolls with dreidel $i+1$. The construction consists of three phases: first, generate certain appropriate partitions of x; second, build advantage tables from the partitions; and third, attempt to construct sets of dreidels corresponding to advantage tables.

The easiest way to understand the construction is to work backwards from a set of dreidels. Consider the set of three four-sided dreidels with faces

$$A_1 = 1 \quad 7 \quad 8 \quad 10$$
$$A_2 = 4 \quad 5 \quad 6 \quad 11$$
$$A_3 = 2 \quad 3 \quad 9 \quad 12$$

For this set of dreidels, A_i wins 9 of the 16 equally likely rolls with A_{i+1}. For example, face one of A_1 dominates no face of A_2, but the other faces of A_1 each dominate three faces of A_2. Similarly, the first three faces of A_2 dominate only two faces of A_3, and the fourth face of A_2 dominates three faces of A_3. This information can be succinctly summarized in what we call an advantage table. For this example, the advantage table is:

$$0 \quad 3 \quad 3 \quad 3$$
$$2 \quad 2 \quad 2 \quad 3$$
$$1 \quad 1 \quad 3 \quad 4$$

Since A_i was constructed to win 9 of the 16 equally likely rolls with A_{i+1}, each row must add to 9. Since each entry in the table represents the number of faces of a dreidel dominated by a certain face, no number in the table may be larger than 4, the number of sides on a dreidel in this example. Also, since we list the sides of dreidels in increasing order, the entries in the rows of an advantage table must be in non-decreasing order. Finally, note that in column j the smallest entry must be less than j. To see that this is true, consider the numbers on the faces of the set of dreidels (the $A_{i,j}$'s) as a matrix. The smallest number in column j of this matrix cannot dominate more than $j-1$ faces of its successor dreidel, for if it did, then the number for face j of the successor dreidel would have to be smaller than this number which was assumed the smallest in column j; a contradiction. (Note that the choice of which dreidel to label A_1 is arbitrary; but once that choice is made, the other names (A_2 and A_3) are given by the dominance relations that obtain.)

For the case of d dreidels with s sides and an advantage of x out of s^2 between adjacent dreidels, the rows of an advantage table consist of partitions of x into s integral parts arranged from left-to-right in non-decreasing order, where no part is less than zero or greater than s. The partitions of 9 into 4 parts subject to these constraints are:

$$P_1 = 0 \quad 1 \quad 4 \quad 4 \qquad P_3 = 0 \quad 3 \quad 3 \quad 3 \qquad P_6 = 1 \quad 2 \quad 3 \quad 3$$
$$P_2 = 0 \quad 2 \quad 3 \quad 4 \qquad P_4 = 1 \quad 1 \quad 3 \quad 4 \qquad P_7 = 2 \quad 2 \quad 2 \quad 3$$
$$P_5 = 1 \quad 2 \quad 2 \quad 4$$

(An interesting side problem, not explored here, is how many such partitions of x exist.)

One can now easily list all matrices that could be advantage tables for d dreidels selecting d (not necessarily distinct) partitions of x as rows, subject to the constraint that column j of the matrix have an entry which is less than j. In listing all such matrices, we are not interested in those whose rows are cyclic permutations of one another. Thus, if the matrix with rows corresponding to partitions P_1, P_3, P_5 were listed, the matrices P_3, P_5, P_1 and P_5, P_1, P_3 should be omitted. When this is done for the example we have been considering, twenty-two potential advantage tables are found.

Now that we have seen how to generate possible advantage tables, we wish to develop a procedure that will, if possible, assign integers to the sides of the dreidels to realize the pattern of the advantage table. If this is not possible, it should tell us so as quickly as possible. We present an algorithm that accomplishes this with reasonable dispatch.

With each row i of an advantage table we associate a counter C_i and a pointer L_i which tells us at which element of the row the counter is pointing. All counters are initialized to 0 and all pointers to 1, so that every counter points at the left most element of its row. A counter is said to be "satisfied" if the number it holds (the contents of C_i) is equal to the element of the advantage table at which it points. Let N be an integer initialized to 1. The algorithm has 6 steps:

1. Find a counter C_i such that C_i is satisfied and C_{i-1} is not satisfied. If no such counter exists, no set of dreidels exists for this advantage table, and the algorithm halts. (Remember: the rows in an advantage table are ordered cyclically, so $C_0 = C_d$.)
2. Assign the integer N to the side corresponding to the element at which this counter is pointing. (That is, if $L_i = j$, assign N to side $A_{i,j}$.)
3. Increment N by 1.
4. Increment L_i by 1. (Move counter i one position to the right.)
5. Increment C_{i-1} by 1 (increment the unsatisfied counter above C_i).
6. Repeat steps 1–5 until either all counters move off the right end of their rows ("moved off" counters are considered to be *not* satisfied) or the algorithm halts in step 1.

Note that L_i is always exactly one greater than C_{i-1} for all i, but the algorithm seems easier to understand if we introduce the pointers explicitly. Each pass through the algorithm assigns a number to one side of one of the dreidels so in $d \cdot s$ passes we are guaranteed to exit from step 6 unless we fail earlier.

Because in step 5 we increment the counter of the row preceding the row to which we assign the integer N, each counter tallies the number of sides on the succeeding dreidel that have been "taken care of" (by having integers assigned to them). Since we are assigning integers in increasing order, we cannot assign a value to face j of dreidel i until $\alpha_{i,j}$ sides of its successor have been assigned smaller integers, where $(\alpha_{i,j})$ is the advantage table. Thus it is that we wish to work with satisfied counters. But now consider two successive satisfied counters. If the lower one is processed first, it will mean that another side of the second of the pair of dreidels will be assigned. But since the upper counter is satisfied, we know that we have already assigned just exactly enough sides to the second dreidel. Consequently, we look for a satisfied counter C_i such that C_{i-1} is not satisfied.

TABLE 3 shows an application of this algorithm; asterisks are used to mark satisfied counters.

Advantage Table

$P_1 =$	0	1	4	4
$P_5 =$	1	2	2	4
$P_6 =$	1	2	3	3

Faces of Dreidels

$D_1 =$	0	3		
$D_2 =$	2	5	6	
$D_3 =$	1	4		

Counters and their Pointers at the beginning of each Pass

C_1	0*	0	0	1*	1	1	2	3
L_1	1	2	2	2	3	3	3	3
C_2	0	0	1*	1	1	2*	2*	2
L_2	1	1	1	2	2	2	3	4
C_3	0	1*	1	1	2*	2	2	2
L_3	1	1	2	2	2	3	3	3
Pass	1	2	3	4	5	6	7	8

A time history of the assignment algorithm. In pass 8 no satisfied counter is found, so the algorithm fails.

TABLE 3

Some results of this algorithm applied to sets of dreidels with up to seven sides are tabulated in TABLE 4.

The set of four six-sided dreidels (dice) presented in TABLE 4 is given by our algorithm as the only set with an advantage of 2/3, yet they look almost nothing like Efron's set presented by Gardner with the same advantage. The reason for this is that we have chosen to make each side unique. To convert our set of four six-sided dreidels to Efron's dice we employ a technique called compression. If on one particular dreidel two successive integers (N and $N+1$) appear, replace $N+1$ by N, $N+2$ by $N+1$ and so on for all integers (on all the dreidels) larger than N. Repeat until no successive integers appear on the same dreidel. This converts the set given in TABLE 4 to the following table:

Number of Sides	Number of Dreidels	Advantage	Number of Sets	First Advantage Table	Faces of First Set of Dreidels
3	3	5/9	2	0 2 3 1 1 3 1 2 2	1 5 9 3 4 8 2 6 7
4	3	10/16	0	———	———
4	3	9/16	6	0 1 4 4 0 3 3 3 1 2 2 4	1 4 10 11 2 7 8 9 3 5 6 12
4	4	10/16	5	0 2 4 4 1 3 3 3 2 2 2 4 1 1 4 4	1 9 12 13 5 8 10 11 4 6 7 16 2 3 14 15
5	4	16/25	4	0 1 5 5 5 0 4 4 4 4 2 3 3 3 5 2 2 2 5 5	1 3 15 16 17 2 11 12 13 14 6 8 9 10 20 4 5 7 18 19
5	3	15/25	2	0 0 5 5 5 3 3 3 3 3 1 2 2 5 5	1 3 11 12 13 6 7 8 9 10 2 4 5 14 15
5	3	16/25	0	———	———
6	4	24/36	1	0 0 6 6 6 6 4 4 4 4 4 4 3 3 3 3 6 6 2 2 2 6 6 6	1 2 16 17 18 19 10 11 12 13 14 15 6 7 8 9 23 24 3 4 5 20 21 22
6	3	23/36	0	———	———
7	4	33/49	0	———	———
7	5	33/49	> 16	0 0 5 7 7 7 7 3 5 5 5 5 5 5 3 4 4 4 4 7 7 3 3 3 3 7 7 7 1 2 2 7 7 7 7	1 3 20 23 24 25 26 13 16 17 18 19 21 22 9 11 12 14 15 34 35 6 7 8 10 31 32 33 2 4 5 27 28 29 30

TABLE 4

1	1	5	5	5	5
4	4	4	4	4	4
3	3	3	3	7	7
2	2	2	6	6	6

which is the same as Efron's, except that we start with 1, and he starts with 0.

The authors wish to thank the University Computing Center for a grant of computer time to carry out this study.

References

[1] M. Gardner, The paradox of the nontransitive dice, Scientific American, 223 (1970) 110–111.
[2] ———, On the paradoxical situations that arise from nontransitive relations, Scientific American, 231 (1974) 120–125.
[3] Z. Usiskin, Max-Min probabilities in the voting paradox, Ann. of Math. Stat., (1964) 857–862.
[4] J. E. Pomeranz, and R. L. Weil, The cyclical majority problem, CACM, 13, 4, (1970) 251–254.

Matrices Derived from Finite Abelian Groups

Some representation theory in reverse: information about certain matrices is found by relating them to finite groups.

Roger Chalkley
University of Cincinnati

In this paper, we investigate a special class of matrices called G-matrices which will be defined in terms of a finite abelian group G of order n. All matrices considered here have size $n \times n$ and their components belong to a field F which contains a primitive nth root of unity. The set of G-matrices with the operations of matrix addition and multiplication forms a commutative ring which is isomorphic to the ring of $n \times n$ diagonal matrices over F. To specify an isomorphism, we shall use the fundamental theorem for finite abelian groups to construct a nonsingular matrix M whose components are powers of a fixed primitive nth root of unity. We shall find that, for each G-matrix A, both $M^{-1}AM$ and MAM^{-1} are diagonal matrices; and, for each diagonal matrix D, both $M^{-1}DM$ and MDM^{-1} are G-matrices. After a general study of other matrices which have properties similar to those of M, we develop a formula to express the characteristic polynomial of any G-matrix as a product of linear factors over F.

Let $g_1, g_2, \ldots, g_n$ be an enumeration for the elements of G. If A is an $n \times n$ matrix over F given by $A = [a_{rs}]$ and if there exists a function σ from G to F which satisfies $a_{rs} = \sigma(g_r^{-1} g_s)$, for $r = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, n$, then we say that A is a ***G*-matrix**.

By the basis theorem for finite abelian groups ([**5**] or [**6**]), there exist elements $h_1, h_2, \ldots, h_q$ in G with orders $m_1, m_2, \ldots, m_q$ such that, for each g in G, there are unique integers $\alpha_1, \alpha_2, \ldots, \alpha_q$ which satisfy

$$g = h_1^{\alpha_1} h_2^{\alpha_2} \cdots h_q^{\alpha_q} \quad \text{and} \quad 0 \leqq \alpha_i < m_i \quad \text{for} \quad i = 1, 2, \ldots, q. \tag{1}$$

Since there are m_i values for α_i, we have

$$n = m_1 m_2 \cdots m_q, \tag{2}$$

where n is the order of G.

We now begin our construction of the matrix M. Let ζ be a primitive nth root of unity in F. For $i = 1, 2, \ldots, q$, let ψ_i be the map of G into F which is defined in terms of (1) by

$$\psi_i(h_1^{\alpha_1} h_2^{\alpha_2} \cdots h_q^{\alpha_q}) = \zeta_i^{\alpha_i}, \quad \text{where} \quad \zeta_i = \zeta^{(n/m_i)}.$$

As in (1), there exist integers α_{ij} such that

$$g_j = h_1^{\alpha_{1j}} h_2^{\alpha_{2j}} \cdots h_q^{\alpha_{qj}}, \quad \text{for} \quad j = 1, 2, \ldots, n, \tag{3}$$

where $G = \{g_1, g_2, \ldots, g_n\}$. We next define a map χ_j of G into F by

$$\chi_j = \psi_1^{\alpha_{1j}} \psi_2^{\alpha_{2j}} \cdots \psi_q^{\alpha_{qj}}, \quad \text{for} \quad j = 1, 2, \ldots, n. \tag{4}$$

Thus, for each g in G, $\chi_j(g) = (\psi_1(g))^{\alpha_{1j}}(\psi_2(g))^{\alpha_{2j}} \cdots (\psi_q(g))^{\alpha_{qj}}$.

PROPOSITION 1. *The maps $\chi_1, \chi_2, \ldots, \chi_n$ are homomorphisms of G into the multiplicative group F^* of nonzero elements of F and satisfy*

$$\chi_r(g_s) = \chi_s(g_r), \quad \text{for} \quad r = 1, 2, \ldots, n \quad \text{and} \quad s = 1, 2, \ldots, n, \tag{5}$$

and

$$\sum_{j=1}^{n} \chi_j(g) = \begin{cases} n & \text{if} \quad g = e, \\ 0 & \text{if} \quad g \neq e, \end{cases} \tag{6}$$

where e is the unit element of G and g is any element of G.

Proof. For each r and s, we note that

$$\chi_r(g_s) = \zeta_1^{\alpha_{1s}\alpha_{1r}} \zeta_2^{\alpha_{2s}\alpha_{2r}} \cdots \zeta_q^{\alpha_{qs}\alpha_{qr}} = \chi_s(g_r).$$

Since $\chi_1, \chi_2, \ldots, \chi_n$ are easily verified to be n distinct homomorphisms of G into F^*, (6) is a restatement of a well-known result (e.g., [**4**], pp. 467–468). Alternatively, we can deduce (6) by a direct computation based on (3) and (4). This concludes the proof.

Now we are ready to construct the nonsingular matrix M and to exhibit its inverse L. Let L and M be the $n \times n$ matrices over F whose components λ_{rs} and μ_{rs}, respectively, of row index r and column index s are $\lambda_{rs} = \chi_r(g_s^{-1})/n$ and $\mu_{rs} = \chi_r(g_s)$. (Since F contains a primitive nth root of unity, the characteristic of F does not divide n.) Also, let $\delta_{rs} = 1$ when $r = s$; and, $\delta_{rs} = 0$ when $r \neq s$.

PROPOSITION 2. *The matrix M is nonsingular and $M^{-1} = L$.*

Proof. We use (5) to obtain $\chi_r(g_s^{-1}) = (\chi_r(g_s))^{-1} = (\chi_s(g_r))^{-1} = \chi_s(g_r^{-1})$. With (6), the (r, s) component of LM equals

$$\sum_{j=1}^{n} \lambda_{rj}\mu_{js} = \frac{1}{n} \sum_{j=1}^{n} \chi_j(g_r^{-1} g_s) = \delta_{rs}.$$

Hence, we have $LM = I$ and $M^{-1} = L$. This completes the proof.

Now suppose P is an $n \times n$ permutation matrix. That is, P can be obtained from the $n \times n$ identity matrix I by a permutation of the columns of I. The transpose P^T of P is an $n \times n$ permutation matrix. If D is an $n \times n$ diagonal matrix, then $P^T D P$ is a diagonal matrix. We have $P^T I P = I$ and $P^{-1} = P^T$.

PROPOSITION 3. *The matrix N defined by $N = (1/n)M^2$ is an $n \times n$ permutation matrix such that $N^{-1} = N$. Moreover, if A is a G-matrix, then A^T is a G-matrix and $A^T = NAN$.*

Proof. For the (r, s) component ν_{rs} of N, we find

$$\nu_{rs} = \frac{1}{n} \sum_{j=1}^{n} \mu_{rj}\mu_{js} = \frac{1}{n} \sum_{j=1}^{n} \chi_j(g_r g_s).$$

By (6), we note that $\nu_{rs} = 1$ when $g_r g_s = e$ and $\nu_{rs} = 0$ when $g_r g_s \neq e$. Thus, N is a permutation matrix. With (5), we deduce $M^T = M$, $N^T = N$, and $N^{-1} = N$.

Suppose A is a G-matrix. Set $H = NAN$. Let the (r, s) components of A and H be a_{rs} and η_{rs}. Then $a_{rs} = \sigma(g_r^{-1} g_s)$, for some map σ of G into F, and

$$\eta_{rs} = \sum_{j=1}^{n} \sum_{k=1}^{n} \nu_{rj}\sigma(g_j^{-1} g_k)\nu_{ks}.$$

We have $\nu_{rj}\nu_{ks} = 0$ unless $g_r g_j = e = g_k g_s$; hence,

$$\eta_{rs} = \sigma((g_r^{-1})^{-1} g_s^{-1}) = \sigma(g_s^{-1} g_r) = a_{sr}$$

and $H = A^T$. Let τ be the map of G into F defined by $\tau(g) = \sigma(g^{-1})$ for each g in G. With $\eta_{rs} = \tau(g_r^{-1} g_s)$, H is a G-matrix. This completes the proof.

THEOREM 1. *Suppose $n \times n$ matrices A and D over F satisfy $AM = MD$. Then, A is a G-matrix if and only if D is a diagonal matrix; and, A is a diagonal matrix if and only if D is a G-matrix.*

Proof. Our initial goal is to show that A is a G-matrix if and only if D is diagonal. Let a_{rs} and d_{rs} be the (r, s) components of A and D. Suppose A is a G-matrix with $a_{rs} = \sigma(g_r^{-1} g_s)$. For $D = M^{-1}AM$, we find

$$d_{rs} = \sum_{j=1}^{n} \sum_{k=1}^{n} \lambda_{rj} a_{jk} \mu_{ks} = \frac{1}{n} \sum_{j=1}^{n} \chi_j(g_r^{-1}) \sum_{k=1}^{n} \sigma(g_j^{-1} g_k) \chi_k(g_s).$$

As g_k ranges over G, so does $g_j^{-1} g_k$; thus, with (5), we have

$$\sum_{k=1}^{n} \sigma(g_j^{-1} g_k) \chi_k(g_s) = \sum_{k=1}^{n} \sigma(g_j^{-1} g_k) \chi_s(g_k)$$

$$= \sum_{k=1}^{n} \sigma(g_j^{-1} g_k) \chi_s(g_j^{-1} g_k) \chi_s(g_j) = \chi_j(g_s) \sum_{k=1}^{n} \sigma(g_k) \chi_s(g_k).$$

We use the two preceding formulas and (6) to obtain

$$d_{rs} = \frac{1}{n} \sum_{j=1}^{n} \chi_j(g_r^{-1} g_s) \sum_{k=1}^{n} \sigma(g_k) \chi_s(g_k) = \delta_{rs} \sum_{k=1}^{n} \sigma(g_k) \chi_s(g_k). \tag{7}$$

This yields $d_{rs} = 0$ when $r \neq s$. Thus, D is a diagonal matrix.

Now suppose D is diagonal. With $d_{rs} = d_{rr} \delta_{rs}$ and $A = MDM^{-1}$, we find

$$a_{rs} = \sum_{j=1}^{n} \sum_{k=1}^{n} \mu_{rj} d_{jj} \delta_{jk} \lambda_{ks} = \sum_{k=1}^{n} \mu_{rk} d_{kk} \lambda_{ks}$$

$$= \frac{1}{n} \sum_{k=1}^{n} \chi_k(g_r) d_{kk} \chi_k(g_s^{-1}) = \frac{1}{n} \sum_{k=1}^{n} d_{kk} (\chi_k(g_r^{-1} g_s))^{-1}.$$

We define a map σ of G into F by

$$\sigma(g) = \frac{1}{n} \sum_{k=1}^{n} d_{kk} (\chi_k(g))^{-1}, \quad \text{for each } g \text{ in } G.$$

Thus, with $a_{rs} = \sigma(g_r^{-1} g_s)$, A is a G-matrix.

To verify the second part of the theorem, we deduce from Proposition 3 that D is a G-matrix if and only if NDN is a G-matrix. But, from the preceding paragraphs, NDN is a G-matrix if and only if the matrix $A = M^{-1}(M^2 DM^{-2})M = M^{-1}(NDN)M$ is diagonal. This completes the proof.

PROPOSITION 4. *Suppose C is a nonsingular $n \times n$ matrix over F such that, for each $n \times n$ matrix D, $C^{-1}DC$ is diagonal if and only if D is diagonal. Then there exist an $n \times n$ permutation matrix P and an $n \times n$ diagonal matrix E which satisfy $C = PE$.*

Proof. Let D be any diagonal matrix. Let b_{rs}, c_{rs}, and d_{rs} be the (r, s) components of C^{-1}, C, and D, respectively. When $r \neq s$, the (r, s) component of $C^{-1}DC$ is

$$0 = \sum_{j=1}^{n} \sum_{k=1}^{n} b_{rj} d_{jj} \delta_{jk} c_{ks} = \sum_{k=1}^{n} b_{rk} d_{kk} c_{ks}.$$

We set $d_{tt} = 1$ and $d_{kk} = 0$ for $k \neq t$ to obtain $b_{rt} c_{ts} = 0$ whenever $r, s, t = 1, 2, \ldots, n$ and $r \neq s$. Similarly, CDC^{-1} is a diagonal matrix and $c_{rt} b_{ts} = 0$ whenever $r, s, t = 1, 2, \ldots, n$ and $r \neq s$. Suppose $c_{jk} \neq 0$. By

the preceding conditions, only the (k, j) component in either the jth column or the kth row of C^{-1} can be nonzero; since C^{-1} is nonsingular, we have $b_{kj} \neq 0$. Similarly, only the (j, k) component in the jth row or the kth column of C is nonzero. Thus, since C is nonsingular, each row of C has precisely one nonzero component and each column of C has precisely one nonzero component. This yields the conclusion of the proposition and completes the proof.

PROPOSITION 5. *Let π be a permutation of $\{1, 2, \ldots, n\}$ and let ϕ be the bijection of G onto G defined by $\phi(g_j) = g_{\pi(j)}$ for $j = 1, 2, \ldots, n$. Then a necessary and sufficient condition for the correspondence*

$$g_r^{-1} g_s \leftrightarrow g_{\pi(r)}^{-1} g_{\pi(s)} \tag{8}$$

of G onto G to be a well-defined bijection is that, for each x', y', x'', y'' in G, $\phi(x')\phi(y') = \phi(x'')\phi(y'')$ if and only if $x'y' = x''y''$.

Proof. Any element of G is expressible as $g_r^{-1} g_s$, for some r and s. Thus, for a bijection of G onto G to be well defined by (8), it is necessary and sufficient that for each j, k, r, s in $\{1, 2, \ldots, n\}$, $g_j^{-1} g_k = g_r^{-1} g_s$ if and only if $g_{\pi(j)}^{-1} g_{\pi(k)} = g_{\pi(r)}^{-1} g_{\pi(s)}$. This is equivalent to $g_{\pi(k)} g_{\pi(r)} = g_{\pi(j)} g_{\pi(s)}$ if and only if $g_k g_r = g_j g_s$. To complete the proof, we note that this last statement can be rewritten as the condition of the proposition.

A permutation for which the condition of Proposition 5 is satisfied will henceforth be called **admissible**. If M_1 is a nonsingular $n \times n$ matrix over F, we say that M_1 **transforms** (the set of) ***G*-matrices onto** (the set of) **diagonal matrices** when, for any $n \times n$ matrix A over F, $M_1^{-1} A M_1$ is a diagonal matrix if and only if A is a G-matrix. Theorem 1, for instance, shows that both M and M^{-1} transform G-matrices onto diagonal matrices.

THEOREM 2. *Let M_1 be a nonsingular $n \times n$ matrix over F. Then:*

(a) *M_1 transforms G-matrices onto diagonal matrices if and only if there exist an $n \times n$ permutation matrix P and a nonsingular $n \times n$ diagonal matrix E over F such that $M_1 = MPE$;*

(b) *Both M_1 and M_1^{-1} transform G-matrices onto diagonal matrices if and only if there exists an admissible permutation π, a nonzero element c in F, and a homomorphism χ of G into F^* such that*

$$M_1 = MPE, \quad P = [\delta_{r,\pi(s)}], \quad \text{and} \quad E = [c\chi(g_r)\delta_{rs}]. \tag{9}$$

Proof. (a) Suppose M_1 transforms G-matrices onto diagonal matrices. Set $C = M^{-1} M_1$. We use Theorem 1 to verify that C satisfies the hypothesis of Proposition 4. Thus, we obtain $C = PE$ and $M_1 = MPE$.

Now, suppose $M_1 = MPE$. Let A be an $n \times n$ matrix over F. Set $D = M^{-1} A M$, $D_1 = P^T D P$, and $D_2 = E^{-1} D_1 E$. We find $D_2 = M_1^{-1} A M_1$. By Theorem 1, A is a G-matrix if and only if D is diagonal; but, D is diagonal if and only if D_2 is diagonal. Thus, M_1 transforms G-matrices onto diagonal matrices.

(b) Suppose both M_1 and M_1^{-1} transform G-matrices onto diagonal matrices. By (a), a permutation π of $\{1, 2, \ldots, n\}$ and suitable nonzero elements $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ in F exist such that

$$M_1 = MPE, \quad P = [\delta_{r,\pi(s)}], \quad \text{and} \quad E = [\varepsilon_r \delta_{rs}].$$

Let $A = [a_{rs}]$ and $B = [b_{rs}]$ be $n \times n$ matrices over F which satisfy

$$B = M_1^{-1}(MAM^{-1})M_1 = E^{-1} P^T A P E. \tag{10}$$

We obtain

$$b_{rs} = \sum_{i,j,k,t} \varepsilon_r^{-1} \delta_{ri} \delta_{j,\pi(i)} a_{jk} \delta_{k,\pi(t)} \varepsilon_t \delta_{ts} = \varepsilon_r^{-1} \varepsilon_s a_{\pi(r),\pi(s)}$$

for $r = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, n$. By (10), Theorem 1, and our hypothesis, B is a G-matrix if and

only if A is a G-matrix. First, we let A be the G-matrix with $\sigma(g) = 1$, for each g in G, to obtain a map χ of G into F^* such that $\chi(g_r^{-1} g_s) = \varepsilon_r^{-1} \varepsilon_s$ for $r = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, n$. Thus, for each map σ of G into F, there exists a map τ of G into F with

$$\tau(g_r^{-1} g_s) = \chi(g_r^{-1} g_s)\sigma(g_{\pi(r)}^{-1} g_{\pi(s)}), \quad \text{for} \quad r, s = 1, 2, \ldots, n; \tag{11}$$

also, for each map τ of G into F, there exists a map σ of G into F which satisfies (11). This implies that a bijection of G onto G is well defined by (8). From Proposition 5, π is an admissible permutation. For $r = s$, we find $\chi(e) = 1$. Also, given x and y in G, we write $x = g_i^{-1} g_j$ and $y = g_j^{-1} g_k$ to obtain $\chi(xy) = \chi(x)\chi(y)$. Thus, χ is a homomorphism of G into F^* and $\varepsilon_r (\chi(g_r))^{-1} = \varepsilon_s (\chi(g_s))^{-1}$ for $r = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, n$. Set $c = \varepsilon_1 (\chi(g_1))^{-1}$. With $\varepsilon_r = c\chi(g_r)$ for $r = 1, 2, \ldots, n$, we obtain (9).

Finally, suppose M_1 is given by (9). From (a), M_1 transforms G-matrices onto diagonal matrices. Let B be any $n \times n$ matrix over F. Set $D = M_1 B M_1^{-1}$ and $A = M^{-1} D M$. With the notation $A = [a_{rs}]$ and $B = [b_{rs}]$, we obtain (10) and $b_{rs} = \chi(g_r^{-1} g_s) a_{\pi(r), \pi(s)}$ for $r = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, n$. Since π is admissible, we use Proposition 5 to conclude that B is a G-matrix if and only if A is a G-matrix. But, by Theorem 1, A is a G-matrix if and only if D is diagonal. Thus, M_1^{-1} transforms G-matrices onto diagonal matrices. This completes the proof.

When a representation (9) exists for M_1, it is necessarily unique. Namely, suppose P_1, P_2 are permutation matrices and E_1, E_2 are nonsingular diagonal matrices such that $P_1 E_1 = P_2 E_2$. Then, the permutation matrix $P_1^{-1} P_2$ is equal to the diagonal matrix $E_1 E_2^{-1}$. This yields $P_1 = P_2$ and $E_1 = E_2$. As in [**4**], p. 209, the homomorphisms of G into F^* are linearly independent over F. Thus, the representation for E in (9) is also unique.

It follows from Theorem 2 that both M^{-1} and $(M^{-1})^{-1}$ transform G-matrices onto diagonal matrices. From Proposition 3, we obtain $M^{-1} = \left(\frac{1}{n}\right) NM = MN\left(\frac{1}{n}\right)$. Let π be the permutation such that $g_{\pi(j)} = g_j^{-1}$, for $j = 1, 2, \ldots, n$. The proof of Proposition 3 yields $N = [\delta_{r, \pi(s)}]$. In Proposition 5, the map ϕ associated with π satisfies $\phi(g) = g^{-1}$, for each g in G; it is an automorphism of G.

In general, each automorphism of G specifies an admissible permutation by the correspondence between ϕ and π in Proposition 5. Similarly, for any fixed element g_0 in G, the multiplication ϕ defined by $\phi(g) = g_0 g$, for each g in G, specifies an admissible permutation π. There may exist admissible permutations which are not specified by automorphisms or multiplications; for example, such occur when G is a cyclic group of order 4. In Proposition 5, if ϕ_1 and ϕ_2 correspond to π_1 and π_2, then the composite $\phi_1 \circ \phi_2$ corresponds to the composite $\pi_1 \circ \pi_2$. Thus, the set of admissible permutations is easily verified to be a subgroup of the symmetric group of all permutations of $\{1, 2, \ldots, n\}$.

PROPOSITION 6. *Under matrix addition and multiplication, the G-matrices form a commutative ring which is isomorphic to the ring of $n \times n$ diagonal matrices over F. Moreover, if A is a nonsingular G-matrix, then A^{-1} is a G-matrix.*

Proof. To each G-matrix A, we assign the diagonal matrix $M^{-1}AM$. This specifies a one-to-one correspondence between the two sets of matrices which preserves the operations of matrix addition and multiplication. Since the $n \times n$ diagonal matrices over F form a commutative ring, the same is true for the G-matrices. Let A be a nonsingular G-matrix. Then, $M^{-1}AM$ and its inverse $M^{-1}A^{-1}M$ are diagonal. Thus, A^{-1} is a G-matrix. This completes the proof.

(Each matrix $M_1 = MPE$ of Theorem 2 likewise specifies an isomorphism. But, for any G-matrix A, we have $M_1^{-1} A M_1 = P^T (M^{-1} A M) P$. Thus, there are only $n!$ distinct isomorphisms of this type.)

PROPOSITION 7. *Suppose $M_1 = MPE$, as in Theorem 2. Then, an $n \times n$ matrix A over F is a G-matrix if and only if each column vector of M_1 is an eigenvector of A.*

Proof. Set $D = M_1^{-1}AM_1$. From $AM_1 = M_1D$, we note that D is a diagonal matrix if and only if each column vector of M_1 is an eigenvector of A. From this, the conclusion follows by an application of Theorem 2.

PROPOSITION 8. *Suppose $\theta_1, \theta_2, \ldots, \theta_n$ are elements of F and $f(X) = (X-\theta_1)(X-\theta_2)\cdots(X-\theta_n)$. Then, there exists a G-matrix A such that* $\det(XI - A) = f(X)$.

Proof. Set $D = [\theta_r\delta_{rs}]$ and $A = MDM^{-1}$. Then, A is a G-matrix and $\det(XI-A) = \det(M(XI-D)M^{-1}) = \det(XI-D) = f(X)$.

PROPOSITION 9. *For a given map σ of G into F, let A be the G-matrix defined by $A = [\sigma(g_r^{-1}g_s)]$ and let $\xi_1, \xi_2, \ldots, \xi_n$ be defined by*

$$\xi_s = \sum_{k=1}^{n} \sigma(g_k)\chi_s(g_k), \quad \text{for} \quad s = 1, 2, \ldots, n. \tag{12}$$

Then, $\det A = \xi_1\xi_2\cdots\xi_n$ *and*

$$\det(XI - A) = \prod_{s=1}^{n} (X - \xi_s). \tag{13}$$

Proof. For $D = M^{-1}AM$, we use (7) to obtain $D = [\delta_{rs}\xi_s]$. This yields

$$\det(XI - A) = \det(XI - D) = \prod_{s=1}^{n} (X - \xi_s)$$

and $\det A = \det D = \xi_1\xi_2\cdots\xi_n$.

By Proposition 9, the characteristic polynomial of any G-matrix is expressible as a product of linear factors over F. However, formula (12) is encumbered with $\chi_1, \chi_2, \ldots, \chi_n$. To obtain a better computational procedure, we must choose an appropriate enumeration $g_1, g_2, \ldots, g_n$. This we do by means of the next two propositions.

PROPOSITION 10. *Suppose $f_1, f_2, \ldots, f_n$ and $g_1, g_2, \ldots, g_n$ are two enumerations for the elements of G, let σ be a map of G into F, and set $B = [\sigma(f_r^{-1}f_s)]$, $A = [\sigma(g_r^{-1}g_s)]$, and $P = [\delta_{r,\pi(s)}]$, where π is the permutation such that $f_j = g_{\pi(j)}$ for $j = 1, 2, \ldots, n$. Then $B = P^TAP$ and* $\det(XI - B) = \det(XI - A)$. *Moreover, if π is admissible, then B is a G-matrix* **[6]** *relative to $g_1, g_2, \ldots, g_n$.*

Proof. The (r, s) component of P^TAP is

$$\sum_{j=1}^{n}\sum_{k=1}^{n} \delta_{j,\pi(r)}\sigma(g_j^{-1}g_k)\delta_{k,\pi(s)} = \sigma(g_{\pi(r)}^{-1}g_{\pi(s)}) = \sigma(f_r^{-1}f_s).$$

This yields $B = P^TAP$ from which we obtain $\det(XI - B) = \det(XI - A)$. Next, suppose π is admissible. By Proposition 5, there exists a bijection υ of G onto G such that

$$f_r^{-1}f_s = g_{\pi(r)}^{-1}g_{\pi(s)} = \upsilon(g_r^{-1}g_s), \quad \text{for} \quad r, s = 1, 2, \ldots, n.$$

We have $B = [(\sigma \circ \upsilon)(g_r^{-1}g_s)]$, where the composite $\sigma \circ \upsilon$ of σ and υ is a map of G into F. Thus, B is a G-matrix. This completes the proof.

In the context introduced for formulas (1) and (2), let R be the set of q-tuples $(\alpha_1, \alpha_2, \ldots, \alpha_q)$ with integral components which satisfy $0 \leqq \alpha_i < m_i$, for $i = 1, 2, \ldots, q$. Given a q-tuple of R, an integer k is defined by

$$k = 1 + \sum_{i=1}^{q} \alpha_i \left(\prod_{j=1}^{i-1} m_j \right), \quad \text{where} \quad \prod_{j=1}^{0} m_j = 1. \tag{14}$$

Then

$$1 \leqq k \leqq 1 + \sum_{i=1}^{q} (m_i - 1)\left(\prod_{j=1}^{i-1} m_j\right) = n.$$

Set $S = \{1, 2, \ldots, n\}$. Thus, (14) defines a map of R into S.

PROPOSITION 11. *The map defined by* (14) *is a bijection of R onto S.*

Proof. Suppose k is in S. We use the Euclidean algorithm to obtain

$$\begin{aligned} k - 1 &= \gamma_1 m_1 + \alpha_1, \qquad 0 \leqq \alpha_1 < m_1, \\ \gamma_1 &= \gamma_2 m_2 + \alpha_2, \qquad 0 \leqq \alpha_2 < m_2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \gamma_{q-1} &= \gamma_q m_q + \alpha_q, \qquad 0 \leqq \alpha_q < m_q. \end{aligned}$$

We find

$$k - 1 = \alpha_1 + \alpha_2 m_1 + \cdots + \alpha_q (m_1 m_2 \cdots m_{q-1}) + \gamma_q (m_1 m_2 \cdots m_q).$$

With $k - 1 < n$, we must have $\gamma_q = 0$. Thus, (14) is a surjective map of R onto S. To establish directly that (14) is one-to-one, suppose

$$\sum_{i=1}^{q} \alpha_i \left(\prod_{j=1}^{i-1} m_j\right) = \sum_{i=1}^{q} \beta_i \left(\prod_{j=1}^{i-1} m_j\right), \tag{15}$$

where $0 \leqq |\alpha_i - \beta_i| < m_i$ for $i = 1, 2, \ldots, q$. We rewrite (15) to deduce that m_i divides $\alpha_i - \beta_i$ and $\alpha_i = \beta_i$, for $i = 1, 2, \ldots, q$. This completes the proof.

THEOREM 3. *Let $g_1, g_2, \ldots, g_n$ be the elements of G specified in terms of formula* (1), *for $k = 1, 2, \ldots, n$, by*

$$g_k = h_1^{\alpha_1} h_2^{\alpha_2} \cdots h_q^{\alpha_q} \quad \text{and} \quad k = 1 + \sum_{i=1}^{q} \alpha_i \left(\prod_{j=1}^{i-1} m_j\right). \tag{16}$$

For a map σ of G into F, set $A = [\sigma(g_r^{-1} g_s)]$. Then

$$\det(XI - A) = \prod_{(\rho)} \left(X - \sum_{(\alpha)} \sigma(g_k) \rho_1^{\alpha_1} \rho_2^{\alpha_2} \cdots \rho_q^{\alpha_q}\right), \tag{17}$$

where each q-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_q)$ in R specifies k and g_k by (16), *the summation is extended over all n of the q-tuples in R, and the product is extended over all n of the q-tuples $(\rho_1, \rho_2, \ldots, \rho_q)$ in which the ith component is an m_ith root of unity in F.*

Proof. For $i = 1, 2, \ldots, q$, let ζ_i be a primitive m_ith root of unity in F and let ψ_i be the homomorphism of G into F^* with

$$\psi_i(h_i) = \zeta_i \quad \text{and} \quad \psi_i(h_j) = 1 \quad \text{for} \quad j \neq i.$$

We use Proposition 11 to define a homomorphism χ_s of G into F^* by

$$\chi_s = \psi_1^{\beta_1} \psi_2^{\beta_2} \cdots \psi_q^{\beta_q} \quad \text{and} \quad s = 1 + \sum_{i=1}^{q} \beta_i \left(\prod_{j=1}^{i-1} m_j\right), \tag{18}$$

for $s = 1, 2, \ldots, n$. When s ranges from 1 to n, the formula

$$(\chi_s(h_1), \chi_s(h_2), \ldots, \chi_s(h_q)) = (\zeta_1^{\beta_1}, \zeta_2^{\beta_2}, \ldots, \zeta_q^{\beta_q})$$

specifies all n of the q-tuples $(\rho_1, \rho_2, \ldots, \rho_q)$. Since our notation is consistent with (3) and (4), the earlier results for χ_s are applicable. With (16), (12), and (13), we obtain

$$\xi_s = \sum_{(\alpha)} \sigma(g_k)(\chi_s(h_1))^{\alpha_1}(\chi_s(h_2))^{\alpha_2}\cdots(\chi_s(h_q))^{\alpha_q}$$

and (17). This completes the proof.

From the preceding formulas, we find

$$\det A = \prod_{(\rho)} \left(\sum_{(\alpha)} \sigma(g_k)\rho_1^{\alpha_1}\rho_2^{\alpha_2}\cdots\rho_q^{\alpha_q} \right). \tag{19}$$

There are alternate ways to write (17) and (19). For (19), we have

$$\det A = \prod_{(\beta)} \left(\sum_{(\alpha)} \sigma(g_k)\zeta_1^{\alpha_1\beta_1}\zeta_2^{\alpha_2\beta_2}\cdots\zeta_q^{\alpha_q\beta_q} \right), \tag{20}$$

where the q-tuples $(\alpha_1, \alpha_2, \ldots, \alpha_q)$ and $(\beta_1, \beta_2, \ldots, \beta_q)$ range over R and k is specified by (16).

We proceed to describe the matrix A of Theorem 3 in sufficient detail to enable us to derive further results. Let us introduce n variables $X_1, X_2, \ldots, X_n$ and define matrices $V_1, V_2, \ldots, V_q$ as follows. Let V_1 be the $m_1 \times m_1$ circulant matrix in which the sth component of the first row is X_s, for $s = 1, 2, \ldots, m_1$. Suppose V_i has been defined as a matrix of size $u \times u$ where $u = m_1 m_2 \cdots m_i$, the sth component of the first row of V_i is X_s for $s = 1, 2, \ldots, u$, and various permutations of the components of the first row of V_i yield the other rows of V_i. To define V_{i+1} when $i < q$, set $v = m_{i+1}$; for $j = 1, 2, \ldots, v$, let W_{ij} be the $u \times u$ matrix obtained from V_i by the replacement of $X_1, X_2, \ldots, X_u$ in V_i with

$$X_{1+(j-1)u}, X_{2+(j-1)u}, \ldots, X_{u+(j-1)u};$$

form the $v \times v$ block-circulant matrix whose first row is

$$(W_{i1}, W_{i2}, \ldots, W_{iv});$$

and, let V_{i+1} be the resulting matrix of size $w \times w$, with $w = uv$, in which the sth component of the first row is X_s, for $s = 1, 2, \ldots, w$. We find that various permutations of the components of the first row of V_{i+1} yield the other rows of V_{i+1}. If $i + 1 < q$, then the procedure is repeated.

THEOREM 4. *The matrix A of Theorem 3 results when $\sigma(g_1), \sigma(g_2), \ldots, \sigma(g_n)$ are substituted for $X_1, X_2, \ldots, X_n$ in V_q.*

Proof. For $i = 1, 2, \ldots, q$, set $u_i = m_1 m_2 \cdots m_i$ and let G_i be the subgroup of G generated by $\{h_1, h_2, \ldots, h_i\}$. Then, the order of G_i is u_i and the elements of G_i are the first u_i of the elements $g_1, g_2, \ldots, g_n$ from (16).

For $i = 1, 2, \ldots, q$, let T_i be the array of size $u_i \times u_i$ in which the (r, s) component is $g_r^{-1} g_s$ for $r = 1, 2, \ldots, u_i$ and $s = 1, 2, \ldots, u_i$. We shall establish that the (r, s) component of T_i is g_k if and only if the (r, s) component of V_i is X_k.

The (r, s) component of T_1 is $(h_1^{r-1})^{-1} h_1^{s-1} = h_1^{s-r}$. Since T_1 has the pattern of an $m_1 \times m_1$ circulant matrix, the statement about T_i and V_i is true for $i = 1$. Suppose it to be true for a particular integer i with $i < q$. Set $h = h_{i+1}$ and $m = m_{i+1}$. The array T_{i+1} is represented in block form by

$$\begin{bmatrix} T_i & T_i h & \cdots & T_i h^{m-1} \\ T_i h^{m-1} & T_i & \cdots & T_i h^{m-2} \\ \vdots & \vdots & \vdots\,\vdots\,\vdots & \vdots \\ T_i h & T_i h^2 & \cdots & T_i \end{bmatrix}$$

where $T_i h^j$ is the array obtained from T_i when each component of T_i is multiplied by h^j. With (16), the supposition for T_i, and the definition of V_{i+1}, we conclude: the (r, s) component of T_{i+1} is g_k if and

only if the (r, s) component of V_{i+1} is X_k. Thus, the statment about T_i and V_i is true for $i = q$. Since the (r, s) components of T_q and A are $g_r^{-1}g_s$ and $\sigma(g_r^{-1}g_s)$, the matrix A results when, for $k = 1, 2, \ldots, n$, X_k is replaced in V_q by $\sigma(g_k)$. This completes the proof.

We now illustrate Theorems 3 and 4 in the case where G is the Klein four-group. Here, we have $n = 4$, $q = m_1 = m_2 = 2$, $g_1 = e$, $g_2 = h_1$, $g_3 = h_2$, $g_4 = h_1h_2$,

$$V_1 = \begin{bmatrix} X_1 & X_2 \\ X_2 & X_1 \end{bmatrix}, \qquad W_{11} = \begin{bmatrix} X_1 & X_2 \\ X_2 & X_1 \end{bmatrix}, \qquad W_{12} = \begin{bmatrix} X_3 & X_4 \\ X_4 & X_3 \end{bmatrix},$$

$$V_2 = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 \\ X_2 & X_1 & X_4 & X_3 \\ X_3 & X_4 & X_1 & X_2 \\ X_4 & X_3 & X_2 & X_1 \end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_2 & a_1 & a_4 & a_3 \\ a_3 & a_4 & a_1 & a_2 \\ a_4 & a_3 & a_2 & a_1 \end{bmatrix},$$

where $a_i = \sigma(g_i)$, for $i = 1, 2, 3, 4$. From (17), we obtain the formula

$$\det(XI - A) = \prod_{(\rho)} (X - (a_1 + a_2\rho_1 + a_3\rho_2 + a_4\rho_1\rho_2)),$$

in which (ρ_1, ρ_2) ranges over $(1, 1)$, $(-1, 1)$, $(1, -1)$, and $(-1, -1)$. In [3], the matrix A was used to derive formulas for the roots of a quartic equation. Circulant matrices served a similar purpose in [1] and [2].

For a cyclic group G of order n, where $q = 1$ and $m_1 = n$, the G-matrix A of Theorem 3 is circulant and a well-known formula for its determinant results from (19) or (20).

References

[1] R. Chalkley, Cardan's formulas and biquadratic equations, this MAGAZINE, 47 (1974) 8–14.
[2] ———, Circulant matrices and algebraic equations, this MAGAZINE, 48 (1975) 73–80.
[3] ———, Quartic equations and tetrahedral symmetries, this MAGAZINE, 48 (1975) 211–215.
[4] S. Lang, Algebra, Addison–Wesley, Reading, Massachusetts, 1965, pp. 209, 467–468.
[5] C. C. MacDuffee, An Introduction to Abstract Algebra, Wiley, New York, 1950, pp. 72–77.
[6] O. Schreier and E. Sperner, Introduction to Modern Algebra and Matrix Theory, Chelsea, New York, 1959, pp. 260–271.

Divide et impera

> If it were always necessary to reduce everything to intuitive knowledge, demonstration would often be insufferably prolix. This is why mathematicians have had the cleverness to divide the difficulties and to demonstrate separately the intervening propositions. And there is art also in this; for as the mediate truths (which are called *lemmas*, since they appear to be a digression) may be assigned in many ways, it is well, in order to aid the understanding and memory, to choose of them those which greatly shorten the process, and appear memorable and worthy in themselves of being demonstrated.
>
> — GOTTFRIED WILHELM LEIBNITZ

NOTES

A Direct Attack on a Birthday Problem

SAMUEL GOLDBERG
Oberlin College

In a recent article in this MAGAZINE, Nymann [3] computed the probability $P(n, k)$ that in a group of n people at least one pair have the same birthday with at least one member of such a pair among the first k people. After presenting a table for $k(n)$, the smallest value of k for which $P(n, k) \geqq 1/2$, he concludes with the conjecture that "the last entry in the table $[k(254) = 1]$ seems to show that it takes at least 254 people before the expected number of [different] birthdays represented is 183." (Of course, the number 183 is at least half of all possible birthdays in the assumed 365-day year.)

The truth of Nymann's conjecture is easily established from well-known formulas [1, p. 239, Exercise 17; p. 493]. Nevertheless, it may be of interest and pedagogically useful to see how the relevant results can be developed from first principles so they become accessible to beginning students of probability. That is the purpose of this note.

We adopt the standard assumptions for the now well-known birthday problem. (See Feller [1, p. 33] or Mosteller [2, p. 46] for statements and solutions of the basic problem.) Let X_n denote the number of *different* birthdays among n people. Our first task is to develop a formula for $E(X_n)$, the mean or expected value of the random variable X_n. It will then be easy to verify the conjecture that $n = 254$ is the smallest value of n for which $E(X_n) \geqq 183$.

Although it is often not the preferred method of attack (see below), beginners are taught to find the mean of a random variable directly from its probability distribution. Moreover we have here an example where the classroom technique of "experiment with small numbers and see if a pattern can be guessed" turns out to be particularly appropriate. So we proceed by first obtaining the probability distributions of the random variables X_n for $n = 1, 2, 3, 4$. These are presented in TABLE 1 and it will suffice to illustrate how to compute a couple of entries, say $P(X_3 = 2)$ and $P(X_4 = 3)$. When $n = 3$, there are three patterns in which we can have two different birthdays (xxy, xyx, yxx) and each of these occurs $(365)(1)(364)$ times among the $(365)^3$ possible equiprobable orderings of three birthdays. Dividing gives the tabular entry for $P(X_3 = 2)$. Similarly, to determine $P(X_4 = 3)$ we note that there

	Value of X_n			
	1	2	3	4
X_1	1			
X_2	$\frac{1}{365}$	$\frac{364}{365}$		
X_3	$\frac{1}{(365)^2}$	$\frac{3(364)}{(365)^2}$	$\frac{(364)(363)}{(365)^2}$	
X_4	$\frac{1}{(365)^3}$	$\frac{7(364)}{(365)^3}$	$\frac{(6)(364)(363)}{(365)^3}$	$\frac{(364)(363)(362)}{(365)^3}$

Probability distributions of the number X_n of different birthdays among n people.

TABLE 1

are six patterns (*xxyz, xyxz, xyzx, yxxz, yxzx, yzxx*) in which $n = 4$ persons can have three different birthdays and that each of these occurs (365)(1)(364)(363) times among the $(365)^4$ possible orderings of four birthdays. Using the distributions in TABLE 1 and the definition of mean value, we simplify the resulting fractions by writing 364 as (365 − 1), 363 as (365 − 2), etc, to obtain the following suggestive results:

$$E(X_1) = 1$$

$$E(X_2) = 2 - \frac{1}{365}$$

$$E(X_3) = 3 - \frac{3}{365} + \frac{1}{(365)^2}$$

$$E(X_4) = 4 - \frac{6}{365} + \frac{4}{(365)^2} - \frac{1}{(365)^3}.$$

In no class over a number of years has there been any failure to "see" the pattern here, that we have alternating terms and binomial coefficients. The generally accepted "guess" is, of course, that

$$E(X_n) = \sum_{k=1}^{n} (-1)^{k+1} \frac{\binom{n}{k}}{(365)^{k-1}}.$$

Rewriting by adding and subtracting the missing $k = 0$ term and then factoring $(365)^{n-1}$ from the denominator of each term, we have $E(X_n) = (365)^{1-n}\left[(365)^n - \Sigma_{k=0}^{n}\binom{n}{k}(-1)^k(365)^{n-k}\right] = (365)^{1-n}[(365)^n - (365-1)^n]$. Hence our "guess" becomes the simple formula

$$E(X_n) = 365 - (364)^n/(365)^{n-1} \qquad (n = 1, 2, 3, \ldots).$$

Finding the smallest value of n so that $E(X_n) \geqq 183$ is equivalent to finding the smallest n for which $(364/365)^{n-1} \leqq 1/2$. Turning to a table of common logarithms, we quickly find that $n = 254$, as conjectured. (See Mosteller [2, p. 48] for a solution to the very closely related "birthmate" problem.)

To be complete, our classroom unit requires only that the "guessed" formula for $E(X_n)$ be proved correct. This we do most simply by first noting that each day is somebody's birthday or nobody's birthday. It turns out to be easier to focus on the birthdayless days. Define the random variables Y_k to have the value 1 if day number k is the birthday of nobody among the n people under consideration, and to have the value 0 otherwise, for $k = 1, 2, \ldots, 365$. The total number of birthdayless days is then the sum $(Y_1 + Y_2 + \cdots + Y_{365})$. Since for each k

$$E(Y_k) = P(Y_k = 1) = (364/365)^n,$$

the expected number of birthdayless days is $365(364/365)^n$. Subtracting this from 365 yields the expected number of different birthdays among n people and completes the proof that our formula for $E(X_n)$ is correct. And this demonstration allows one final point to be made as the bell rings: that one can often most efficiently compute the mean value of a random variable *without* ever knowing its probability distribution.

References

[1] W. Feller, An Introduction to Probability Theory and its Applications, vol. 1, 3rd edition, Wiley, New York, 1968.

[2] F. Mosteller, Fifty Challenging Problems in Probability with Solutions, Addison-Wesley, Reading, Mass., 1965.

[3] J. E. Nymann, Another generalization of the birthday problem, this MAGAZINE, 48 (1975) 46–47.

Distribution of Orders of Abelian Groups

JONATHAN M. KANE
University of Wisconsin

In a recent paper, Gallian [1] gives an account of some computer related projects in group theory. In one of these projects a study was made (see TABLE 1) to determine, for each integer k, the percentages of integers in a given interval which are orders of exactly k Abelian groups. For example, the data in Table 1 shows that in the interval from 50001 to 50500, 19.80% of the integers correspond to orders of two non-isomorphic Abelian groups while 2.00% correspond to orders of six non-isomorphic Abelian groups. One quickly notices that the percentages in TABLE 1 are, to a large extent, independent of the interval of integers analyzed. The purpose of this paper is to explain this occurrence.

No. of groups	1 – 10000	50001 – 50500	500001 – 505000	500001 – 550000	900001 – 901000	999001 – 1000000
1	60.83	61.00	60.82	60.80	60.90	60.80
2	20.08	19.80	20.04	20.06	19.90	19.70
3	7.44	7.20	7.38	7.42	7.70	7.70
4	2.20	2.20	2.22	2.23	2.40	2.30
5	3.21	3.40	3.20	3.20	3.00	3.20
6	1.46	2.00	1.42	1.45	1.20	1.60
7	1.51	1.60	1.50	1.47	1.60	1.40
8	0.08	0.00	0.14	0.10	0.10	0.10
9	0.22	0.20	0.28	0.22	0.30	0.10
10 or more	2.97	2.60	3.00	3.05	2.90	3.10

Percentages of integers in given ranges that are the order of the given number of distinct Abelian groups.

TABLE 1

The number of Abelian groups of a given order can easily be determined by applying the fundamental theorem of finite abelian groups [2, Sec. 214]: if $r = p_1^{e_1}p_2^{e_2}p_3^{e_3}\cdots p_s^{e_s}$ where the p_i's are distinct primes, the number of Abelian groups with order r is $g(r) = \alpha(e_1)\alpha(e_2)\alpha(e_3)\cdots\alpha(e_s)$, where $\alpha(e)$ is the number of ways e can be written as the sum of positive integers. Some values of the function α (often called the **partition function**) are given in TABLE 2.

The question posed by Gallian's paper can now be restated in number theoretic terms: For a given integer k, what percent of the integers in a given interval satisfy the equations $g(r) = k$? Letting Z be the set of positive integers, define $A_k = \{r \in Z: g(r) = k\}$. To answer the preceding question we will,

e:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$\alpha(e)$:	1	2	3	5	7	11	15	22	30	42	56	77	101	135	176	213	297	385	490	627

Number of ways, $\alpha(e)$, of writing the positive integer e as a sum of positive integers.

TABLE 2

for convenience, analyze A_k as a whole rather than over finite intervals. To do this, we introduce the following concept of density. If $X \subset Z$, let $N(X, n)$ be the number of elements in X less than or equal to n. Then define the **density** of X to be $D(X) = \lim_{n\to\infty} N(X, n)/n$ if the limit exists. Now let us consider the density of the sets A_k.

From the definition of $g(r)$, it is clear that $g(r) = 1$ if and only if no prime appears more than once in the prime factorization of r. For $g(r)$ to equal 2, exactly one prime factor of r must be repeated. One can use TABLE 2 to construct in this manner the solution sets for $g(r) = k$ for any k; see TABLE 3. It follows from this table that, for example, $2^3 \cdot 3^2 \cdot 5 \cdot 13$ is an element of A_6 and $3^2 \cdot 7 \cdot 11^2 \cdot 19^2 \cdot 23$ is an element of A_8. The structure of A_k may vary from trivial (e.g., A_{13} is empty) to complex. For example, since $30 = \alpha(9) = \alpha(2)\,\alpha(3)\,\alpha(4) = \alpha(2)\,\alpha(7)$, an element is in A_{30} if it has exactly one prime repeated nine times, or if it has exactly one prime appearing twice, another appearing 3 times, and a third appearing four times, or again if it has one prime appearing twice and another seven times.

k	Description of the elements of A_k
1	No prime factor can be repeated.
2	Exactly 1 prime factor must appear twice.
3	Exactly 1 prime factor must appear 3 times.
4	Exactly 2 prime factors must appear twice each.
5	Exactly 1 prime factor must appear 4 times.
6	Exactly 1 prime factor must appear twice and a second must appear 3 times.
7	Exactly 1 prime factor must appear 5 times.
8	Exactly 3 prime factors must appear twice each.
9	Exactly 2 prime factors must appear 3 times each.
10	Exactly 1 prime factor must appear twice and a second must appear 4 times.
11	Exactly one prime factor must appear 6 times.
12	Exactly 2 prime factors must appear twice each and a third must appear three times.
13	A_{13} is empty since 13 is a prime that is not $\alpha(e)$ for any e.
14	Exactly 1 prime factor must appear twice and a second must appear 5 times.

Criteria that an integer be in A_k, the set of integers n such that there exist exactly k Abelian groups of order n.

TABLE 3

Beginning now with A_1, if $P_i = \{r \in Z: p_i^2 | r\}$ where p_i is the ith prime, $A_1 = Z - \bigcup_{i\in Z} P_i = \bigcap_{i\in Z}(Z - P_i)$. It is clear that the density of $P_i = D(P_i) = 1/p_i^2$ and that $D(Z - P_i) = 1 - 1/p_i^2$. Intuitively it seems that $D(A_1) = \prod_{i\in Z}(1 - 1/p_i^2)$ but one needs a couple of lemmas to make this rigorous.

LEMMA 1. *$D(X \cup Y) = D(X) + D(Y) - D(X \cap Y)$ when any three of these terms exist.*

Proof. Clearly, $N(X \cup Y, n) = N(X, n) + N(Y, n) - N(X \cap Y, n)$ so

$$\begin{aligned} D(X \cup Y) &= \lim_{n\to\infty} N(X \cup Y, n)/n \\ &= \lim_{n\to\infty} N(X, n)/n + \lim_{n\to\infty} N(Y, n)/n - \lim_{n\to\infty} N(X \cap Y, n)/n \\ &= D(X) + D(Y) - D(X \cap Y). \end{aligned}$$

The sets $P_1, P_2, P_3, \ldots$ are in a sense independent in that each P_i is a set of multiples of a p_i^2 and the p_i's are relatively prime. This allows the P_i's to satisfy the hypothesis of the next lemma.

LEMMA 2. *Let $\{X_i\}$ be a sequence of subsets of Z. If the X_i's have the property that for any $m \in Z$*

$$D\left(\bigcup_{j\leq m} X_j \cap X_{m+1}\right) = D\left(\bigcup_{j\leq m} X_j\right) \cdot D(X_{m+1}),$$

then for any $m \in Z$

$$D\left(\bigcup_{j\leq m} X_j\right) = 1 - \Pi_{j\leq m}[1 - D(X_j)].$$

Proof. This is trivially true for $m = 1$. Assuming that

$$D\left(\bigcup_{j\leq m} X_j\right) = 1 - \Pi_{j\leq m}[1 - D(X_j)],$$

apply Lemma 1 to get

$$\begin{aligned} D\left(\bigcup_{j\leq m+1} X_j\right) &= D\left(\bigcup_{j\leq m} X_j\right) + D(X_{m+1}) - D\left[\bigcup_{j\leq m} X_j \cap X_{m+1}\right] \\ &= 1 - \Pi_{j\leq m}[1 - D(X_j)] + D(X_{m+1}) - D\left(\bigcup_{j\leq m} X_j\right) \cdot D(X_{m+1}) \\ &= 1 - \Pi_{j\leq m+1}[1 - D(X_j)]. \end{aligned}$$

Thus the lemma is proved by induction.

Another important property of the P_i's is that for each $n \in Z$, $N(P_i, n) = [n/p_i^2] \leqq n/p_i^2$ so $N(P_i, n)/n \leqq 1/p_i^2$ for all n. Thus, since $\Sigma_{i\in Z} 1/p_i^2$ converges, this will allow the next lemma to be applied to the sequence $\{P_i\}$.

LEMMA 3. *Let $\{X_i\}$ be a sequence of subsets of Z. If for each $i \in Z$ there is an s_i such that $s_i \geqq N(X_i, n)/n$ for all n and $\Sigma_{i\in Z} s_i < \infty$, then $D(\bigcup_{i\in Z} X_i) = \lim_{j\to\infty} D(\bigcup_{i\leq j} X_i)$.*

Proof. Let $Y_1 = X_1$ and $Y_{i+1} = X_{i+1} - \bigcup_{j\leq i} X_j$. Then the Y_i's are disjoint and Lemma 1 shows that $D(\bigcup_{i\leq j} X_i) = \Sigma_{i\leq j} D(Y_i)$. Since $\Sigma_{i\leq j} D(Y_i)$ is bounded by 1, $\Sigma_{i\in Z} D(Y_i)$ converges. Thus $\lim_{j\to\infty} D(\bigcup_{i\leq j} X_i) = \Sigma_{i\in Z} D(Y_i) = \Sigma_{i\in Z} \lim_{n\to\infty} N(Y_i, n)/n$. Now $s_i \geqq N(X_i, n)/n \geqq N(Y_i, n)/n$ for all n, so the dominated convergence theorem [3, Sec. 1.34] can be applied to obtain

$$\lim_{j\to\infty} D\left(\bigcup_{i\leq j} X_i\right) = \lim_{n\to\infty} \Sigma_{i\in Z} N(Y_i, n)/n = \lim_{n\to\infty} N\left(\bigcup_{i\in Z} Y_i, n\right) \Big/ n = D\left(\bigcup_{i\in Z} X_i\right).$$

Now since $A_1 = Z - \bigcup_{i\in Z} P_i$, $D(A_1) = 1 - D(\bigcup_{i\in Z} P_i)$. Applying Lemmas 2 and 3 to $D(\bigcup_{i\in Z} P_i)$ yields

$$D\left(\bigcup_{i\in Z} P_i\right) = \lim_{j\to\infty} D\left(\bigcup_{i\leq j} P_i\right) = \lim_{j\to\infty}\left[1 - \prod_{i\leq j}(1 - 1/p_i^2)\right] = 1 - \prod_{i\in Z}(1 - 1/p_i^2).$$

As was cleverly shown by Euler [4, Sec. 9.11] $D(A_1) = \Pi_{i\in Z}(1 - 1/p_i^2) = 1/(\Sigma_{i\in Z} 1/i^2) = 6/\pi^2$ which is approximately 60.793%. This corresponds very nicely with the data from the first row of TABLE 1. $D(A_1)$ can obviously be approximated by computing $N(A_1, n)/n$ for large n. Also, since $D(A_1)$ in the limit does not depend on the behavior of A_1 in any finite interval, any interval (not just those beginning with 1) can be used to approximate $D(A_1)$. Since the product $\Pi_{i\in Z}(1 - 1/p_i^2)$ converges rapidly, the approximations of $D(A_1)$ made by taking averages over long intervals such as those in TABLE 1 end up being sharp approximations and, therefore, the averages are, to a large part, independent of the intervals.

$D(A_2)$ is only slightly more complicated to calculate than $D(A_1)$. By TABLE 3, the elements in A_2

k	$D(A_k)$	Approximation in %
1	$\frac{6}{\pi^2}$	60.792
2	$\frac{6}{\pi^2}\Sigma\frac{1}{p_i}\left(\frac{1}{p_i+1}\right)$	20.059
3	$\frac{6}{\pi^2}\Sigma\frac{1}{p_i^2}\left(\frac{1}{p_i+1}\right)$	7.412
4	$\frac{3}{\pi^2}\left(\Sigma\frac{1}{p_i}\frac{1}{p_i+1}\right)^2-\frac{3}{\pi^2}\Sigma\frac{1}{p_i^2}\frac{1}{(p_i+1)^2}$	2.207
5	$\frac{6}{\pi^2}\Sigma\frac{1}{p_i^3}\frac{1}{p_i+1}$	3.207
6	$\frac{6}{\pi^2}\Sigma\frac{1}{p_i}\frac{1}{p_i+1}\Sigma\frac{1}{p_i}\frac{1}{p_i+1}-\frac{6}{\pi^2}\Sigma\frac{1}{p_i^3}\frac{1}{(p_i+1)^2}$	1.446
7	$\frac{6}{\pi^2}\Sigma\frac{1}{p_i^4}\frac{1}{p_i+1}$	1.474
8	$\frac{1}{\pi^2}\left(\Sigma\frac{1}{p_i}\frac{1}{p_i+1}\right)^3-\frac{3}{\pi^2}\Sigma\frac{1}{p_i}\frac{1}{p_i+1}\Sigma\frac{1}{p_j^2}\frac{1}{(p_j+1)^2}+\frac{2}{\pi^2}\Sigma\frac{1}{p_i^3}\frac{1}{(p_i+1)^3}$	0.107
9	$\frac{3}{\pi^2}\left(\Sigma\frac{1}{p_i^2}\frac{1}{p_i+1}\right)^2-\frac{3}{\pi^2}\Sigma\frac{1}{p_i^4}\frac{1}{(p_i+1)^2}$	0.216

Precise formulas and numerical approximations for the densities of the sets A_k of integers n for which there are exactly k Abelian groups of order n.

TABLE 4

are those positive integers whose factorization contain exactly one prime squared. Let $Q_j = \{r \in Z: p_i^2 | r \text{ and } p_i^3 \nmid r\}$. Then $A_2 = \bigcup_{j\in Z}(Q_j - \bigcup_{i\neq j} P_i)$. For any set X, let X^* denote $Z - X$. Then $A_2 = \bigcup_{j\in Z}[Q_j \cap (\bigcup_{i\neq j} P_i)^*]$. Clearly, $D(Q_j) = 1/p_j^2 - 1/p_j^3$ and $D[(\bigcup_{i\neq j} P_i)^*] = [\Pi_{i\in Z}(1-1/p_i^2)]/(1-1/p_j^2) = 6/(1-1/p_j^2)\pi^2$. Thus, by arguing as above, one obtains the formula

$$D(A_2) = \Sigma_{j\in Z}\frac{6}{\pi^2}\left(\frac{1}{p_j}\right)\left(\frac{1}{p_j+1}\right).$$

This sum has an approximate value of 20.059%.

The same procedure will yield the formulas for $D(A_3)$, $D(A_5)$, and $D(A_7)$. Formulas for $D(A_4)$, $D(A_6)$, $D(A_8)$, and $D(A_9)$ which are listed in TABLE 4 are slightly more difficult and are left for the reader to prove. All the formulas yield numbers which agree closely with the values in TABLE 1.

References

[1] J. A. Gallian, Computers in group theory, this MAGAZINE, 49 (1976) 69–73.
[2] I. N. Herstein, Topics in Algebra, Xerox, Lexington, 1975.
[3] W. Rudin, Real and Complex Analysis, McGraw–Hill, New York, 1974.
[4] ———, Functional Analysis, McGraw–Hill, New York, 1973.

I

A Generalization of a Putnam Problem

C. C. CLEVER
K. L. YOCOM
South Dakota State University

The following problem appeared on the 1973 Putnam Examination: *Let $a_1, a_2, \ldots, a_{2n+1}$ be integers such that, if any one of them is removed, those remaining can be divided into two sets of n having equal sums. Prove $a_1 = a_2 = \cdots = a_{2n+1}$.* A proof may be based on special properties of integers. (Show that

the given integers are either all even or all odd. Then if they are all even they may be divided by 2 while if they are all odd they may be increased by 1 without destroying the property of the problem.) In generalizing the problem, we developed a different proof which is an interesting application of linear algebra. We begin with two generalizations of the problem, which we prove by means of a lemma concerning matrices. Then we state and prove a further generalization as our main theorem.

GENERALIZATION 1. *Let $x_1, x_2, \ldots, x_{2n+1}$ be complex numbers such that, if any one of them is removed, those remaining can be divided into two sets of n having equal sums; then $x_1 = x_2 = \cdots = x_{2n+1}$.*

GENERALIZATION 2. *Let $x_1, x_2, \ldots, x_{2n}$ be complex numbers such that, if any one of them is removed, those remaining can be divided into two sets having equal sums; then $x_1 = x_2 = \cdots x_{2n} = 0$.*

LEMMA. *If A is an n by n matrix having zeros on the main diagonal and all ± 1's off the diagonal, then A is nonsingular if n is even and the rank of A is at least $n-1$ if n is odd.*

Proof. In the expansion of $\det A$, each term is 0, 1 or -1 and the number, d_n, of nonzero terms in the expansion is the number of permutations of order n which leave no element fixed. Such permutations are commonly called derangements and it is well known [1, p. 31] that $d_1 = 0$, $d_2 = 1$ and $d_{n+2} = (n+1)(d_n + d_{n+1})$ for $n \geqq 1$. It follows inductively that d_n is even for n odd and d_n is odd for n even. Thus if n is even, $\det A \neq 0$ while if n is odd, each principal submatrix of A of order $n-1$ has a nonzero determinant. This completes the proof of the lemma.

Proof of 1. Let $x = \operatorname{col}(x_1, x_2, \cdots, x_{2n+1})$ and let A be a $2n+1$ by $2n+1$ matrix having zeros on the main diagonal and exactly n entries equal to 1 and n equal to -1 in each row. Then the components of a solution vector x of $Ax = 0$ satisfy the hypotheses of Generalization 1. Since $x_0 = \operatorname{col}(1, 1, \ldots, 1)$ is one such solution vector, A is singular and by the lemma, A has rank $2n$. Thus all solutions are of the form $x = cx_0 = \operatorname{col}(c, c, \ldots, c)$.

Proof of 2. Let $x = \operatorname{col}(x_1, x_2, \ldots, x_{2n})$ and let A be a $2n$ by $2n$ matrix with zeros on the main diagonal and ± 1's off the diagonal. Then A is nonsingular by the lemma and hence $Ax = 0$ has only the trivial solution $x = 0$.

THEOREM. *Let k and n be positive integers satisfying $n > 2$ and $1 \leqq k \leqq n-2$.*

(a) *If $n - k = 2m$, an even integer, and $x_1, x_2, \ldots, x_n$ is a sequence of complex numbers such that, if any k of them are removed, those remaining can be divided into two sets of m having equal sums, then $x_1 = x_2 = \cdots = x_n$.*

(b) *If $n - k = 2m + 1$, an odd integer, and $x_1, x_2, \ldots, x_n$ is a sequence of complex numbers such that, if any k of them are removed, those remaining can be divided into two sets having equal sums, then $x_1 = x_2 = \cdots = x_n = 0$.*

Proof. The theorem is true for $k = 1$ by 1 and 2 above. Now proceed inductively on k, assuming the theorem true for $k = 1, 2, \ldots, K-1 < n-2$. First suppose $n - K = 2m$ in which case we are to establish (a) for $k = K$. Let x_i and x_j be any two designated elements of the sequence with $i \neq j$. Remove any $K-1$ elements of the sequence but leave x_i and x_j (this is possible since $n - K + 1 \geqq 3$). Then we are left with a sequence of length $2m+1$ satisfying the hypotheses of 1 and hence $x_i = x_j$. Thus $x_1 = x_2 = \cdots = x_n$. Similarly if $n - K = 2m + 1$ we must establish (b) for $k = K$. Again, remove $K-1$ elements of the sequence but this time leave some designated element x_i. The remaining sequence of length $n - K + 1 = 2m$ satisfies the hypotheses of 2 and hence $x_i = 0$. Thus $x_1 = x_2 = \cdots = x_n = 0$.

Reference

[1] H. J. Ryser, Combinatorial Mathematics, MAA Carus Monograph No. 14, 1963.

Solid Polyomino Constructions

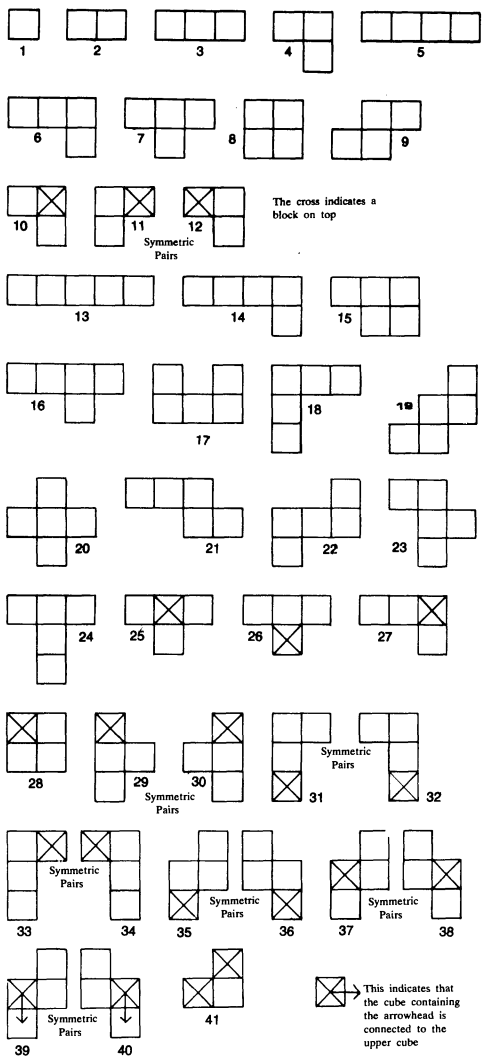
SCOTT L. FORSETH, student
Oak Grove Junior High School,
Clearwater, Florida

A plane polyomino of order n is defined by Solomon W. Golomb [4] as a set of n congruent squares that are simply connected, edge-to-edge. A three-dimensional, or solid, polyomino of order n is a set of n congruent cubes that are simply connected, face-to-face. These are easily made by gluing like cubes together in all permissible patterns. No formula has been discovered which gives the number of different polyominoes of order n, plane or solid. The solid ones of orders one to five are shown in FIGURE 1, each being identified by a number. The set contains 41 pieces and a total of 186 individual or unit cubes.

Martin Gardner has discussed polyominoes in his Mathematical Games section of *Scientific American* and also in each of his three books [1, 2, 3]. Chapters thirteen in [1] and [3] are concerned with plane polyominoes. Chapter six in [2] deals with the well-known Soma cube of Piet Hein consisting of the pieces numbered 4, 6, 7 and 9–12 in FIGURE 1. Gardner points out that there are more than 230 essentially different ways of stacking these pieces in a $3 \times 3 \times 3$ cube (or simply, a 3-cube) but to date no one knows precisely how many.

A similar cube occurs in Steinhaus [5, p. 168] where pieces numbered 6, 11, 12, 30, 33 and 40 are used. It is easy to prove that there are just two solutions to the Steinhaus cube. The 30 and the 40 go together successfully in just one way. Then the 33 has only two possible positions after each of which the positions of 6, 11 and 12 are uniquely determined.

Steinhaus remarks: "Two pieces are congruent by symmetry. (Was it possible to avoid it?)". He refers to numbers 11 and 12. We do not know his answer but it is certainly possible to avoid using either 11 or 12. Indeed, a 3-cube can be constructed using, in addition to the four polyominoes 6, 30, 33, 40, any of the pairs 8, 11 or 8, 12 or 9, 11 or 9, 12 or, for that matter,



Representations of all polyominoes of orders one through five. They use a total of 186 unit cubes: 9 in the four trivial polyominoes (Nos. 1–4) of orders 1–3; 32 in the eight polyominoes (Nos. 4–12) of order 4; and 145 in the 29 polyominoes (Nos. 13–41) of order five. No general formula is known that permits easy extension of this pattern.

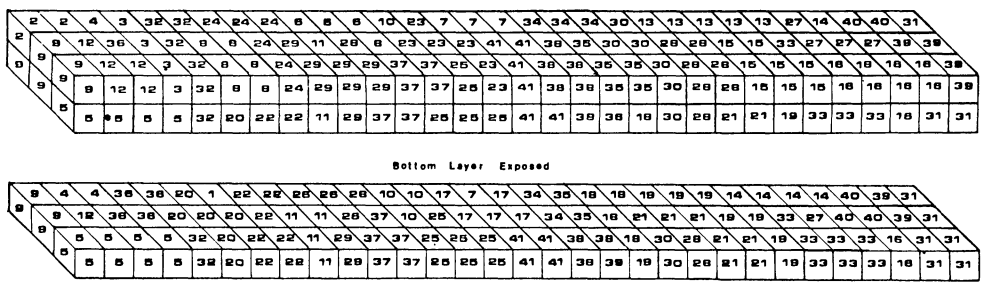
FIGURE 1

two 11's or two 12's. It is remarkable that, in the last two cases, the configuration of the 6, 30, 33 and 40 is the same. That is, two 11's and again two 12's can be placed in congruent arrangements so as to fill the same space. This space can be visualized easily by placing the 15 flat on the table with the "extra" unit cube protruding from the unit cube in the northwest corner of the square and then covering the southwest, northwest and northeast cubes of the square with the 4.

Many 3-cubes can be formed by choosing various combinations of the polyominoes of orders one to five. The following seem to be of special interest. We discard (by necessity) the number 5 piece and note that the remaining seven polyominoes of order four contain a total of 28 individual cubes. To reduce this number to the required 27 we substitute either number 3 or 4 for one of the seven numbered 6–12. Since there are two ways of choosing a three and seven ways of discarding a four, there are fourteen cases altogether. Thirteen of them stack into a 3-cube. One case is impossible: numbers 3, 6 and 8–12 cannot be assembled into a cube.

Other questions arise. Is it possible to build a 4-cube using only polyominoes of orders four and five? There are just two ways of getting the necessary 64 unit cubes: (a) by using one of order four (eight cases) and the other twelve of order five and (b) by using six of order four (28 cases) and the other eight of order five. All thirty-six cases are possible, each in several ways. Or again, is it possible to construct a 5-cube using only polyominoes of order five? The answer is yes and in many ways.

There are some interesting constructions involving parallelepipeds. If we discard six of the 186 total unit cubes, there remain $180 = 2^2 \times 3^2 \times 5$. There are many ways to remove polyominoes with a total of six unit cubes. We handicapped ourselves in trying to stack a $4 \times 5 \times 9$ by discarding the 1, 2 and 3 which are the simplest to fit in. The handicap was too great and we were unsuccessful. But we were able to construct two different ones by discarding the 2 and 8 and another by discarding the 1 and 28. We did not consider other arrangements of the factors of 180 but we did investigate the $2 \times 3 \times 5$ parallelepipeds using just the prime factors and omitting the 1, 2 and 3. They can be made with six 5's but the more interesting case uses two 5's and five 4's. The two 5's can be selected in 406 ways and the stacking is possible in each!



An arrangement of all 41 polyominoes of order one through five in a rectangular block of size $2 \times 3 \times 31 = 186$.

FIGURE 2

If we discard four unit cubes, there remain $182 = 2 \times 7 \times 13$ and we built one omitting the 8. By discarding the 2 we get $184 = 2^3 \times 23$ and we made a $2 \times 4 \times 23$.

But the case of maximum interest, and difficulty, uses all polyominoes of orders one to five. After many trials we found a $2 \times 3 \times 31$ (= 186); see FIGURE 2. Indeed we made two, so the problem does not have a unique solution.

The author wishes to acknowledge the encouragement and guidance of Cletus Oakley of Haverford College in the preparation of this paper.

References

[1] Martin Gardner, The Scientific American Book of Mathematical Puzzles and Diversions, Simon and Schuster, New York, 1959.

[2] Martin Gardner, The Second Scientific American Book of Mathematical Puzzles and Diversions, Simon and Schuster, New York, 1961.
[3] ———, New Mathematical Diversions from Scientific American, Simon and Schuster, New York, 1966.
[4] Solomon Golomb, Polyominoes, Charles Scribner's, New York, 1965.
[5] Hugo Steinhaus, Mathematical Snapshots, Oxford University Press, New York, 1969.

Continuity of Coordinate Functionals

PAUL MILNES
University of Western Ontario

Let V be an n-dimensional normed linear space over the complex numbers and let $\{x_1, x_2, \ldots, x_n\} \subset V$ be a basis for V. Then, for each i, $1 \leqq i \leqq n$, the coordinate functional, f_i corresponding to x_i, which is defined for $x = \Sigma_1^n a_j x_j \in V$ by $f_i(x) = a_i$, is continuous. It is natural to ask for an analogue of this result in infinite dimensional normed linear space, and one is known. Let V be a Banach space with a Schauder (or topological) basis $\{x_n\}_{n=1}^{\infty}$: each $x \in V$ has a unique representation $x = \Sigma_1^{\infty} a_n x_n$, the series converging in V. (Many Banach spaces are known to have Schauder bases; however Enflo [2] has recently constructed a separable infinite dimensional Banach space which does not have a Schauder basis.) Then the coordinate functionals $f_i(x) = f_i(\Sigma_1^{\infty} a_n x_n) = a_i$ are continuous (see [3; 9.6, p. 115], for example).

One might wonder about a similar result using a Hamel (or algebraic) basis instead of a Schauder basis. Bronson and Steiner [1] have shown by specific example that a Hamel basis coordinate functional can be discontinuous. We now show that discontinuity of such functionals is not unusual. Let V be a Banach space with Schauder basis $\{x_n\}$ and extend $\{x_n\}$ to a Hamel basis $\{x_n\} \cup \{y_\gamma\}_{\gamma \in I}$: each $x \in V$ has a unique representation $x = \Sigma b_n x_n + \Sigma b_\gamma x_\gamma$, where only a finite number of the b_n's and b_γ's are non-zero. It follows from the density in V of the linear span of $\{x_n\}$ that no coordinate functional corresponding to any of the y_γ's can be continuous. And the (Hamel basis) coordinate functional corresponding to an x_m will be continuous if and only if $a_{\gamma,m}$ in the Schauder basis expansion $y_\gamma = \Sigma a_{\gamma,n} x_n$ is zero for all $\gamma \in I$. For, when this latter condition holds, the Hamel basis coordinate b_m of $x \in V$ is the same as its Schauder basis coordinate a_m, and is a continuous function of x. And if $y_\gamma = a_{\gamma,m} x_m + \Sigma_{n \neq m}^{\infty} a_{\gamma,n} x_n$ with $a_{\gamma,m} \neq 0$, then $\{a_{\gamma,m} x_m + \Sigma_{n=k}^{\infty} a_{\gamma,n} x_n\}_{k=m+1}^{\infty}$ gives a sequence converging to $a_{\gamma,m} x_m$ whose members all have Hamel basis coordinate b_m equal to zero, while $a_{\gamma,m} x_m$ has that coordinate equal to $a_{\gamma,m}$, which is not equal to zero.

(If $V = \ell_1$, $x_1 = (1,0,0,\ldots) \in \ell_1$ and $y_k = (10^k, 1/k, 1/k^2, 1/k^3, \ldots) \in \ell_1$, $k = 2,3,4,\ldots$, one can see that the set $\{x_1, y_2, y_3, y_4, \ldots\}$ is linearly independent in ℓ_1 and hence can be extended to a Hamel basis $\mathcal{H}$ for ℓ_1, and that the Hamel basis coordinate functional corresponding to x_1 is not continuous. This is the example of Bronson and Steiner in [1]; the discontinuity of the functional can be demonstrated using the methods outlined above at least if $\mathcal{H}$ is assumed to contain the remaining members of the usual Schauder basis for ℓ_1, namely, $\{(0,1,0,0,\ldots),(0,0,1,0,0,\ldots),\ldots\}$.)

Thus, one can have a Hamel basis with any finite number of coordinate functionals continuous and all the others discontinuous.

References

[1] R. Bronson and G. Steiner, A note on conjugate spaces, this MAGAZINE, 46 (1973) 158.
[2] P. Enflo, A counterexample to the approximation problem in Banach spaces, Acta Math., 130 (1973) 309–317.
[3] H. H. Schaefer, Topological Vector Spaces, Springer-Verlag, New York, 1971.

*n*th Root Groups

ROBERT E. KENNEDY
ROBERT W. BUSBY
Central Missouri State University

We say that an element a in a group G, written multiplicatively with identity e, has an **nth root** if there is an x in G such that $x^n = a$. If, for a given positive integer n, every element of G has an nth root we call G an **nth root group.** A little experimentation with examples will show that an nth root group need not be an mth root group if $m \neq n$. The purpose of this paper is to give some conditions which, for a given n, imply that a group is an nth root group. For finite groups the situation is, as we shall see, fairly simple and well known. For infinite groups, however, satisfactory criteria are much more difficult, especially for nonabelian groups.

The following proposition completely characterizes finite nth root groups. It is probably well known to most mathematicians (for instance, condition 4 of Proposition 1 is Theorem 2 of [1]), and its proof uses techniques which are available to the beginning student of group theory. For completeness, however, we will include an outline of its proof.

PROPOSITION 1. *The following statements are equivalent for a finite group G of order r.*

(1) *G is an nth root group.*
(2) *If $x^n = y^n$, where x and y belong to G, then $x = y$.*
(3) *If $x^n = e$, where x belongs to G, then $x = e$.*
(4) *The order of G and n are relatively prime.*
(5) *For each positive integer t, and each $a \in G$, there exists $x \in G$ such that*

$$a = (x)^{n^t}.$$

Proof. Define $f: G \to G$ by $f(x) = x^n$. Then f is a mapping which is onto if G is an nth root group, and thus also one-to-one since G is a finite set. Hence whenever $x^n = y^n$, it follows that $x = y$. So (1) implies (2), and (2) obviously implies (3). To see that (3) implies (4) let p be a prime integer which divides both n and r, so $n = pt$. Since p divides r, by the first Sylow theorem there exists $y \neq e$ in G such that $y^p = e$, and in turn $y^n = (y^p)^t = e$, a contradiction. Thus there cannot exist a prime which divides both n and r, i.e., n and r are relatively prime. Letting $t = 1$ we see that (5) implies (1) so it remains to show that (4) implies (5). Now if n and r are relatively prime, there exist integers u, v such that $un + vr = 1$. Therefore if t is any positive integer, it follows that $1 = 1^t = (un + vr)^t = (un)^t + kr = u^t n^t + kr$ for some integer k. Thus since $a^r = e$ for any a in G,

$$a = a^{(u^t n^t + kr)} = (a^{u^t})^{n^t}$$

for each positive integer t. Set $x = a^{u^t}$ and see that (4) implies (5), completing the proof.

It is necessary that the order of G be finite with respect to conditions (2) and (3) of Proposition 1. For example, consider the group of nonzero complex numbers C under ordinary multiplication. Each element in C has n distinct nth roots for each positive integer n. Thus conditions (2) and (3) fail to hold even though C under multiplication is an nth root group. In the case of a finite nth root group, it should be pointed out that condition (2) of Proposition 1 states that each element has a unique nth root.

Let x be an element of a finite nth root group G. Then a sequence $x_0, x_1, \ldots, x_j, \ldots$ of elements of G may be generated by taking successive nth roots of x, i.e., $x_0 = x$ and for each positive integer m, $x_m^n = x_{m-1}$. The set of elements which make up this sequence will be denoted by $N(x)$. (The finiteness

of G is important here to guarantee the uniqueness of each successive nth root.) Since $N(x)$ has a finite number of elements we may write $N(x) = \{x_0, x_1, \ldots, x_{k-1}\}$ where $x_k = x_t$ for some integer t such that $0 \leqq t < k$. Hence by taking successive nth powers, we have that $x_{k-t} = x_0$. But this implies that $t = 0$, and so $x_k = x$. Thus

$$(x)^{n^k} = x$$

and it follows that the order of x divides $n^k - 1$. We may summarize this in the following result:

PROPOSITION 2. *Let x be an element of a finite nth root group G. Then $N(x)$ has k elements if, and only if, k is the smallest positive integer such that the order of x divides $n^k - 1$.*

The results so far depend on the finiteness of G. We can obtain a different criterion (Proposition 3) for G to be an nth root group in terms of the following two subsets of G: $G_n = \{x^n \mid x \in G\}$ and $E = \{x \in G \mid x^n = e\}$. This criterion permits of extensions to the infinite case. In particular it holds unaltered (Proposition 4) for the infinite abelian case and applies in general with a technical modification (Proposition 5).

PROPOSITION 3. *Let G be a finite group. G is an nth root group if, and only if, G_n is an nth root group and $E \subset G_n$.*

Proof. The "only if" follows immediately since $G = G_n$ when G is an nth root group. To prove the converse, let r be the order of G. Assume that there exists a prime integer p such that p divides both n and r. Then there is a y in G where $y \neq e$ and $y^p = e$. Hence $y^n = e$ and $y \in E$. It follows then that $y \in G_n$ and we have that p divides n and the order of G_n. But by Proposition 1, this contradicts the assumption that G_n is an nth root group. We must conclude that no such prime p exists and that n and r are relatively prime. By Proposition 1, G is an nth root group.

PROPOSITION 4. *Let G be an abelian group. G is an nth root group if, and only if, G_n is an nth root group and $E \subset G_n$.*

Proof. Assume that G_n is an nth root group and $E \subset G_n$. Let $x \in G$. Then $x^n \in G_n$ and there exists $y \in G_n$ such that $x^n = y^n$. Since G is abelian it follows that $(xy^{-1})^n = e$ and so $xy^{-1} \in E$. Therefore there is a $z \in G$ such that $xy^{-1} = z^n$, and we have that $x = z^n y = z^n t^n = (zt)^n$ where $t^n = y$. Hence G is an nth root group. The "only if" is again clear since $G = G_n$ when G is an nth root group.

PROPOSITION 5. *Let G be a group with the following properties*: (1) *G_n is an nth root group*, (2) *$E \subset G_n$ and* (3) *$(xy)^n = x^n y^n$ for all $x, y \in G$. Then G is an nth root group.*

Proof. In the proof of Proposition 4, replace the condition that G is abelian by condition (3).

The converse of Proposition 5 fails to hold. For example, let G be the multiplicative group of nonsingular $m \times m$ matrices over the complex number field. G is a square root group but condition (3) is not true since G is not abelian. In addition, neither conditions (1) and (3) nor (2) and (3) are sufficient for a group to be an nth root group, as is shown by, respectively, the multiplicative group of nonzero reals, and the additive group of integers written multiplicatively.

Although the above results are all accessible to a beginning student of group theory, we must admit that we have been unable to find a proof or a counterexample to Proposition 4 for the infinite nonabelian case. This we leave as an open problem.

Reference

[1] W. R. Utz, Square roots in groups, AMER. MATH. MONTHLY, 60 (1953) 185–186.

A Strange Ultrametric Geometry

GEORGE AKST
New Mexico State University

The main purpose of this paper is to answer a question raised by Dence in [2]. In that article, he defines a certain kind of metric, called an ultrametric, on the rationals Q and questions whether it can be extended to $Q \times Q$, while still maintaining its special properties. We do just that in this article, creating thereby an unusual geometry that is very close to a model of Euclidean geometry in the sense that most of the Euclidean axioms are satisfied. However, it is quite different from the ordinary model of Euclidean geometry in many ways. Thus it shows that when attempting proofs in axiomatic geometries, one must not rely too much on 'common sense', but rather on straightforward deductions form the axioms and previous theorems.

To construct our model, we need to introduce the concept of an ultrametric distance. Recall that a **metric** on S is a function $d: S \times S \to [0, \infty)$ such that for $x, y, z \in S$, (1) $d(x, y) = 0$ iff $x = y$; (2) $d(x, y) = d(y, x)$; and (3) $d(x, z) \leqq d(x, y) + d(y, z)$. Hence, on the real line R, ordinary distance (i.e., $d(a, b) = |a - b|$) is a metric. To get an ultrametric distance, we replace (3) by the condition (3′) $d(x, z) \leqq \max(d(x, y), d(y, z))$. The reader will easily notice that (3′) is a strictly stronger condition than (3) and that the ordinary distance function on R is not an ultrametric.

We now describe a particular ultrametric distance function on the set Q of rational numbers. Given $x \in Q$, $x \neq 0$, write $x = 2^k(a/b)$ where a and b are odd integers. Then the exponent k is uniquely determined. We then define $|x|_2 = (1/2)^k$ and $d(x, y) = |x - y|_2$. If $x = 0$, then $|x|_2 = 0$. For example, if $x = 5/12 = 2^{-2}(5/3)$, then $|x|_2 = (1/2)^{-2} = 4$. The verification that this function satisfies the ultrametric inequality can be found in [1].

(There is nothing special about the number 2 in this definition. Any prime number p will work: if $x = p^k(a/b)$, where a and b are relatively prime to p, then we can define $|x|_p = (1/p)^k$. And even this is a specific case in the general theory of ultrametric spaces. A nice discussion of some of these general abstract ideas can be found in [2].)

Dence raised the question as to whether or not we can non-trivially extend this ultrametric to two dimensions. Let us first examine a few possible ways of extending, all of which fail. If $A = (a_1, a_2)$, $B = (b_1, b_2) \in Q \times Q$, define $D_1(A, B) = d(a_1, b_1) + d(a_2, b_2)$ and $D_2(A, B) = d(a_1, b_1)$ (where d is the ultrametric distance on Q described above). Then both of the above functions fail to satisfy the ultrametric distance inequality; in fact D_2 is not even metric.

A more familiar way to extend a metric is exhibited by $D_3(P_1, P_2) = (d(p_1, p_2)^2 + d(q_1, q_2)^2)^{1/2}$. The reader will recognize that this is the same procedure as is used in extending the ordinary distance function on R to the Cartesian plane. Unfortunately, though D_3 is a metric, it also does not satisfy the ultrametric inequality. To see this, let $A = (1/2, 1/2)$, $B = (0, 1)$, $C = (1/3, 2/3)$; then

$$D_3(A, B) = (d(1/2, 0)^2 + d(1/2, 1)^2)^{\frac{1}{2}} = \sqrt{4+4} = \sqrt{8},$$

$$D_3(B, C) = (d(0, 1/3)^2 + d(1, 1/4)^2)^{\frac{1}{2}} = \sqrt{1+16} = \sqrt{17},$$

$$D_3(A, C) = (d(1/2, 1/3)^2 + d(1/2, 1/4)^2)^{\frac{1}{2}} = \sqrt{4+16} = \sqrt{20}.$$

Therefore, $D_3(A, C) \not\leqq \max(D_3(A, B), D_3(B, C))$.

A function on $Q \times Q$ which does satisfy the ultrametric inequality is given by $D(A, B) = \max(d(a_1, b_1), d(a_2, b_2))$. The function D clearly satisfies (1) and (2). To verify (3′), let $A = (a_1, a_2)$, $B = (b_1, b_2)$, $C = (c_1, c_2) \in Q \times Q$. Then

$$
\begin{aligned}
D(A, C) &= \max(d(a_1, c_1), d(a_2, c_2)) \\
&\leqq \max(\max(d(a_1, b_1), d(b_1, c_1)), \max(d(a_2, b_2), d(b_2, c_2))) \\
&= \max(d(a_1, b_1), d(b_1, c_1), d(a_2, b_2), d(b_2, c_2)) \\
&= \max(\max(d(a_1, b_1), d(a_2, b_2)), \max(d(b_1, c_1), d(b_2, c_2))) \\
&= \max(d(A, B), d(B, C)).
\end{aligned}
$$

The unit circle in this geometry is rather interesting. It intersects the cooordinate axes at all rational points whose numerators and denominators do not contain a factor of 2 when expressed in lowest form. Elsewhere it contains all points of the form $(a/b, c/d)$, in lowest form, where a and c are not both even. Thus the "unit circle" is spread all over the plane and is in fact dense in the ordinary topology of Euclidean 2-space.

In order to talk about a geometric model in the classical sense we must have at least points and lines, concepts that do not depend on the distance function. In our model, we let $Q \times Q$ be the set of points. But then we raise the question: what does a line in $Q \times Q$ look like? A natural idea is to define a line in $Q \times Q$ as the intersection of a line in $R \times R$ (which of course is $\{(x, y): y = mx + b, m, b \in R\}$ or $\{(x, y): x = b, b \in R\}$) with $Q \times Q$. However, the line $y = x + \sqrt{2}$ has no solutions in $Q \times Q$, and so we would not want to consider this as a line. Thus we would have to add to the above definition the qualification "provided this intersection is nonempty." Yet even this is not good enough, since the line $y = \sqrt{2}\,x$ has only one solution in $Q \times Q$; we would not want to consider a single point as a line, since many geometries insist that a line contain at least two points. Thus we are led to the following definition: *A* **line** in $Q \times Q$ is the intersection of a line in $R \times R$ with $Q \times Q$, provided this intersection contains at least two points.

At first, this definition may sound very crude. But it turns out that a line defined in this manner must contain infinitely many points and in fact will be dense in the original line from $R \times R$. To see this, let $(p, q), (r, s) \in Q \times Q$ be two distinct points on the line. If $p = r$, then the equation of the line is given by $x = r$. If $p \neq r$, then the equation of the line is given by $(y - p)/(x - q) = (s - q)/(r - p)$ or

$$
y = \frac{s - q}{r - p}\, x + p - \frac{q(s - q)}{r - p}.
$$

Now that we have a suitable definition of a line, we can talk about parallel lines. We define two lines to be **parallel** if their intersection is void. One of the foremost questions one can ask about any geometry is whether or not it satisfies the parallel postulate; that is, through a given point not on a given line, is there a unique line parallel to the given line? One direction of this postulate is clear for our model. Since the parallel postulate hold in $R \times R$, given a point and a line in $Q \times Q$, we can find a line in $R \times R$, through that point and parallel to the given line. Now, since it is parallel to the given line in $Q \times Q$, it must have rational slope and since it goes through another point with rational coordinates, it must be a line in $Q \times Q$. However, can there be more? In other words, is it possible for an extraordinary line containing p (l_2 in FIGURE 1) to avoid intersecting the given line by passing through it at a hole — that is, at a point with at least one irrational coordinate?

In fact, this cannot happen: *the parallel postulate holds in* $Q \times Q$. To show this we need only show that it is not possible for two lines in $Q \times Q$ with different slopes to intersect at a point not in $Q \times Q$. This follows (except for lines with slope ∞, a case we leave to the reader) from the fact that the intersection of the lines $y = m_1 x + b_1$ and $y = m_2 x + b_2$, where $m_1, m_2, b_1, b_2 \in Q$, $m_1 \neq m_2$, is a point of $Q \times Q$.

Many books on geometry include fallacious proofs in order to emphasize that one must not rely on pictures when attempting axiomatic proofs. The faults in many of these 'proofs' lie in altering the diagram just slightly so that intersections occur in the wrong places. One of the most common of these is the 'proof' that every triangle is isosceles [3]. In light of this false demonstration, it is surprising and

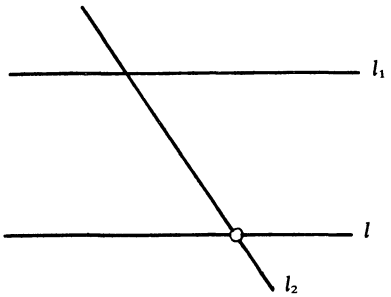


FIGURE 1

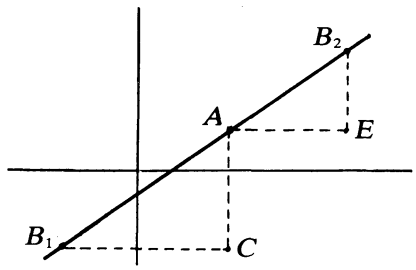


FIGURE 2

interesting that *in the model $Q \times Q$ all triangles are isosceles.* For suppose A, B, C are the vertices of the triangle. If it were not isosceles, then one of the sides, say AB, would be strictly larger than the other two. But then, $D(A, B) \not\leqq \max(D(A, C), D(C, B))$, contradicting the ultrametric inequality.

Many axiom systems for Euclidean geometry contain an axiom to the effect that once a point is fixed on a line and distances are measured from that point, there is a one-to-one correspondence between those distances and the real number line. At the least, one would like that given a point A on a line, there exists a point B on the line such that the distance from A to B is 1. The latter is clearly true in $Q \times Q$ for all lines of slope 0 or ∞, but also holds true for all lines in $Q \times Q$. We of course cannot hope to have a one-to-one correspondence with the real numbers since distances in $Q \times Q$ take on only values of the form $(1/2)^k$. Thus, the following statement is about the strongest one can hope to expect: *Let l be a line and A a point on l. Then for any integer k, there exists a point B on l such that $D(A, B) = (1/2)^k$.*

In proving this, we will consider only lines l having positive slope. The cases of negative, zero, or infinite slope are left to the reader. Suppose A has coordinates (a_1, a_2) and consider the points C: $(a_1, a_2 - 2^k)$ and E: $(a_1 + 2^k, a_2)$. We use d for our ultrametric distance function in Q and e for ordinary Euclidean distance. Let B_1, B_2 be as in FIGURE 2. Notice that $d(A, C) = d(A, E) = (1/2)^k$. We claim that either $d(B_1, C) \leqq (1/2)^k$ or $d(B_2, E) \leqq (1/2)^k$. Once this is shown, since the two-dimensional ultrametric is defined as the maximum length of the legs of the triangle, the proof will be complete.

Suppose that $e(B_2, E) = 2^s(a/b)$, where a and b are odd integers. Since the two triangles, ΔAB_1C and ΔAB_2E, are similar, we have

$$\frac{2^s(a/b)}{2^k} = \frac{2^k}{e(B_1, C)},$$

or $e(B_1, C) = 2^{2k-s}(b/a)$. To prove the claim we must show that $(1/2)^k \geqq \min((1/2)^s, (1/2)^{2k-s})$. If $k \leqq s$, then $(1/2)^k \geqq (1/2)^s$, and we're done. Otherwise, $k > s$, so $k < 2k - s$. Hence, $(1/2)^k > (1/2)^{2k-s}$.

Two concepts common in Euclidean geometry — betweenness and area — appear incapable of definition in our model. The idea of betweenness is very easily expressed via the Euclidean metric. We say of three distinct collinear points A, B, C, that B is between A and C if $e(A, C) = e(A, B) + e(B, C)$. This definition coincides with the intuitive meaning of one point being between two others. But in our model $Q \times Q$ with the ultrametric d, it is easy to show that the corresponding "definition" of betweenness would not be well defined.

A second incomplete feature of this model is its lack of a natural area function. In the most general sense, we would like an area function to be a map from the set of all polygonal regions in the model to $[0, \infty)$ such that if A is the disjoint union of B and C, then $\text{area}(A) = \text{area}(B) + \text{area}(C)$ and if A is congruent to B, then $\text{area}(A) = \text{area}(B)$. If we define the area of a rectangle as the product of the lengths of two consecutive sides, the unit square will have, as expected, area 1. But if we divide the unit square into quarters, then each quarter has area 4. We leave it as an open problem to determine whether or not any area function can be defined using the ultrametric.

The study of models, such as $Q \times Q$, with strange and unusual properties serves several purposes. First, it makes clear that things aren't always as they seem and forces one to stick closer to the deductive reasoning process. Second, it illuminates many interesting ideas such as independence of axioms, consistency, and completeness (see [4], for instance). Finally, it is a tactic that forces one to question certain preconceived ideas that may hinder mathematical growth.

The author would like to thank the referee and the editors for their help and suggestions in improving this paper.

References

[1] George Bachman, Introduction to p-Adic Numbers and Valuation Theory, Academic Press, New York, 1964.
[2] T. P. Dence, Another Euclidean geometry, this MAGAZINE, 47 (1974) 125–132.
[3] D. C. Kay, College Geometry, Holt, Rinehart, and Winston, New York, 1969.
[4] C. R. Wylie, Jr., Foundations of Geometry, McGraw-Hill, New York, 1964.

Lattice Points in Convex Sets

P. R. SCOTT
University of Adelaide

Let K be a closed convex region in the plane, having area $A(K)$. In [2], Minkowski showed that if K is symmetric about the origin O, and $A(K) \geqq 4$, then K contains a non-zero point of the integral lattice. Ehrhart showed in [1] that the same conclusion follows if the center of gravity of K lies at O, and $A(K) \geqq 9/2$. We prove here an extension of Ehrhart's theorem: *If the center of gravity G of K does not lie at O, but the chord of K determined by G and O is bisected by O, and $A(K) \geqq 9/2$, then K contains a non-zero point of the integral lattice.*

To prove this, we let AB be the chord of K which is determined by G and O, and which has O as midpoint. Suppose that AB partitions K into two regions K_1, K_2, and let a, b be support lines to K at A, B respectively. If a, b are not parallel, we assume that they intersect on the K_1-side of AB. Let K_1' denote the reflection of K_1 in O, and set $K^* = K_1 \cup K_1'$. Then by construction K^* is closed, convex and symmetric about O. To complete the proof we use, as did Ehrhart, a result sometimes known as

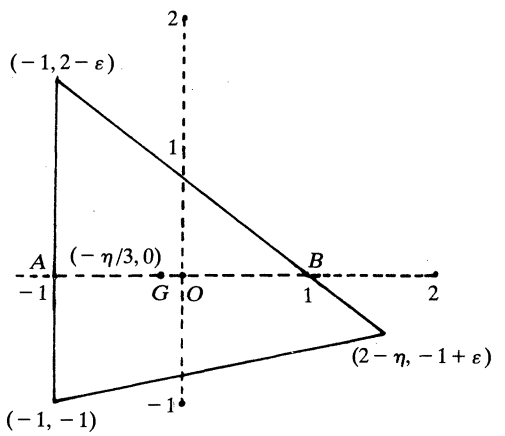


FIGURE 1

'Winternitz' Theorem' (see [3]): If a bounded convex region in the plane is divided into two parts by a line l that passes through its center of gravity, then the ratio of the areas of the two parts always lies between 4/5 and 5/4. Since $A(K) \geqq 9/2$, it follows that $A(K_1) \geqq (4/9) \cdot (9/2)$; thus $A(K^*) \geqq 4$. Now from Minkowski's Theorem, K^* contains a non-zero point of the integral lattice. Using the symmetry of K^* we deduce that K_1, and so K, contains a non-zero point of the integral lattice.

We observe that the proof of this extension can easily be adapted to establish Ehrhart's theorem. For, in the case where G and O coincide, it is sufficient to establish the existence of a chord of K which is bisected by O; this can be achieved by a simple continuity argument.

Finally, we note that our result is, as are those of Minkowski and Ehrhart, best possible. For, let ε be a small positive number, and set $\eta = \varepsilon/(2-\varepsilon)$. Then the triangle (see FIGURE 1) having vertices $(-1, -1)$, $(-1, 2-\varepsilon)$, $(2-\eta, -1+\varepsilon)$ contains no interior lattice points other than O. The center of gravity G of the triangle has coordinates $(-\eta/3, 0)$, and the chord determined by O and G has O as its midpoint. The area of the triangle is $(3-\varepsilon)(3-\eta)/2$; as $\varepsilon \to 0$, this approaches the bound 9/2 of the theorem.

References

[1] E. Ehrhart, Une généralisation du théorème de Minkowski, Comptes Rendus, 240 (1955) 483–485.
[2] H. Minkowski, Geometrie der Zahlen, Leipzig and Berlin, 1896.
[3] I. M. Yaglom and V. G. Boltyanskii, Convex Figures, Holt, Rinehart and Winston, 1961.

Primitive Roots without Quadratic Reciprocity

ALBERT WILANSKY
Lehigh University

Baum [1] has given interesting and useful criteria for certain primitive roots. The purpose of this note is to derive his results without use of quadratic reciprocity. We need prove only his first theorem, since Baum himself proves the second result from the first without making new use of quadratic reciprocity.

THEOREM. *If p and q are odd primes, $p = 2q+1$, then if $q \equiv 1 \pmod 4$, $q+1$ is a primitive root modulo p, while if $q \equiv 3 \pmod 4$, q is a primitive root modulo p.*

Proof. We use the Legendre Symbol $(a \mid b)$ as in [1], and the well-known results which state that:

$$(1) \qquad (2 \mid p) = (-1)^{(p^2-1)/8} \quad \text{and} \quad (-1 \mid p) = (-1)^{(p-1)/2}.$$

First, we assume that $q \equiv 1 \pmod 4$, so $2q+2 \equiv 1 \pmod p$ and

$$(1 \mid p) = (2q+2 \mid p) = (2 \mid p)(q+1 \mid p).$$

Applying (1) we get that $(q+1 \mid p) = -1$. Since $q+1$ is not a quadratic residue, it is a primitive root as in [1].

Next, if $q \equiv 3 \pmod 4$, it follows that $2q \equiv -1 \pmod p$ and $(-1 \mid p) = (2q \mid p)$. Applying (1) we get that $(q \mid p) = -1$ and, as in [1], q is a primitive root.

Reference

[1] J. D. Baum, A note on primitive roots, this MAGAZINE, 38 (1965) 12–14.

Trigonometric Power Series

JOHN STAIB
Drexel University

Most modern calculus texts have little to say about the power series for $\tan x$ (or $\sec x$). A few coefficients are given, the difficulty of obtaining more through the formula $a_n = f^{(n)}(0)/n!$ is made clear, and perhaps a reference is provided. If the student pursues the reference, he will very likely be confronted with a somewhat involved connection between these coefficients and the Bernoulli (or Euler) numbers. The following single path to both the tangent and secant coefficients is more direct and it serves as a good exercise in the method of generating functions.

We begin by letting $f(x) = \sec x + \tan x$, and then observe that $f(x) = \cos x \cdot f'(x)$. Introducing power series for each of these latter functions, we have

$$\sum_0^\infty a_n x^n = \sum_0^\infty b_n x^n \cdot \sum_0^\infty c_n x^n.$$

Identifying coefficients of like terms, and noting that $b_n = 0$ when n is odd, we obtain

$$a_n = \sum_{j=0}^{[n/2]} b_{2j} c_{n-2j}.$$

Replacing b_{2j} by its known value, and noting that $c_n = (n+1)a_{n+1}$, we arrive at

$$a_n = \sum_{j=0}^{[n/2]} \frac{(-1)^j}{(2j)!} (n+1-2j) a_{n+1-2j}.$$

Although this equation can be used to generate the a_k, a better recursion formula is possible for the related numbers A_k, where $A_k = (k!)a_k$. These A_k turn out to be positive integers. To see this we make the appropriate substitutions and then multiply through by $n!$ to obtain

$$A_n = \sum_{j=0}^{[n/2]} (-1)^j \binom{n}{2j} A_{n+1-2j}.$$

Finally, we solve for A_{n+1}. For $n = 0$ and 1, we obtain $A_1 = A_0$ and $A_2 = A_1$. For $n > 1$, we have

$$A_{n+1} = A_n + \binom{n}{2} A_{n-1} - \binom{n}{4} A_{n-3} + \cdots,$$

where the sum ends either with A_1 or A_2. Then, beginning with $A_0 = (0!)a_0 = f(0) = 1$, we may generate the A_k:

$$1, 1, 1, 2, 5, 16, 61, 272, 1385, \ldots.$$

It follows that

$$\sec x + \tan x = 1 + x + \frac{1}{2!} x^2 + \frac{2}{3!} x^3 + \frac{5}{4!} x^4 + \frac{16}{5!} x^4 + \cdots.$$

But $\sec x$ is even and $\tan x$ is odd. Therefore,

$$\sec x = 1 + \frac{1}{2!} x^2 + \frac{5}{4!} x^4 + \frac{61}{6!} x^6 + \frac{1385}{8!} x^8 + \cdots \quad \text{and}$$

$$\tan x = x + \frac{2}{3!} x^3 + \frac{16}{5!} x^5 + \frac{272}{7!} x^7 + \cdots.$$

The technique used above can be exploited to produce a variety of similar exercises for the student in advanced calculus. One has only to seize on another of the many identities involving $\tan x$ or $\sec x$. For example, the identity $\tan x \cos x = \sin x$ leads to

$$\sum_{k=0}^{n} (-1)^k \binom{2n+1}{2k+1} A_{2k+1} = 1,$$

while $\sec x \cos x = 1$ leads to

$$\sum_{k=0}^{n} (-1)^k \binom{2n}{2k} A_{2k} = 0 \quad \text{for} \quad n > 0.$$

However, it should be noted that these formulas lead separately to the coefficients of $\tan x$ and $\sec x$. The special beauty of the particular identity used here is that it forcefully demonstrates that the somewhat obscure and hardly related coefficients of $\tan x$ and $\sec x$ are in fact the odd and even subsequences of the same sequence!

For futher insight into this matter, it is instructive to look at $\sin x$ and $\cos x$. Here some connection between their coefficients appears from the very beginning. This connection is ultimately established by the identity $e^{ix} = \cos x + i \sin x$. But what is more pertinent here is that the coefficients of $\sin x$ and $\cos x$ can be found in a manner paralleling that used above for $\tan x$ and $\sec x$.

Taking $f(x) = \cos x + i \sin x$ we note first that $f'(x) = if(x)$. Introduction of $\Sigma_{n=0}^{\infty} a_n x^n$ for $f(x)$ leads to the recursion formula

$$a_{n+1} = \frac{i}{n+1} a_n,$$

and from $a_0 = f(0) = 1$ we get our start. Thus we obtain

$$\cos x + i \sin x = 1 + ix - \frac{1}{2!} x^2 - \frac{i}{3!} x^3 + \cdots,$$

from which the usual series are obtained by matching real and imaginary parts. Note that the relatedness between $\sin x$ and $\cos x$, despite appearances, is not so direct as that between $\tan x$ and $\sec x$. That is, we cannot say that the coefficients of $\sin x$ and $\cos x$ form the odd and even subsequences of the same sequence but only that this is true for $i \sin x$ and $\cos x$.

Finally, the "togetherness" of $\tan x$ and $\sec x$ suggests that the function $(\tan x + \sec x)$ should have an existence of its own independent of its relationship to $\tan x$ and $\sec x$. This turns out to be the case: the numbers A_n give the solution to the so-called zigzag problem. A **zigzag** is a permutation of 1 through n in which the listed numbers successively rise and fall. For example, for $n = 7$, one zigzag is $(7, 1, 4, 2, 6, 3, 5)$. This zigzag starts with a "fall." By replacing each number by its complement relative to $n + 1 = 8$, we obtain $(1, 7, 4, 6, 2, 5, 3)$, which is a zigzag beginning with a "rise." In this way, all zigzags (for a given n) may be paired. Thus, there are just as many zigzags beginning with a rise as with a fall. And, curiously, the count of either class is A_n. The mathematical link between the zigzag count and the coefficients of $(\tan x + \sec x)$ is elegantly established in [**1**, pp. 64–69]; see also [**2**, p. 258]. However, this proof seems not to be based on any intuitive considerations. After all, why should that number which gives half the count of all the zigzags arising from $(1, 2, \ldots, n)$ be the same number as that which gives the nth derivative evaluated at 0 for the function $(\tan x + \sec x)$? Perhaps some reader could suggest why this "ought" to be.

References

[1] Heinrich Dorrie, 100 Great Problems of Elementary Mathematics, Dover, New York, 1965.

[2] L. Comtet, Advanced Combinatorics, D. Reidel, Dordrecht, 1974.

PROBLEMS

DAN EUSTICE, Editor
LEROY F. MEYERS, Associate Editor
The Ohio State University

Proposals

To be considered for publication, solutions should be mailed before December 1, 1976.

978. For $\lambda > 0$, let

$$(1 - x - y + axy)^{-\lambda} = \sum_{m,n=0}^{\infty} c_{m,n}^{(\lambda)} x^m y^n.$$

Show that $c_{m,n}^{(\lambda)} \geqq 0$ for all m, n if and only if $a \leqq 1$. [*L. Carlitz, Duke University.*]

979. Define $P(m, n)$ to be the number of permutations of the first n natural numbers for which m is the first number whose position is left unchanged. Clearly $P(1, n) = (n-1)!$ for all n. Show that, for $m = 1, 2, \ldots, n-1$,

$$P(m+1, n) = P(m, n) - P(m, n-1).$$

[*Mike Chamberlain and John Hawkins, University of Santa Clara.*]

980. Show that in a perspective drawing of a straight railroad track which is at right angles to the image plane the reciprocals of the images of the ties form an arithmetic progression. [*Peter Ungar, New York University.*]

981. Show that if a smooth curve in R^3 has the property that each principal normal line passes through a fixed point, then the curve must be an arc of a circle. [*Steven Jordan, University of Illinois at Chicago Circle.*]

982. Let $f(n)$ be the sum of all the positive divisors of the positive integer n, including 1 and n. Let A denote the set of all rational numbers of the form $f(n+1)/f(n)$. Determine the closure of A in the set of real numbers. [*Roy DeMeo, Jr., Franklin Square, New York.*]

983. Are there arbitrarily long sequences of consecutive integers no two of which have the same number of prime divisors? [*Bernardo Recáman, Bogotá, Colombia.*]

ASSISTANT EDITORS: DON BONAR, *Denison University*; WILLIAM A. MCWORTER, JR., *The Ohio State University. We invite readers to submit problems believed to be new. Proposals should be accompanied by solutions, when available, and by any information that will assist the editors. Solutions to published problems should be submitted on separate, signed sheets. An asterisk (*) will be placed by a problem to indicate that the proposer did not supply a solution. A problem submitted as a Quickie should be one that has an unexpected succinct solution. Readers desiring acknowledgement of their communications should include a self-addressed stamped card. Send all communications to this department to Dan Eustice, The Ohio State University. 231 W. 18th Ave., Columbus, Ohio 43210.*

984. Let $(b_1, b_2, \ldots, b_n)$ be a non-zero element of $\mathbb{R}^n$. For which n, $2 \leqq n \leqq 8$, is it true that one can choose an orthogonal basis for $\mathbb{R}^n$ from the collection

$$\{(\pm b_{\pi(1)}, \pm b_{\pi(2)}, \ldots, \pm b_{\pi(n)}) : \pi \in P_n\},$$

where P_n is the set of all permutations of $(1, 2, \ldots, n)$? [*Peter Ørno, The Ohio State University.*]

985. Let $Q_k = 1/(k+2)! + 2/(k+3)! + 3/(k+3)! + \cdots$. Show that Q_k is transcendental for all positive integers k, but rational for $k = 0$. [*Jeffrey Shallit, Princeton University.*]

986. Show that there exists a constant c such that $a + b < n + c \ln n$, for all positive integers a, b, and n for which $n!/(a!\,b!)$ is an integer. [*Paul Erdös, Hungarian Academy of Science.*]

987. Let f be differentiable with f' continuous on $[a, b]$. Show that if there is a number c in $(a, b]$ such that $f'(c) = 0$, then we can find a number ξ in (a, b) such that

$$f'(\xi) = \frac{f(\xi) - f(a)}{b - a}.$$

[*Sidney Penner, Bronx Community College.*]

Quickies

Solutions to Quickies appear at the conclusion of the Problems section.

Q635. Prove that for any prime, p, there exists an infinite number of values of m for which p is a divisor of $2^{m+1} + 3^m - 17$. [*Erwin Just, Bronx Community College.*]

Q636. Does there exist a triangle such that the tangents of its angles are of the form x, $1 + x$, and $1 - x$? [*Richard L. Francis, Southeast Missouri State University.*]

Q637. Bisect a line segment with a straight edge given only a line parallel to it. [*Bertram Ross, University of New Haven.*]

Solutions

A Popular Inequality **May 1975**

937. Which is greater: e^π or $(e^e \cdot \pi^e \cdot \pi^\pi)^{1/3}$? [*Norman Schaumberger, Bronx Community College.*]

Solution: Let $k = e^e \pi^e \pi^\pi / e^{3\pi}$. If $k > 1$, then $e^e \pi^e \pi^\pi > e^{3\pi}$. Replace π with x to define

$$F(x) = \frac{e^e x^e x^x}{e^{3x}} \quad \text{and} \quad n(x) = \log_e F(x).$$

Differentiating with respect to x, we find:

$$n'(x) = \frac{e}{x} + \log_e x - 2,$$

$$n''(x) = \frac{1}{x} - \frac{e}{x^2}.$$

Letting $x = e$ and substituting into the equations, we obtain:

$$n(e) = 0, \qquad n'(e) = 0, \qquad n''(e) = 0.$$

Observe that when $x > e$,

$$n''(x) = \frac{1}{x}\left(1 - \frac{e}{x}\right) > 0, \text{ because } \frac{1}{x} > 0 \text{ and } 1 - \frac{e}{x} > 0.$$

Thus, when $x > e$, $n'(x)$ is increasing. Since $n'(e) = 0$, if $x > e$, then $n'(x)$ is always positive. Therefore, if $x > e$, $n(x)$ is increasing. Thus, since $\pi > e$, $n(\pi) > n(e)$. However, $n(e) = 0$, so $n(\pi) > 0$ and $(e^e \cdot \pi^e \cdot \pi^\pi)^{1/3} > e^\pi$.

ROBERT SCHERRER, Student
St. Louis University High School
St. Louis, Missouri

Also solved by M. Ahuja, Geoffrey Akst, John T. Annulis & Dale Burton, Merrill Barnebey, Stephen Baron, Tim Cherney, Raphael T. Coffman, Romae J. Cormier, Winston Crawley & Ann Doris & Ron Wilson, Clayton W. Dodge, Ragnar Dybvik (Norway), Thomas Elsner, Donald C. Fuller, Michael Goldberg, Richard A. Groeneveld, Robert M. Hashway, Richard Johnsonbaugh, Ralph Jones, Steven Kahan, Lew Kowarski, Sidney Kravitz, Gerhard Metzen (Canada), Henrietta O. Midonick, C. C. Oursler, C. B. A. Peck, Charles F. Pinzka, R. L. Raymond, Lawrence A. Ringenberg, Rose-Hulman Problems Group, Daniel Mark Rosenblum, D. K. Ross (Australia), Jeffrey Shallit, Joseph Silverman, J. M. Stark, Temple University Problem Solving Group, University of Santa Clara Problem Solving Group, T. C. Wales, Edward T. H. Wang (Canada), Kenneth M. Wilke, Ken Yocom, and the proposer. Several solvers used the ubiquitous hand calculators.

A Convergent Example **May 1975**

938. Let Σa_n be an infinite series, and set $s_n = a_1 + \cdots + a_n$. A familiar theorem of Abel says that if the a_n are positive and Σa_n diverges, then $\Sigma(a_n/s_n)$ also diverges. If we allow arbitrary signs, can we make Σa_n diverge to $+\infty$ while $\Sigma(a_n/s_n)$ converges? [*S. C. Geller and W. C. Waterhouse, Cornell University.*]

Solution: The answer is yes. As an example, consider the series with $a_{2k-1} = k^{1/2}$ and $a_{2k} = 1 - (k+1)^{1/2}$. Since $s_{2k-1} = k$ and $s_{2k} = k + 1 - (k+1)^{1/2}$ the series diverges to $+\infty$. Then, since $a_1/s_1 = 1$, $a_{2k}/s_{2k} = -(k+1)^{-1/2}$, and $a_{2k+1}/s_{2k+1} = (k+1)^{-1/2}$, the series $\Sigma(a_n/s_n)$ converges to 1.

M. T. BIRD
San Jose, California

Also solved by Larry Bennett & Ken Yocom, Stephen C. Currier, Donald C. Fuller, J. M. Stark, and the proposers.

Counting in Cubes **May 1975**

939. Consider an $n \times n \times n$ cube consisting of n^3 unit cubes. Using only the unit cubes, determine, in terms of n: (1) the number of possible sizes of rectangular parallelepipeds "imbedded" in the cube, (2) the number of cubes of all sizes "imbedded" in the cube, and (3) the number of rectangular parallelepipeds of all sizes "imbedded" in the cube. [*Richard A. Gibbs, Fort Lewis College.*]

Solution:

(1) (a) There are $\binom{n}{1}$ distinct kinds of cubes.

(b) There are $2\binom{n}{2}$ distinct kinds of non-cubic parallelepipeds with square bases; the factor 2 arises because the longer dimension may be either a side of the base or the height of the parallelepiped.

(c) There are $\binom{n}{3}$ distinct kinds of parallelepipeds with all three dimensions different.

Therefore, the required number is $\binom{n}{1}+2\binom{n}{2}+\binom{n}{3}$ or $(n+2)(n+1)n/6$.

(2) There are n^3 cubes of side 1, $(n-1)^3$ cubes of side 2, ..., down to one cube of side n. So the total number of cubes is $\Sigma_{i=1}^{n} i^3 = (n+1)^2n^2/4$.

(3) A rectangular parallelepiped is defined by inserting two planes in each of the three orthogonal sides of the $n \times n \times n$ cube. There are $n+1$ places in each side where a plane can be inserted. Therefore, the total number of imbedded rectangular parallelepipeds is $\binom{n+1}{2}^3 = (n+1)^3n^3/8$.

JULIUS VOGEL
Boston, Massachusetts

Also solved by J. C. Binz (Switzerland), D. P. Choudhury (India), Thomas Elsner, Michael Goldberg, M. G. Greening (Australia), Kent Harris, Philip Haverstick, Clarence R. Perisho, Temple University Problem Solving Group, Gillian W. Valk, and the proposer.

Time Repeats Itself **May 1975**

940. To elaborate on an old problem of Dudeney [1], let us suppose a clock has minute and hour hands of the same length and indistinguishable. [Ignore the fact that on all clocks the hour hand is the one nearer the clock face.] Of the set of all instants in a 12-hour period, consider the partition:

A = set of all instants when the clock reading would be ambiguous;
B = set of all instants when the reading would not be ambiguous.
Which, if either, of these sets is finite?

[1] H. E. Dudeney, 536 Puzzles and Curious Problems, edited by M. Gardner, Charles Scribner's Sons, New York, 1967, p. 14.

[*Edwin P. McCravy, Midlands Technical Education Center.*]

Solution: Let φ = the angle the minute hand makes with 12 o'clock, and let ψ = the angle the hour hand makes with 12 o'clock. Then

$$\frac{\varphi}{2\pi} = \frac{\psi - \frac{\pi}{6}m}{\frac{\pi}{6}}, \quad \text{for} \quad m = 0, 1, \ldots, 11.$$

The ambiguous instants occur when minute and hour hands can be interchanged, so that

$$\frac{\psi}{2\pi} = \frac{\varphi - \frac{\pi}{6}n}{\frac{\pi}{6}}, \qquad n = 0, 1, \ldots 11 \quad \text{except} \quad n = m.$$

Thus $\psi = 2\pi(12m+n)/143$ and $\varphi = 24\pi(12m+n)/143 - 2\pi m$, where $n, m = 0, 1, \ldots, 11$, and $n \neq m$ give the position of the hands when the reading would be ambiguous. Therefore A is finite with 132 ambiguous instants.

G. W. VALK
Tucker, Georgia

Also solved by Walter Bluger (Canada), Tim Cherney, D. P. Choudhury (India), Steven R. Conrad, Milton Eisner, Thomas Elsner, Abraham L. Epstein, Michael Goldberg, M. G. Greening (Australia), Robert M. Hashway, Karl Heuer, Vaclav Konecny (Czechoslovakia), Lew Kowarski, Henry S. Lieberman, Carl McCarty, Clarence R. Perisho, Bob Prielipp, Lawrence A. Ringenberg, Rose-Hulman Problems Group, Joseph Silverman, Temple University Problem Solving Group, Julius Vogel, Ken Yocom, and the proposer.

Conrad and Prielipp found several references for this problem, especially Amer. Math. Monthly E106 (1935, 110) and E1571 (1964, 91).

Legendre Polynomial **May 1975**

941. Show that each of the following expressions is equal to the nth Legendre polynomial.

$$\text{(i)}\ \frac{1}{n!}\begin{vmatrix} x & 1 & 0 & 0 & \cdots & 0 \\ 1 & 3x & 2 & 0 & \cdots & 0 \\ 0 & 2 & 5x & 3 & \cdots & 0 \\ 0 & 0 & 3 & & & \vdots \\ \vdots & \vdots & \vdots & & & n-1 \\ 0 & 0 & 0 & \cdots & n-1 & (2n-1)x \end{vmatrix} \qquad \text{(ii)}\ \frac{1}{n!}\begin{vmatrix} x & 1 & 0 & 0 & \cdots & 0 \\ 1 & 3x & 1 & 0 & \cdots & 0 \\ 0 & 4 & 5x & 1 & \cdots & 0 \\ 0 & 0 & 9 & & & \vdots \\ \vdots & \vdots & \vdots & & & 1 \\ 0 & 0 & 0 & \cdots & (n-1)^2 & (2n-1)x \end{vmatrix}.$$

[*Stanley Rabinowitz, Maynard, Massachusetts.*]

Solution: If both of the determinants are given the same name, $D_n(x)$, and are expanded by their last columns and one of the two determinants obtained in each case again expanded by the last rows, then the following equation (for both (i) and (ii)) results:

$$nD_n(x) = (2n-1)xD_{n-1}(x) - (n-1)D_{n-2}(x).$$

That $D_n(x)$ is the nth Legendre polynomial now follows from this recursion relation and the fact that $D_1(x)$ and $D_2(x)$ are the first and second degree Legendre polynomials, respectively: x and $(3x^2-1)/2$.

GRAHAM LORD
Université Laval, Québec.

Also solved by A. L. Andrew (Australia), Clyde A. Bridger, D. P. Choudhury (India), Michael Ecker, Donald C. Fuller, Ralph Garfield, M. G. Greening (Australia), Philip Haverstick, Ralph Jones, Vaclav Konecny (Czechoslovakia), Bob Prielipp, A. G. Shannon (Australia), Joseph Silverman, Scott Smith, J. M. Stark, Temple University Problem Solving Group, and the proposer.

A Constant Sum **May 1975**

942. Determine the maximum value of

$$S = \sum_{1 \le i < j \le n} \left(\frac{x_i x_j}{1-x_i} + \frac{x_i x_j}{1-x_j} \right)$$

where $x_i \ge 0$ and $x_1 + x_2 + \cdots + x_n = 1$. [*M. S. Klamkin, University of Waterloo.*]

Solution: We have

$$2S = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{x_i x_j}{1-x_i} + \frac{x_i x_j}{1-x_j}\right) - \sum_{i=1}^{n}\frac{2x_i^2}{1-x_i} = 2\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{x_i x_j}{1-x_i} - \sum_{i=1}^{n}\frac{2x_i^2}{1-x_i}$$

$$= 2\left(\sum_{j=1}^{n} x_j\right)\left(\sum_{i=1}^{n}\frac{x_i}{1-x_i}\right) - 2\sum_{i=1}^{n}\frac{x_i^2}{1-x_i}$$

$$= 2\sum_{i=1}^{n} \frac{x_i - x_i^2}{1 - x_i} = 2\sum_{i=1}^{n} x_i = 2.$$

Thus $S = 1$.

JOSEPH SILVERMAN, Student
Brown University

Also solved by M. Ahuja, J. C. Binz (Switzerland), M. T. Bird, Thomas Elsner, Donald C. Fuller, Mark D. Galit, Ralph Garfield, M. G. Greening (Australia), Richard A. Groeneveld, Robert M. Hashway, Ralph Jones, Vaclav Konecny (Czechoslovakia), Henry S. Lieberman, Graham Lord, J. M. Malone II, Gerhard Metzen (Canada), John D. O'Neill, University of Santa Clara Problem Solving Group, Scott Smith, Temple University Problem Solving Group, Phil Tracy, Ken Yocom, and the proposer.

Answers

Solutions to the Quickies which appear near the beginning of the Problems section.

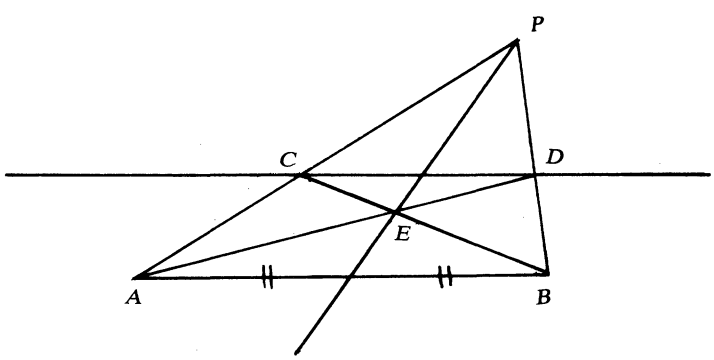
Q635. Fermat's theorem asserts that for any prime p, p will be a divisor of $a^p - a$, and since $a^p - a$ is easily seen to be a divisor of $a^{p^n} - a$, where n is any positive integer, p will also be a divisor of $a^{p^n} - a$. In the given expression replace m by $p^n + 1$ to obtain $2^{p^n+2} + 3^{p^n+1} - 17 = 2^2 \cdot 2^{p^n} + 3 \cdot 3^{p^n} - 17 = 4(2^{p^n} - 2) + 3(3^{p^n} - 3)$. Since the latter expression is divisible by p, it follows that whenever $m = p^{n+1}$, $2^{m+1} + 3^m - 17$ will be divisible by p.

Q636. If A, B, and C are the angles of a triangle, then

$$\tan A + \tan B + \tan C = (\tan A)(\tan B)(\tan C).$$

If $x + (1+x) + (1-x) = x(1+x)(1-x)$, then $x + 2 = x - x^3$, or $x^3 = -2$, or $x = -\sqrt[3]{2}$. As $x < -1$, both x and $(1+x)$ are negative, meaning that the triangle has two obtuse angles. Thus no triangle satisfies the above condition.

Q637. Let the line segment be AB. Take any point P as shown. Draw PA and PB. Then join BC and AD. The intersection of AD and BC is E. The line PE is a bisector of line segment AB. (This result may be found in A. S. Smogorzhevskii, The Ruler in Geometrical Constructions, Blaisdell, Waltham, 1971, p. 52.)



NEWS & LETTERS

RECREATIONAL MATHEMATICS

Miami University in Oxford, Ohio will host a conference on recreational mathematics on Sept. 24-25, 1976. The purpose of the conference is to explore the relationship between recreational mathematics and significant problems, to develop recreational mathematics as a motivating tool for students, and to have fun.

Invited speakers include Leon Bankoff, problems editor of *Pi Mu Epsilon Journal;* Elwyn Berlekamp, expert on games and graphs; Murray Klamkin, originator of many applied and recreational problems; David Klarner, theorist in recreational mathematics; Joseph Madachy, founder and editor of *Journal of Recreational Mathematics;* Robinson Rowe, problem solver in algebra, trigonometry, calculus; C.W. Trigg, leading contributor in recreational mathematics; and Robert Wainwright, editor of *Lifeline,* journal of Conway's game of Life.

Requests for information as well as abstracts of contributed papers should be sent to Prof. Donald O. Koehler, Department of Mathematics, Miami University, Oxford, Ohio 45056.

ERRATUM

It has come to my attention that there was an error in equation (3) of my paper "An Interesting Continued Fraction" (this *Magazine,* January 1975, pp. 207-211). The minus sign separating the two bracketed expressions should be a plus sign, so the line should read:

$$(3)\quad a_n = [\tfrac{1}{2}(b + \sqrt{b^2 - 4})]^n + [\tfrac{1}{2}(b - \sqrt{b^2 - 4})]^n .$$

Jeffrey Shallit, Student
Princeton University
Princeton, N.J. 08540

NEW YORK MATH FAIR

Secondary school students in the eleventh or twelfth grades in the New York area are eligible to participate in the Greater Metropolitan New York Math Fair scheduled for March 6, 1977. The fair consists of talks before a group of judges based on a paper written by the student on some topic of his own choosing.

Further details and application forms may be obtained from

Dr. Theresa J. Barz, Secretary
Math Fair Committee
Department of Mathematics
St. John's University
Jamaica, New York 11439

ODD RECIPROCALS

E.J. Barbeau's inquiry (in "Expressing One as a Sum of Odd Reciprocals", this *Magazine,* January 1976, p. 34) for sets of nine or eleven odd integers whose reciprocals sum to one produced several responses whose methods ranged from pure thought to brute force computation. Five sets of nine integers were reported (3, 5, 7, 9, 11, 15 plus 21, 135, 10395; or 21, 165, 693; or 21, 231, 315; or 33, 45, 385; or 35, 45, 231) as were two sets of eleven (3, 5, 7, 9, 11, 15, 27, 105, 135, 945, 10395 and 3, 5, 7, 9, 15, 21, 25, 35, 45, 125, 325).

Bruce Leasure (Univ. of Illinois, Urbana) and Allan Wm. Johnson, Jr. (Defense Comm. Agency, Washington, D.C.) report that a complete search of all possibilities confirms these five as the only sets of nine odd integers whose reciprocals sum to one. The number of sets of eleven is apparently not known.

Allan Johnson also confirmed, in response to another of Barbeau's inquiries, that one can indeed be expressed as the sum of reciprocals of different positive integers each of which is the product of exactly two primes. He found sets containing as few as 50 integers and showed by direct computation that at least 38 are required. The question of determining the fewest integers required by this type of representation remains open.

The Editors.

SIMPLE, SIMPLER, ...

The article "Hex Must Have a Winner --An Inductive Proof" (this *Magazine*, March 1976, pp. 85-86) reminded me of a topological proof which is perhaps somewhat shorter.

To fix our ideas, suppose black wishes to connect the left and right sides of the board, and white wishes to connect the top and bottom. Let B be the set of black hexagons connected to the left side. Clearly B either contains a black hexagon on the right side or it does not. If it does, then black has completed a chain. A chain connecting the two sides blocks white, who could not also have a chain without some hexagon being both black and white. But, if B does not contain a hexagon on the right side of the board, the hexagons connected to B on the right are all white, are connected, and form a chain from top to bottom. Hence white has won.

Paul B. Johnson
University of California
Los Angeles
California 90024

I was apalled by the complexity of the 'simple' proof that Hex must have a winner. Here is a much simpler proof that I was told many years ago.

Imagine the playing board for the game of Hex to be made out of paper. Whenever white moves, he colors the hexagon of his choice white. Whenever black moves, he cuts out the hexagon of his choice. Repeat this until no one can move any more.

Pick up the playing board in your hands, holding the two 'white' edges. Pull your hands apart. Either the paper stops you, in which case there msut be a path of white squares and so white wins; or nothing stops you, in which case there is a 'path' of cut out squares between the top and the bottom of the board, and so black wins.

Clearly, one of the two must occur; and so someone must win.

Daniel Zwillinger, student
Mass. Inst. Tech.
Cambridge
Massachusetts 02139

A BICENTENNIAL ALPHAMETIC

USA
FIFTY
UNITED
STATES

Rev. Bernard J. Portz, S.J.
Creighton University
Omaha
Nebraska 68178

Editors Note: Please don't submit solutions. We will publish some answers in the next issue.

A COUNTEREXAMPLE

The conjecture stated at the conclusion of Man Keung Siu's note "When is -1 a power of 2?" (this *Magazine*, November 1975, pp. 284-286) is false; the smallest counterexample is 205.

A more illuminating statement of the title question would be "What is the nature (mod 8) of the divisors of $2^x + 1$?" The following facts are easy to prove and well known:

A) If 4 divides x, all divisors of $2^x + 1$ are congruent to 1 (mod 16).

B) If x is odd, all divisors of $2^x + 1$ are congruent to 1 or 3 (mod 8). (For example, $3 \cdot 281$ divides $2^{35} + 1$.)

C) If $x = 4k + 2$, all divisors of $2^x + 1$ are congruent to 1 (mod 4).

Thus if $2^x + 1 = 0 \pmod{d}$ is solvable, d has its prime divisors from just one of A), B), or C). In particular, d cannot have one prime divisor congruent to 3 (mod 8) and another congruent to 5 (mod 8).

J.L. Selfridge
Northern Illinois Univ.
DeKalb
Illinois 60115

Siu showed that $2x \equiv -1 \pmod{d}$ is solvable if d is a power of a prime p and (a) $p \equiv 1 \pmod 8$ plus additional conditions; or (b) $p \equiv 3 \pmod 8$; or (c) $p \equiv 5 \pmod 8$. He expressed interest in a necessary and sufficient condition for the congruence to be solvable modulo some composite d and conjectured that it is solvable if and only if it is solvable modulo each prime divisor of d and all the prime divisors of d belong to exactly one of (a), (b), or (c).

Selfridge's example (above) of 41 (= 205/5) shows that the conjecture is false and that a prime divisor of type (a) can be combined with other types. Also 697 shows that for type (a) primes, solvability modulo each prime does not imply solvability modulo the product. However, the problem is only with divisors of type (a).

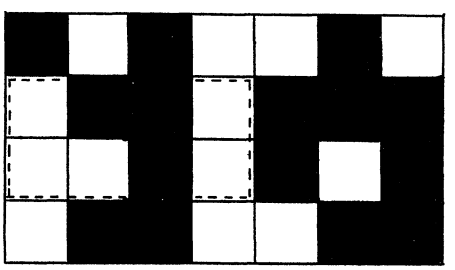
Let $d = p_1^{e_1} \ldots p_r^{e_r}$, r_i be the exponent of 2 mod $p_i^{e_i}$ and $2^{s_k} \| r_k$. $2^x \equiv -1 \pmod{d}$ is solvable if and only if each r_i is even and g.c.d. (r_i, r_s) divides $\frac{1}{2}(r_i - r_j)$; this is true if and only if $s_i = s_j$. For $1 \le i \le j \le n$. This result is an easy consequence of results in elementary number theory and Siu's work. It is useful in constructing examples and in proving, for instance, that if the p_i are restricted to types (b) and (c), then $2^x \equiv -1 \pmod{d}$ is solvable if and only if it is solvable mod each p_i and all the p_i are of the same type.

Joseph B. Dennin, Jr.
University of Maryland
Princess Anne
Maryland 21853

1976 U.S.A. MATHEMATICAL OLYMPIAD

The fifth annual U.S.A. Mathematical Olympiad, which took place on May 4, 1976, consisted of the following five problems. The exam was prepared and directed by Dr. Murray S. Klamkin and Dr. Samuel L. Greitzer.

1. (a) Suppose that each square of a 4 x 7 chessboard, as shown below, is colored either black or white. Prove that with any such coloring, the board must contain a rectangle (formed by the horizontal and verticle lines of the board), such as the one outlined in the figure, whose four distinct corner squares are all of the same color.
(b) Exhibit a black-white coloring of a 4 x 6 board in which the four corner squares of every such rectangle as described above, are not the same color.



2. If A and B are fixed points on a given circle and XY is a variable diameter of the same circle, determine (with proof) the locus of the point of intersection of lines AX and BY. You may assume that AB is not a diameter.

3. Determine (with proof) all integral solutions of $a^2 + b^2 + c^2 = a^2b^2$.

4. If the sum of the lengths of the six edges of a trirectangular tetrahedron $PABC$ (i.e., $\angle APB = \angle BPC = \angle CPA = 90°$) is S, determine (with proof) its maximum volume.

5. If $P(x)$, $Q(x)$, $R(x)$, and $S(x)$ are all polynomials such that

$$P(x^5) + xQ(x^5) + x^2R(x^5) = (x^4 + x^3 + x^2 + x + 1)S(x),$$

prove that $x - 1$ is a factor of $P(x)$.

PICK IFF EULER

It may be of interest to note, in relation to "Triangulations and Pick's Theorem" by R.W. Gaskell, et al. (this *Magazine*, January 1976, pp. 35-37), that, not only does Pick's theorem follow from Euler's formula, but also Euler's formula follows from Pick's theorem. (See DeTemple and Robertson, "The Equivalence of Euler's and Pick's Theorems," *Mathematics Teacher*, 67 (1974) 222-226.) The strategy is this:

1) Observe that a primitive triangle has area 1/2.

2) Show that $V - E + F$ remains unchanged as a (plane) map is triangulated.

3) From $F = T =$ twice area $= 2V_i + V_b - 2$, $2E = 3F + V_b$, and $V = V_i + V_b$ obtain $V - E + F = 1$ for the triangulated map.

Finally, it is worth noting that *both* Pick's theorem and Euler's formula follow from the basic formula $T = 2V_i + V_b - 2$. This result can be established without the use of Euler or Pick by observing that $\pi T = 2\pi V_i + \pi(V_b - 2)$. From this it follows, using the elegant argument of Gaskell, et al., that the area of a primitive triangle is 1/2. With these two facts in hand, both Pick's theorem and Euler's formula can be obtained.

Richard A. Gibbs
Fort Lewis College
Durango, Colorado 81301

1976 NSF-CBMS REGIONAL RESEARCH CONFERENCES

The National Science Foundation has granted through the Conference Board of the Mathematical Sciences eight Regional Research Conferences for college and university mathematics teachers for the summer of 1976. Two more conferences are expected to be funded, but they had not been officially granted at the time this issue went to press.

Date	Host Institution	Subject	Lecturer
May 31-June 4	Univ. of Nebraska	Tansference Methods in Harmonic Analysis	Guido Weiss
May 31-June 4	Univ. of Colorado	Topological Methods in Differential Equations	Charles Conley
June 28-July 2	Univ. of Houston	Nonlinear Diffusion	Donald G. Aronson
June 28-July 2	St. Olaf College	Mathematics of Optimal Facility Locations	Alan J. Goldman
July 5-9	Univ. of Pittsburgh	Complex Manifold Techniques in Relativity	Roger Penrose
July 8-12	Kent State Univ.	Banach Spaces of Smooth Functions	Aleksander Pelczynski
July 26-30	Illinois State Univ.	Diophantine Approximations	Wolfgang Schmidt
August 2-6	Old Dominion Univ.	Computational Fluid Dynamics and Turbulence Theory	Steven A. Orszag

Approximately 25 mathematicians can attend each conference; travel and subsistence allowances are paid by NSF. Inquiries about particular conferences and requests for application forms should be addressed to the departments of mathematics at the host institutions.

CBMS anticipates that the Regional Conferences project will continue in 1977 and that for these the deadline for receipt of proposals by NSF will be November 15, 1976, instead of December 1 as it has been in prior years.

MAA STUDIES IN MATHEMATICS

This series is intended to bring to the mathematical community expository articles at the collegiate and graduate level on recent developments in mathematics.

These numbers are currently available:

1. *Studies in Modern Analysis,* edited by R. C. Buck.
2. *Studies in Modern Algebra,* edited by A. A. Albert.
3. *Studies in Real and Complex Analysis,* edited by I. I. Hirschman, Jr.
4. *Studies in Global Geometry and Analysis,* edited by S. S. Chern.
5. *Studies in Modern Topology,* edited by P. J. Hilton.
6. *Studies in Number Theory,* edited by W. J. LeVeque.
7. *Studies in Applied Mathematics,* edited by A. H. Taub.
8. *Studies in Model Theory,* edited by M. D. Morley.
9. *Studies in Algebraic Logic,* edited by Aubert Daigneault.
10. *Studies in Optimization,* edited by G. B. Dantzig and B. C. Eaves.
11. *Studies in Graph Theory, Part I,* edited by D. R. Fulkerson.
12. *Studies in Graph Theory, Part II,* edited by D. R. Fulkerson.

One copy of each volume in this series may be purchased by individual members of the Association for $5.00 each; additional copies and copies for nonmembers are priced at $10.00 each. Special price for the two-volume set, Volumes 11 and 12: $9.00; for nonmembers the price is $18.00. (Orders for under $10.00 must be accompanied by payment. Prepaid orders will be delivered postage and handling free.)

Orders should be sent to:

MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, D.C. 20036

THE MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, DC 20036

MATHEMATICS MAGAZINE VOL. 49, NO. 3, MAY 1976