MATHEMATICS MAGAZINE

CONTENTS

CODEN: MAMGA8

VOLUME 48 • SEPTEMBER 1975 • NUMBER 4

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

Address Change Requested

MATHEMATICS MAGAZINE

EDITORIAL CORRESPONDENCE should be sent to the EDITORS-ELECT, J. ARTHUR SEEBACH, JR., and LYNN A. STEEN, Department of Mathematics, St. Olaf College, Northfield, Minnesota 55057. Articles should be typewritten and triple-spaced on $8\frac{1}{2}$ by 11 paper. The greatest possible care should be taken in preparing the manuscript. Authors should submit two copies and keep one copy for protection against possible loss of the manuscript. Figures should be drawn on separate sheets in India ink. They should be appropriately lettered and of a suitable size for photographing.

REPRINT PERMISSION should be requested from LEONARD GILLMAN, Mathematical Association of America, University of Texas, Austin, Texas 78712. (See also the copyright notice below.)

NOTICE OF CHANGE OF ADDRESS and other subscription correspondence should be sent to the Executive Director, A. B. WILLCOX, 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 (University of Buffalo), Buffalo, New York 14214.

THE MATHEMATICS MAGAZINE is published by the Mathematical Association of America at Washington, D.C., bi-monthly except July-August. Ordinary subscriptions are $10 per year. Members of the Mathematical Association of America and of Mu Alpha Theta may subscribe at the special rate of $7.00. Back issues may be purchased, when in print, for $2.00.

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

SEVEN GAME SERIES IN SPORTS

RICHARD A. GROENEVELD and GLEN MEEDEN, Iowa State University

1. Introduction. We consider finding a satisfactory probability model for the number of games played in a "best of seven" series. Such series are used in professional baseball, hockey and basketball to determine the winner of a year's competition. The first team to win four games is, of course, declared the series winner.

It is hoped that this example of probability modeling would be helpful for teachers of introductory probability and statistics courses, as an example of an application in the area of sports.

2. Model I. One model which has often been assumed (see [3] and [7]) considers such a series to be the outcome of independent trials between a stronger team, which wins each game with probability $p \geqq 1/2$ and a weaker team with probability $q = 1-p$ of winning any game. Let X be a random variable representing the number of games played in such a series. These assumptions imply that the probability function for X is:

$$f(4) = p^4 + q^4, \quad f(5) = 4pq(p^3 + q^3),$$
$$f(6) = 10p^2q^2(p^2+q^2), \quad \text{and} \quad f(7) = 20p^3q^3. \tag{1}$$

Some intuitively satisfying results follow from these expressions. As pq has a maximum at $p = 1/2$ on $[1/2, 1]$, we see that the maximum value of $f(7)$, i.e. the probability of a seven game series, is 5/16 and occurs when the teams are presumed of equal ability. Similarly $p^4 + q^4$ has its minimum of 1/8 on $[1/2, 1]$ at $p = 1/2$ and maximum of 1 at $p = 1$. Thus a four game series is least likely if the teams are evenly matched and is certain if $p = 1$. Finally, it is straightforward to show that

$$E(X) = 4 + 4pq(1 + 2p + 3p^2 - 10p^3 + 5p^4) = 4 + 4pq(1 + 2pq + 5p^2q^2),$$

so that $E(X)$ has its maximum value of 5.8125 at $p = 1/2$. In words, the more evenly balanced the teams, the longer the series.

To determine whether this one parameter model provides an adequate description of the length of seven game series, we first considered data from the National Hockey League playoffs over the years 1939–1967. This period was chosen for hockey because there was a single stable league over this period. Between 1943–1967 the first and third teams in the final standings and the second and fourth teams played seven game series with the winners playing a seven game series for the Stanley Cup. (From 1939–42 there was one seven game playoff series and the seven game Stanley Cup series). The lengths of 83 series played over this period are given in Table 1. A Chi-square goodness of fit test (see [1]) has been carried out for the model given by I. The row of expected values has been found by numerically computing the maximum likelihood estimate of p, based on this multinomial model and the frequencies observed. This estimate, $\hat{p} = 0.654$ has been used in (1) again, to estimate the probabilities of 4, 5, 6 and 7 game series, with multiplication by 83 yielding the expected numbers. The Chi-square value of 0.360 is exceeded with probability approximately 0.84 in the Chi-square distribution with 2 degrees of Freedom, indicating consistency with the model proposed.

TABLE 1 — Hockey Series

Games Played	*4*	*5*	*6*	*7*
Frequency	15	26	24	18
Expected Number	16.351	24.153	23.240	19.256
$\chi^2 = 0.360$	$\hat{p} = 0.654$	$\bar{x} = 5.542$		

Source: The Encyclopedia of Hockey.

TABLE 2 — World Series (Baseball)

Games Played	*4*	*5*	*6*	*7*
Frequency	12	16	13	25
Expected Number	9.524	17.312	20.084	19.081
$\chi^2 = 5.97$	$\hat{p} = 0.580$	$\bar{x} = 5.772$		

Source: The Sports Encyclopedia: Baseball.

For baseball we have considered the data on the world series between 1903–1973. This is summarized in Table 2. There were 66 seven game series over this period (4 series were the best of 9 games and there was no world series in 1904). The expected numbers have been found using the method of maximum likelihood as before. Here the value of $\chi^2 = 5.97$ would only be exceeded with probability 0.051 in the Chi-square distribution with 2 degrees of freedom, suggesting that model I does not fit baseball data very well. A glance at Table 2 indicates that this model overestimates the number of 6 game series and underestimates the number of 7 game series.

In [3] Mosteller gives a detailed analysis of world series competition. Both the independent trial model and the maximum likelihood estimate at p are given in this paper along with much additional interesting discussion. Mosteller based his analysis on 44 best of seven game series and found that the independent trials model fit quite well. Since then 22 additional series have been played, 14 of which have lasted seven games, while only 2 have stopped at six. Hence, as we have just observed, the independent trial model no longer seems appropriate.

Examining the independent trials model more carefully we observe that

$$\begin{aligned} f(6)+f(7) &= 10p^2q^2(p^2+q^2)+20p^3q^3 \\ &= 10p^2q^2(p^2+2pq+q^2) = 10p^2q^2. \end{aligned}$$

Thus the conditional probability a series lasts exactly 6 games, given it lasts more than 5, is:

$$P(X=6 \mid X>5) = f(6)/10p^2q^2 = p^2+q^2.$$

However, $p^2 + q^2 = 1 - 2(1 - p) \geqq 1/2$. For the baseball data we would estimate this probability as $13/38 = 0.342$, which is inconsistent with the model. In terms of a statistical test, if we test the null hypothesis H_0: $P(X = 6 \mid X > 5) \geqq 1/2$ versus H_1: $P(X = 6 \mid X > 5) < 1/2$ we would compute the value of the test statistics

$$Z = (13 - 38(1/2))/\sqrt{38/4} = -1.947,$$

using the normal approximation to the binomial distribution. Referring to the standard normal distribution we would reject H_0, implicit in model I, at the commonly used 5 percent level of significance. On the other hand, the hockey data of Table 1 is consistent with model I here also. The conditional probability would be estimated as $24/42 = 0.57$ which exceeds 0.5 and for hockey we have $\hat{p}^2 + \hat{q}^2 = 0.55$ (approximately 0.57), where $\hat{q} = 1 - \hat{p}$.

3. Model II. In view of the apparent critical role of the sixth game, we propose here a two parameter model to describe the length of seven game series. The probability function is assumed to have the form:

$$f(4) = p^4 + q^4 \qquad f(5) = 4pq(p^3 + q^3)$$
$$(2) \qquad f(6) = 10p^2q^2\lambda \qquad f(7) = 10p^2q^2(1 - \lambda).$$

The first 5 games are considered to be the result of independent trials as before. The parameter λ represents the conditional probability, $P(X = 6 \mid X > 5)$ which is now constrained only to the interval [0, 1]. In the seventh game we assume the stronger team resumes probability p of winning. Using the method of maximum likelihood to estimate p and λ, we find for hockey $\hat{\lambda} = 0.571$ and $\hat{p} = 0.651$. For baseball the corresponding values are $\hat{\lambda} = 0.342$ and $\hat{p} = 0.614$. Using these estimates in (2) we can again estimate the probabilities of 4, 5, 6 and 7 game series and find the expected number of series of each length assuming model II. These results are displayed in Table 3.

TABLE 3 — Observed and Predicted Length of Series for Model II

Games	*4*	*5*	*6*	*7*
Frequency (Hockey)	15	26	24	18
Expected (Hockey)	16.14	24.02	24.46	18.38
Frequency (Baseball)	12	16	13	25
Expected (Baseball)	10.85	18.08	12.68	24.39

We have again computed the value of the Chi-square test statistic for model II. The values are $\chi^2 = 0.2603$ for hockey and $\chi^2 = 0.3844$ for baseball. Neither value exceeds the 50th percentile of the χ^2 distribution with one degree of freedom (0.455), so that this model appears to provide a satisfactory fit to both the hockey and baseball data. In order to examine the assumption of independence in the first five games in this model we observe that this implies that, in five game series, the losing team has equal chance of obtaining its only victory at games 1, 2, 3 or 4. Data giving the winning game for the series loser in the 26 five game hockey series and 16 five game baseball series appears below:

	Hockey				Baseball			
Game	1	2	3	4	1	2	3	4
Frequency	3	10	6	7	4	5	5	2
Expected	6.5	6.5	6.5	6.5	4	4	4	4
	$\chi^2 = 3.85$				$\chi^2 = 1.5$			

Chi-square values have been computed with the expected values found under the assumption of equal probability of victory occurring in any of the first four games. The appropriate number of degrees of freedom is now equal to 3. Neither computed value is unusual (neither is significant at level 0.25), so that the independence assumption in the first 5 games cannot be refuted.

4. The meaning of winning. The determination of a winning team by a seven game series is not, of course, the same as determining which team is stronger. The weaker team may win. Under model I the probability the stronger team wins the series is given by

$$g_{\mathrm{I}}(p) = p^4 + 4p^4q + 10p^4q^2 + 20p^4q^3.$$

Under model II, the probability that the stronger team wins is now given by

$$\begin{aligned} g_{\mathrm{II}}(p) &= p^4 + 4p^4q + 10p^3q^2\lambda + 10p^2q^2(1-\lambda)p \\ &= p^4 + 4p^4q + 10p^3q^2. \end{aligned}$$

Also

$$\begin{aligned} g_{\mathrm{I}}(p) - g_{\mathrm{II}}(p) &= 10p^3q^2(p + 2pq - 1) \\ &= 10p^3q^3(2p-1) \geqq 0, \quad \text{as} \quad p \geqq 1/2. \end{aligned}$$

Hence model II will (except for $p = 1/2$ or 1) yield a smaller probability of the stronger team winning than model I. Table 4 gives the values of these two functions for several values of p. As this table shows this reduction is never very large.

In the case of hockey, the final standings provided a measure of the stronger team in each series. In 59 of the 83 series, the team finishing higher in the standings won the series (71 percent of the time). This compares with estimates of 80 percent and 76.5 percent using models I and II respectively (with p estimated as before). This again represents some evidence in favor of model II. Along these lines, both models suggest that the weaker team will win a series in 4 straight games with probability q^4. For baseball we would expect $66(0.386)^4 = 1.5$ such series. Our candidate for this happening is the 1914 world series won by the Boston Braves, who came from last place in the National League on July 18 to beat the Philadelphia Athletics in the series in 4 straight games. The Athletics had won 2 of the previous 3 world series and were highly favored.

TABLE 4 — The Probability the Stronger Team Wins

p	$g_I(p)$	$g_{II}(p)$
.5	.500	.500
.55	.618	.593
.60	.710	.683
.65	.800	.765
.70	.874	.837
.75	.929	.897
.80	.967	.942
.85	.988	.973
.90	.997	.991
.95	.9998	.9988
1.00	1.000	1.000

5. Discussion. We have presented a two parameter model which provides a good statistical fit to seven game series in hockey and baseball. The interpretation of the parameter p is clear, — it is the probability that the stronger team will win games 1 through 5 or game 7. The parameter λ represents the conditional probability the series ends in six games given it lasts more than five. Apparently this parameter is very much smaller for baseball than for hockey. One possible explanation of this difference is that the team which has won 2 of the first 5 games can temporarily strengthen itself in game 6 by using its best pitcher. The team which is ahead can afford to save its best pitcher for the last game, if it is played. In hockey, on the other hand, there is no possibility of a similar short term advangtage. Basically the same team plays each game. The parameter λ appears associated with the sport being played.

It is of interest to consider how seven game series in orther sports compare with those in hockey and baseball. The National Basketball Association and the American Basketball Association have had 118 seven game series during the period 1950–1973 in playoffs and finals. The lengths of these series have been:

Games	4	5	6	7
Frequency	13	34	29	42

Source: The Modern Encyclopedia of Basketball and The Official Associated Press Sports Almanac, 1974.

In this case $\chi^2 = 0.966$ indicating a satisfactory fit for model II. Using maximum likelihood estimation in model II we find $\hat{p} = 0.533$ and $\hat{\lambda} = 0.408$. This suggests that basketball series have been played between comparatively well-balanced teams. Additionally the value of $\hat{\lambda} = 0.408$ indicates that the conditional probability of a basketball series ending in 6 games, given it lasts more than 5 games, is less than 0.5. In this respect basketball seems similar to baseball and different from hockey. Superficially, the sports of basketball and hockey would seem more likely to produce similar results in seven game series. Both basketball and hockey are played with essentially the same lineup in each game. A possible explanation of the similar sixth game phenomenon in baseball and basketball is that these are more "individualistic" sports than hockey. Hence the strongest players of the team behind after 5 games may make a major effort in game 6 which cannot be sustained for game 7, while the players on the team ahead after 5 games individually husband resources in game 6. Perhaps readers will have other explanations.

References

1. H. Cramer, Mathematical Methods of Statistics, Princeton University Press, 1946, pp. 416–452.
2. Z. Hollander, Editor, The Modern Encyclopedia of Basketball, Four Winds Press, New York, 1973.
3. F. Mosteller, The World Series Competition, J. Amer. Statist. Assoc., 47 (1952) 355–380.
4. D. S. Neft, R. T. Johnson, R. M. Cohen, and J. A. Deutsch, The Sports Encyclopedia: Baseball, Grosset and Dunlap, New York, 1974.
5. R. A. Styer, The Encylopedia of Hockey, rev. ed., A. S. Barnes, New York, 1973.
6. The Official Associated Press Sports Almanac, Dell, New york, 1974.
7. M. Zelinka, How Many Games to Complete a World Series, in Statistics by Example: Weighing Chances, ed. by F. Mosteller, et al., Addison-Wesley, Reading, 1973.

CLOSEST UNITARY, ORTHOGONAL AND HERMITIAN OPERATORS TO A GIVEN OPERATOR

JOSEPH B. KELLER, Courant Institute of Mathematical Sciences, New York University

1. Introduction. It is often of interest to find an operator U_0, in some specified class of operators $\mathcal{U}$, which is closest to a given operator A. We shall describe various situations in which this kind of problem arises, and then show how to solve it for several choices of $\mathcal{U}$.

1. A simple example of this kind arises when one tries to determine a rotation matrix by measuring or by computing its entries. Because of inevitable experimental or computational errors, the resulting matrix A will generally not be orthogonal. Therefore we may wish to adjust its entries to make it orthogonal. A reasonable way to do so is to change it into that orthogonal matrix U_0 which is closest to A in some appropriate norm. Finding U_0 is a problem of the type described above with $\mathcal{U}$ being the class of orthogonal matrices. This problem was solved by K. Fan and A. J. Hoffman [1], who found the unitary and hermitian matrices closest to a given matrix. Their result is independent of the norm, provided the norm is unitarily invariant, i.e., provided that $\|A\| = \|UA\| = \|AU\|$ when U is unitary.

2. Another example arises in factor analysis of psychological test data and latent structure analysis of sociological data. In both of these cases an m by n data matrix A is supposed to be a product of an m by r matrix B times an r by n matrix C, with $r < m$ and $r < n$. The product BC has rank $\leqq r$, but the rank of A will generally exceed r, so there is no such factorization in general. Therefore it

players of the team behind after 5 games may make a major effort in game 6 which cannot be sustained for game 7, while the players on the team ahead after 5 games individually husband resources in game 6. Perhaps readers will have other explanations.

References

1. H. Cramer, Mathematical Methods of Statistics, Princeton University Press, 1946, pp. 416–452.
2. Z. Hollander, Editor, The Modern Encyclopedia of Basketball, Four Winds Press, New York, 1973.
3. F. Mosteller, The World Series Competition, J. Amer. Statist. Assoc., 47 (1952) 355–380.
4. D. S. Neft, R. T. Johnson, R. M. Cohen, and J. A. Deutsch, The Sports Encyclopedia: Baseball, Grosset and Dunlap, New York, 1974.
5. R. A. Styer, The Encylopedia of Hockey, rev. ed., A. S. Barnes, New York, 1973.
6. The Official Associated Press Sports Almanac, Dell, New york, 1974.
7. M. Zelinka, How Many Games to Complete a World Series, in Statistics by Example: Weighing Chances, ed. by F. Mosteller, et al., Addison-Wesley, Reading, 1973.

CLOSEST UNITARY, ORTHOGONAL AND HERMITIAN OPERATORS TO A GIVEN OPERATOR

JOSEPH B. KELLER, Courant Institute of Mathematical Sciences, New York University

1. Introduction. It is often of interest to find an operator U_0, in some specified class of operators $\mathcal{U}$, which is closest to a given operator A. We shall describe various situations in which this kind of problem arises, and then show how to solve it for several choices of $\mathcal{U}$.

1. A simple example of this kind arises when one tries to determine a rotation matrix by measuring or by computing its entries. Because of inevitable experimental or computational errors, the resulting matrix A will generally not be orthogonal. Therefore we may wish to adjust its entries to make it orthogonal. A reasonable way to do so is to change it into that orthogonal matrix U_0 which is closest to A in some appropriate norm. Finding U_0 is a problem of the type described above with $\mathcal{U}$ being the class of orthogonal matrices. This problem was solved by K. Fan and A. J. Hoffman [1], who found the unitary and hermitian matrices closest to a given matrix. Their result is independent of the norm, provided the norm is unitarily invariant, i.e., provided that $\|A\| = \|UA\| = \|AU\|$ when U is unitary.

2. Another example arises in factor analysis of psychological test data and latent structure analysis of sociological data. In both of these cases an m by n data matrix A is supposed to be a product of an m by r matrix B times an r by n matrix C, with $r < m$ and $r < n$. The product BC has rank $\leqq r$, but the rank of A will generally exceed r, so there is no such factorization in general. Therefore it

is customary to seek the best approximate factorization. This corresponds to finding the matrix U_0 closest to A in the class $\mathcal{U}$ of m by n matrices of rank r, since every matrix in $\mathcal{U}$ can be factored in the desired manner, and every factorizable matrix is in $\mathcal{U}$. This problem was solved by C. Eckart and G. Young [2] and by J. B. Keller [3] using the euclidean norm. L. Mirsky [4] showed that the result is independent of the norm, provided the norm is unitarily invariant.

3. A modification of the above problem also arises in factor analysis, in which the matrix C is given. Then $\mathcal{U}$ is the class of matrices BC where B is an arbitrary m by r matrix and C is given. Alternatively, B may be given and C may be arbitrary. These problems were solved by B. Green [5], J. B. Keller [3] and P. Schönemann [6] using the euclidean norm.

4. A special case of problem 3 is that with $m = n = r$ and $A = I$, the identity matrix. Then the problem is to find a matrix B which makes BC closest to the identity. Such a matrix B we call a generalized left inverse of C. Similarly if B is given and if C makes BC closest to I, we call C a generalized right inverse of B. These problems have been solved by R. Penrose [7].

5. A Fredholm integral equation of the second kind is an equation of the form

$$\phi(x) + \int_a^b K(x, y)\phi(y)dy = f(x).$$

Here $\phi(x)$ is the unknown function, f is a given inhomogeneous term, and K is a given function called the kernel. The equation can be reduced to a system of n linear algebraic equations for n unknowns, and then it can be solved easily, if K is a degenerate kernel of rank n. This is a kernel of the form

$$K_n(x, y) = \sum_{j=1}^{n} g_j(x)h_j(y).$$

Therefore one way to obtain an approximate solution of the integral equation is to approximate its kernel K by a degenerate kernel of rank n. Naturally it is desirable to obtain the best such approximation, and this is again an instance of the general problem described above. The solution is given by R. Courant and D. Hilbert [8].

G. H. Golub [9] has given the solutions of problems 1–4 together with a numerical procedure for computing them.

We shall describe two methods for analyzing such problems. The first is the direct method of showing that a particular U_0 is closer to A than any other U in $\mathcal{U}$ be comparing the norm $\|A - U_0\|$ of $A - U_0$ with $\|A - U\|$. The virtue of this method is its simplicity and generality. It applies to operators on infinite dimensional spaces as well as to matrices of finite dimension. However, it has the disadvantage that the solution U_0 must be known before it can be used. The second method is the usual indirect method which is used to find minima in calculus. It consists in representing U as a matrix, differentiating $\|A - U\|$ with respect to the matrix elements, and equating the derivatives to zero. In doing this, the condition that U is in $\mathcal{U}$ must be taken into account by the use of

Lagrange multipliers. This method leads to an equation which must be satisfied by U_0. Thus it has the virtue that U_0 need not be known in advance. Its disadvantage is that after U_0 has been found, it must still be shown that U_0 does yield the smallest value of $\|A - U\|$.

I wish to thank my colleague A. B. Novikoff for some stimulating discussions of these topics.

2. Closest hermitian and anti-hermitian operators. Let A be a linear operator and $\mathcal{U}$ a set of linear operators which map a unitary space into itself. The euclidean norm $\|U\|$ of an operator U is defined as the positive square root of the trace of U times its adjoint U^+, so that

$$\|U\|^2 = \operatorname{tr} UU^+. \tag{1}$$

This is an inner product norm, so the notion of orthogonal projection can be introduced in $\mathcal{U}$, but we shall not make use of it explicitly. In terms of the matrix elements u_{ij} of U with respect to an orthonormal basis of the space, $\|U\|^2$ is given by

$$\|U\|^2 = \sum_{i,j} |u_{ij}|^2. \tag{2}$$

We define $U_0 \in \mathcal{U}$ to be closest to A if U_0 minimizes the distance $\|A - U\|$ among all $U \in \mathcal{U}$. We shall now prove

THEOREM 1. *Among all hermitian operators, the unique operator closest to A is* $U_0 = \frac{1}{2}(A + A^+)$.

Proof. U is an hermitian operator if $U = U^+$, so obviously $U_0 = \frac{1}{2}(A + A^+)$ is hermitian and so is $V = U - U_0$. Therefore any hermitian operator U can be represented as a sum $U = U_0 + V$ where $V = V^+$. Now (1) yields

$$\begin{aligned} \|A - U\|^2 &= \|A - U_0 - V\|^2 \\ &= \|A - U_0\|^2 + \operatorname{tr}[(A - U_0)V^+ + V(A^+ - U_0^+)] + \|V\|^2. \end{aligned} \tag{3}$$

By using the facts that $U_0 = U_0^+$, $V = V^+$ and that the trace of a product is invariant under cyclic permutation of the factors, we find

$$\begin{aligned} \operatorname{tr}[(A - U_0)V^+ + V(A^+ - U_0^+)] &= \operatorname{tr}[(A - U_0)V + V(A^+ - U_0)] \\ &= \operatorname{tr}[(A + A^+ - 2U_0)V] = 0. \end{aligned} \tag{4}$$

The last equality in (4) follows from the definition of U_0. Upon using (4) in (3) we obtain

$$\|A - U\|^2 = \|A - U_0\|^2 + \|V\|^2 \geqq \|A - U_0\|^2. \tag{5}$$

This proves the theorem since $\|V\|^2 > 0$ unless $V = 0$.

An anti-hermitian operator is an operator U for which $U = -U^+$. By changing a few signs in the proof above, we can prove

THEOREM 2. *Among all anti-hermitian operators, the unique operator closest to A is* $U_0 = \frac{1}{2}(A - A^+)$.

3. Closest unitary and orthogonal operators.

THEOREM 3. *A closest unitary operator to A is any unitary operator U_0 which occurs in a polar decomposition $A = (AA^+)^{\frac{1}{2}} U_0$ of A. If A is invertible then $U_0 = (AA^+)^{-\frac{1}{2}} A$ is the unique closest unitary operator to A.*

In the polar decomposition the positive square root is used. This theorem provides a characterization of the unitary factor U_0 in the polar decomposition.

Proof. Since U_0 is unitary, $U_0 U_0^+ = I$. Let U be any unitary operator and let $V = U - U_0$. From the unitarity of U and U_0 it follows that

$$U_0 V^+ + V U_0^+ + V V^+ = 0. \tag{6}$$

Thus any unitary U can be written as $U = U_0 + V$ where V satisfies (6). Now we can write

$$\begin{aligned} \|A - U\|^2 = \|A - U_0 - V\|^2 &= \|A - U_0\|^2 - \mathrm{tr}[(A - U_0)V^+ + V(A^+ - U_0^+) - VV^+] \\ &= \|A - U_0\|^2 - \mathrm{tr}[AV^+ + VA^+]. \end{aligned} \tag{7}$$

The last equality follows from (6). Next we use the polar decomposition of A in the last term in (7), then use the invariance of the trace of a product under cyclic permutation of the factors and then use (6) to obtain

$$\begin{aligned} \mathrm{tr}[AV^+ + VA^+] &= \mathrm{tr}[(AA^+)^{\frac{1}{2}} U_0 V^+ + V U_0^+ (AA^+)^{\frac{1}{2}}] \\ &= \mathrm{tr}[(AA^+)^{\frac{1}{2}}(U_0 V^+ + V U_0^+)] = -\mathrm{tr}[(AA^+)^{\frac{1}{2}} VV^+]. \end{aligned} \tag{8}$$

Since $(AA^+)^{\frac{1}{2}}$ and VV^+ are both hermitian and nonnegative, the trace of their product is nonnegative.

Upon using (8) in (7) and noting that the last trace in (8) is nonnegative, we obtain

$$\|A - U\|^2 = \|A - U_0\|^2 + \mathrm{tr}[(AA^+)^{\frac{1}{2}} VV^+] \geqq \|A - U_0\|^2. \tag{9}$$

This proves the first part of the theorem. If A is invertible then AA^+ is invertible, so the trace in (9) can vanish if and only if $V = 0$. Therefore inequality holds in (9) unless $U = U_0$. This proves the second part of the theorem.

If A is real then $A^+ = A^T$ where A^T, the transpose of A, is also real. Therefore U_0 is real and orthogonal. As a consequence Theorem 3 yields

THEOREM 4. *A closest orthogonal operator to a real operator A is any orthogonal operator U_0 which occurs in a polar decomposition $A = (AA^T)^{\frac{1}{2}} U_0$ of A. If A is invertible then $U_0 = (AA^T)^{-\frac{1}{2}} A$ is the unique closest orthogonal operator to A.*

If A is not necessarily real then by slightly modifying the proof of Theorem 3 we can prove

THEOREM 5. *A closest orthogonal operator to A is any orthogonal operator U_0 which occurs in a polar decomposition* $\mathrm{Re}\,A = (\mathrm{Re}\,A\ \mathrm{Re}\,A^T)^{\frac{1}{2}} U_0$ *of* $\mathrm{Re}\,A$. *If* $\mathrm{Re}\,A$ *is invertible then* $U_0 = (\mathrm{Re}\,A\ \mathrm{Re}\,A^T)^{-\frac{1}{2}} \mathrm{Re}\,A$.

4. Generalized left and right inverses. In the introduction we defined a generalized left inverse of A as an operator U_0 which minimizes $\|UA - I\|$. Now we shall prove

THEOREM 6. *U_0 is a generalized left inverse of A if it is a solution of the equation $U_0AA^+ = A^+$. If A is invertible then $U_0 = A^{-1}$.*

Proof. Any operator U can be written as $U = U_0 + V$ where $V = U - U_0$. Then

$$\begin{aligned}\|UA - I\|^2 &= \|(U_0 + V)A - I\|^2 \\ &= \|U_0A - I\|^2 + \|VA\|^2 + \operatorname{tr}[(U_0A - I)A^+V^+ - VA(A^+U_0^+ - I)].\end{aligned} \tag{10}$$

The trace in (10) vanishes in view of the equation satisfied by U_0 and (10) yields, since $\|VA\|^2 \geqq 0$,

$$\|UA - I\|^2 = \|U_0A - I\|^2 + \|VA\|^2 \geqq \|U_0A - I\|^2. \tag{11}$$

This proves the theorem. Similarly we can prove

THEOREM 7. *U_0 is a generalized right inverse of A if it is a solution of the equation $A^+AU_0 = A^+$. If A is invertible then $U_0 = A^{-1}$.*

Penrose [7] has shown for matrices, i.e., operators on a finite dimensional space, that there exists a unique U_0, called a generalized inverse of A, which satisfies both $U_0AA^+ = A^+$ and $A^+AU_0 = A^+$. From the theorems above, this U_0 is both a generalized left and genralized right inverse of A. Furthermore he has shown that U_0B is the unique best approximate solution of the equation $AX = B$. The best approximate solution is defined to be that X which minimizes $\|AX - B\|$, and if there is more than one minimizer, then it has the least value of $\|X\|$.

5. Lagrange multiplier method. We shall now illustrate the indirect method by using it to find a closest unitary matrix U to the matrix A of dimension n. We first introduce the matrix Λ of Lagrange multipliers λ_{ij} and consider the function $\varepsilon(U)$ defined by

$$\varepsilon(U) = \|A - U\|^2 + \frac{1}{2}\sum_{i,j,k} [\lambda_{ik}(u_{ij}u^*_{kj} - \delta_{ik}) + \lambda^*_{ik}(u^*_{ij}u_{kj} - \delta_{ik})]. \tag{12}$$

We now differentiate (12) with respect to u_{ij} and set $\partial\varepsilon/\partial u_{ij} = 0$. This yields

$$-a^*_{ij} + \frac{1}{2}\sum_k (\lambda_{ik}u^*_{kj} + \lambda^*_{ki}u^*_{kj}) = 0. \tag{13}$$

In differentiating (12) we used the fact that $\|U\|^2 = n$ to avoid differentiating $\|U\|^2$. Differentiation of (12) with respect to u^*_{ij} yields the complex conjugate of (13). If $\Lambda + \Lambda^+$ is nonsingular, the solution of (13) is

$$U^* = 2(\Lambda + \Lambda^+)^{-1}A^*. \tag{14}$$

To determine Λ we use (14) in $UU^+ = I$ and obtain

$$4(\Lambda^* + \Lambda^T)^{-1}AA^+(\Lambda^T + \Lambda^*)^{-1} = I. \tag{15}$$

Multiplication of (15) on the left and on the right by $(\Lambda^T + \Lambda^*)$ yields

$$4AA^+ = (\Lambda^T + \Lambda^*)^2. \tag{16}$$

The solution of (16) is

$$\Lambda^T + \Lambda^* = 2(AA^+)^{\frac{1}{2}}. \tag{17}$$

Then (14) and (17) yield the solution, which we denote by U_0,

$$U_0 = (AA^+)^{-\frac{1}{2}}A. \tag{18}$$

This result holds only if $(AA^+)^{\frac{1}{2}}$ is nonsingular, which is the case if and only if A is nonsingular.

The result (18) for the closest unitary matrix to A is ambiguous because the choice of square root has not been determined. To determine it we use (18) to evaluate $\|A - U_0\|^2$ and find

$$\|A - U_0\|^2 = \|A\|^2 + n - \mathrm{Tr}(AU_0^+ + U_0A^+) = \|A\|^2 + n - 2\,\mathrm{Tr}(AA^+)^{\frac{1}{2}}. \tag{19}$$

We see from (19) that to minimize $\|A - U_0\|$ we must choose the square root which maximizes $\mathrm{Tr}(AA^+)^{\frac{1}{2}}$. This is just the positive square root. The proof that U_0 given by (18) is actually closest to A is given in Section 3.

In case A is real and U is orthogonal, we may seek U in either the class with determinant plus one or that with determinant minus one. Then we must restrict the square root in (18) so that $\det U_0 = +$ or -1 and maximize $\mathrm{Tr}(AA^+)^{\frac{1}{2}}$ subject to this restriction.

Research supported by the National Science Foundation under Grant No. NSF GP-32996X.

References

1. K. Fan and A. J. Hoffman, Some metric inequalities in the space of matrices, Proc. Amer. Math. Soc., 6 (1955) 111–116.

2. C. Eckart and G. Young, The approximation of one matrix by another of lower rank, Psychometrika, 1 (1936) 211–218.

3. J. B. Keller, Factorization of matrices by least squares, Biometrika, 49 (1962) 239–242.

4. L. Mirsky, Symmetric gauge functions and unitarily invariant norms, Quart. J. Math., Oxford, (2), 11 (1960) 50–59.

5. B. Green, The orthogonal approximation of an oblique structure in factor analysis, Psychometrika, 17 (1952) 429–440.

6. P. Schoenemann, A generalized solution of the orthogonal procrustes problem, Psychometrika, 31 (1966) 1–10.

7. R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc., 51 (1955) 406–413; 52 (1956) 17–19.

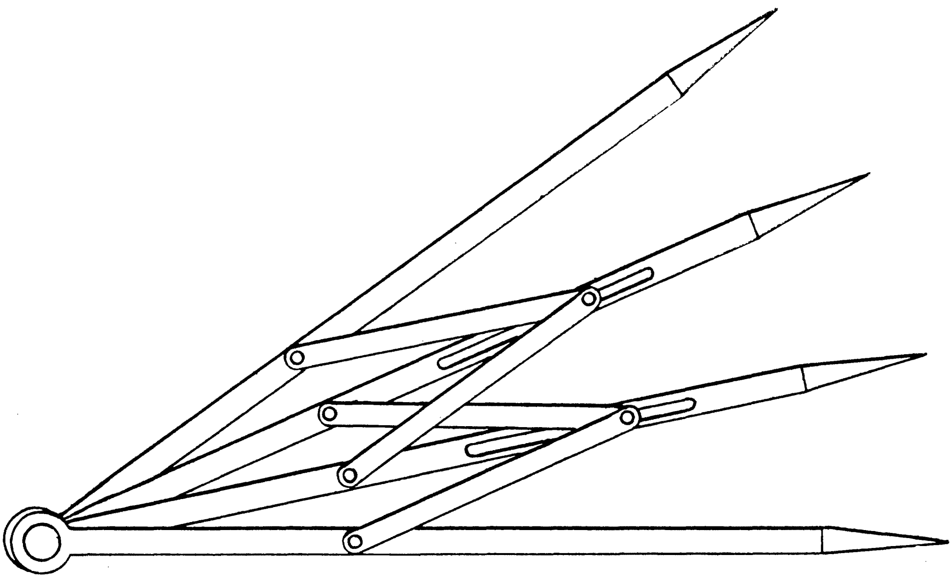
8. R. Courant and D. Hilbert, Methods of Mathematical Physics, (1), Interscience, New York, 1953, p. 161.

9. G. H. Golub, Least squares, singular values and matrix approximations, Apl. Mat., 13 (1968) 44–51.

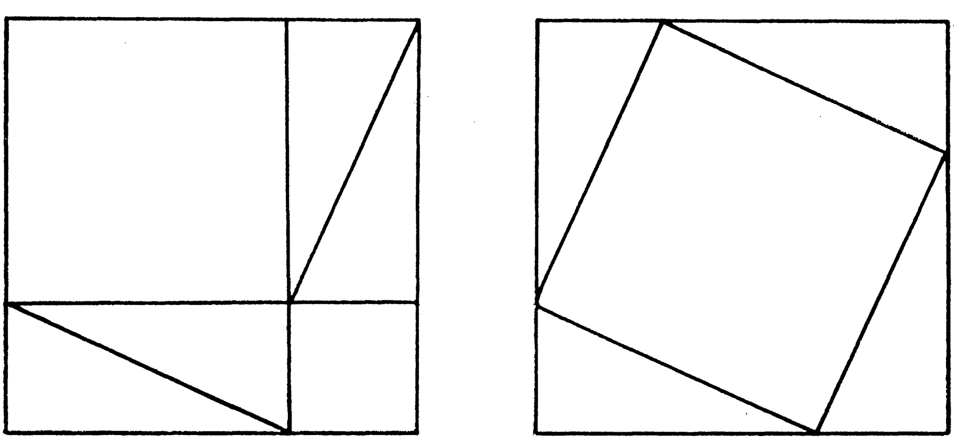
TWO MATHEMATICAL PAPERS WITHOUT WORDS

RUFUS ISAACS, The Johns Hopkins University

ON TRISECTING AN ANGLE



A PROOF OF THE PYTHAGOREAN THEOREM



MINI-PROFILES

KATHARINE O'BRIEN, Portland, Maine

Archimedes, stick in hand,
Traced his tombstone in the sand.

Khayyàm laid cubics on the line—
But better known for a jug of wine.

Fibonacci couldn't sleep—
Counted rabbits instead of sheep.

Fermat found margins a handy place,
But all too soon ran out of space.

Évariste Galois fought a duel—
Fate was ruthless, fate was cruel.

Hamilton crossed a Dublin bridge—
Carved graffiti on its ridge.

Emmy Noether—Adam's rib—
Antedating women's lib.

Gödel—giant-stride agility—
Decided undecidability.

Bourbaki keeping fit and nifty—
Component parts retire at fifty.

A MATRIX WITTICISM

*J. R. S. '74, Lehigh University

Let A, B be 2×2 matrices of real numbers; we define $t(A)$, the trace of A, to be $a_{11} + a_{22}$; and $\Delta(A)$, the determinant of A as usual. The result referred to in the title is the following:

THEOREM. $t(AB) = t(A) \cdot t(B)$ *iff* $\Delta(A + B) = \Delta(A) + \Delta(B)$.

This follows from the formula:

$$t(AB) - t(A) \cdot t(B) = \Delta(A) + \Delta(B) - \Delta(A + B)$$

which may be checked directly. It was discovered in the course of an investigation on the sum of nilpotent matrices which culminated in the discovery that the following facts for a pair of nilpotent matrices A, B, $(A \neq 0)$ are equivalent: i. $A + B$ is nilpotent, ii. AB is nilpotent, iii. $B = mA$, iv. $BA = AB$, v. $AB = 0$.

* *Note.* J. R. S. is the Junior Research Seminar for high school students, sponsored in 1974 by NSF Grant GY 11343

CONDITIONAL EXPECTATION OF THE DURATION IN THE CLASSICAL RUIN PROBLEM

FREDERICK STERN, San Jose State University

1. In the classical ruin problem a point begins a random walk on the line at an integer z and moves regularly to adjacent integers, with probability p $(0 < p < 1)$ to the higher integer and probability $q = 1 - p$ to the lower integer. The walk stops when the point reaches either 0 or the integer a $(0 < z < a)$ for the first time. The probability, q_z, of absorption at zero with starting point z is given (Feller (1968) pp. 344–349) by

(1) $$q_z = 1 - \frac{z}{a} \qquad \text{if} \qquad p = q = \frac{1}{2}$$

and

(2) $$q_z = \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^z}{\left(\frac{q}{p}\right)^a - 1} \qquad \text{if } p \neq q.$$

The expected duration of the walk is

$$D_z = z(a - z) \qquad \text{if} \qquad p = q = \tfrac{1}{2}$$

and

$$D_z = \frac{z}{q-p} - \frac{a}{q-p} \cdot \frac{1 - \left(\frac{q}{p}\right)^z}{1 - \left(\frac{q}{p}\right)^a}.$$

An extension of these results consists in deriving expressions for the conditional expectation of the number of steps before stopping given absorption either at zero or at a. If E_z is the conditional expectation for the number of steps given that the random walk stops at the point zero, and if F_z is the conditional expectation for the duration given that stopping occurs at a, then with $p_z = 1 - q_z$, (p_z being the probability of absorption at a),

(3) $$D_z = p_z F_z + q_z E_z,$$

the expected duration being equal to the weighted sum of the conditional expectations. Below we derive explicit expressions (5) and (6) if $p = 1/2$ and (8) and (9) if $p \neq 1/2$ for E_z and F_z respectively.

We then prove the curious result that if $a = 2z$ (that is, if the walk begins at the midpoint of the interval) then for any choice of p both conditional expectations E_z and F_z are equal to the (unconditional) expected duration D_z, given by z^2 if $p = 1/2$ and by

$$z(p-q)^{-1}\left[1-\left(\frac{q}{p}\right)^z\right]\left[1+\left(\frac{q}{p}\right)^z\right]^{-1} \quad \text{if } p \neq \tfrac{1}{2}.$$

In Table I we present some illustrative numerical evaluations.

TABLE I.
Absorption Probabilities, Expectation and Conditional Expectations

				Probability of absorption		Expected	Conditional Expected duration given stopping	
				at 0	at a	Duration	at 0	at a
p	q	z	a	q_z	p_z	D_z	E_z	F_z
0.5	0.5	9	10	.100	.900	9.0	33.0	6.3
		90	100	.100	.900	900	3300	633
		99	100	.010	.990	99.0	3333	66.3
0.49	0.51	9	10	.119	.881	9.5	32.9	6.3
		90	100	.336	.664	1179	2654	433.8
		99	100	.040	.960	150	2686	44.2
0.45	0.55	9	10	.210	.790	11.0	31.1	5.7
		90	100	.866	.134	766	869	100
		99	100	.182	.818	172	900	10
0.40	0.60	9	10	.339	.661	12.0	26.8	4.4
		90	100	.983	.017	441	448	50
		99	100	.333	.667	162	475	5.0

Classical random walk on integers starting at z, terminating at 0 or a.

2. If we call $u_{z,n}$ the probability that the random walk beginning at z stops at zero in n steps, then since $q_z = \sum_{n=1}^{\infty} u_{z,n}$, the product $q_z E_z$, which we rename $D_{z,0}$ satisfies

$$D_{z,0} = q_z E_z = \sum_{n=1}^{\infty} n u_{z,n}.$$

In the case $p = q = 1/2$, $D_{z,0}$ satisfies a simple difference equation

(4) $$D_{z,0} = \tfrac{1}{2} D_{z+1,0} + \tfrac{1}{2} D_{z-1,0} + q_z$$

with q_z given by equation (1) and the boundary conditions

$$D_{0,0} = D_{a,0} = 0.$$

A particular solution to this equation is $z^2(z/3a - 1)$. The general solution to the associated homogeneous equation $D_{z,0} = \frac{1}{2} D_{z+1,0} + D_{z-1,0}$ is $A + Bz$ so that the unique solution of equation (4) meeting the boundary conditions is

$$D_{z,0}=\frac{z}{3a}(2a-z)(a-z)$$

and hence

(5) $$E_z=\frac{z}{3}(2a-z),\qquad p=q=\frac{1}{2}.$$

For F_z if $p=q=1/2$, we rename p_zF_z as $D_{z,a}$ and note that $D_{z,a}=D_{a-z,0}$ or $D_{z,a}=(z/3a)(a^2-z^2)$ and

(6) $$F_z=\tfrac{1}{3}(a^2-z^2),\qquad p=q=\tfrac{1}{2}.$$

If $p\neq q$, $D_{z,0}$ satisfies

$$D_{z,0}=pD_{z+1,0}+qD_{z-1,0}+q_z$$

with q_z given by equation (2), and $D_{0,0}=D_{a,0}=0$. Rather than proceeding as above, we use an alternate method of finding $D_{z,0}$. The generating function

$$U_z(s)=\sum_{n=0}^{\infty}u_{z,n}s^n$$

is (Feller (1968) p. 350)

$$U_z(s)=\frac{\lambda_1^a(s)\lambda_2^z(s)-\lambda_1^z(s)\lambda_2^a(s)}{\lambda_1^a(s)-\lambda_2^a(s)}$$

where

$$\lambda_1(s)=\frac{1}{2ps}[1+(1-4pqs^2)^{\frac{1}{2}}],\qquad \lambda_2(s)=\frac{1}{2ps}[1-(1-4pqs^2)^{\frac{1}{2}}]$$

so that $D_{z,0}=U_z'(1)$. Differentiating $U_z(s)$, setting $s=1$ and noting for $p>q$, $\lambda_1(1)=1$, $\lambda_2(1)=q/p$, $\lambda_1'(1)=(q-p)^{-1}$, $\lambda_2'(1)=q/p(p-q)^{-1}$ and for $p<q$, the same results hold with the subscripts 1 and 2 interchanged, we get with $r=q/p\neq 1$,

(7) $$D_{z,0}=\frac{(p-q)^{-1}}{1-r^a}\left[z(r^z+r^a)+2a\left(\frac{r^{a+z}-r^a}{1-r^a}\right)\right].$$

Since $D_{z,0}=q_zE_z=((r^z-r^a)/(1-r^a))E_z$, we have

(8) $$E_z=\frac{(p-q)^{-1}}{r^z-r^a}\left[z(r^z+r^a)+2a\left(\frac{r^{a+z}-r^a}{1-r^a}\right)\right].$$

With regard to F_z, recall that $D_{z,a}=p_zF_z=((1-r^z)/(1-r^a))F_z$. The expectation $D_{z,a}$ is given, however, by $D_{a-z,0}$ if in addition the roles of p and q are interchanged in equation (7):

$$D_{z,a}=\frac{(q-p)^{-1}}{1-r^{-a}}\left[(a-z)(r^{z-a}+r^{-a})+2a\left(\frac{r^{z-2a}-r^{-a}}{1-r^{-a}}\right)\right]$$

$$=\frac{(q-p)^{-1}}{r^a-1}\left[(a-z)(r^z+1)+2a\left(\frac{r^z-r^a}{r^a-1}\right)\right].$$

We have finally

$$F_z = \frac{(p-q)^{-1}}{1-r^z}\left[(a-z)(r^z+1)+2a\left(\frac{r^z-r^a}{r^a-1}\right)\right]. \tag{9}$$

3. In the symmetrical case, $p=q=1/2$ it is not at all surprising that when $a=2z$ or in other words when the random walk starts at a point equidistant from 0 and a, the conditional expectations E_z and F_z are equal. In this case, as results (5) and (6) show they both equal z^2. That this equality persists even when $p \neq q$ is somewhat surprising. To prove it, we set $a=2z$ in result (8) so that

$$E_z = \frac{z}{(p-q)}\left[\frac{1+r^z}{1-r^z}-\frac{4r^z}{1-r^{2z}}\right]=\frac{z}{(p-q)}\frac{1-r^z}{1+r^z}.$$

Using $a=2z$ in result (9) we get that the formula for F_z also reduces to the same expression. By relation (3) or direct calculation it is clear that if $a=2z$ this expression also gives D_z, the unconditional expected duration. In short, if $a=2z$,

$$D_z = E_z = F_z = z(p-q)^{-1}(1-r^z)(1+r^z)^{-1}$$

for $p \neq q$ and $r=q/p$.

Reference

1. W. Feller, An Introduction to Probability Theory, 3rd ed., Wiley, New York, 1968.

ON SUMS OF CONSECUTIVE kth POWERS, $k=1,2$

JOHN A. EWELL, Northern Illinois University

1. Introduction. For each given triple of natural numbers j, k, m, let the functional value $f(j,k,m)$ be defined by the following equation:

$$f(j,k,m)=\sum_{i=0}^{j}(m+i)^k.$$

(Here, the term "natural number" will mean a "strictly positive integer".) Further, for given k, let the set M_k be defined by:

$$M_k=\{n: n\in \mathbf{Z}^+ \text{ and } \exists\, j,m\,(j,m\in \mathbf{Z}^+) \text{ and } n=f(j,k,m)\}.$$

We now raise the question: "Which numbers belong to M_k?" It is the purpose of this paper to supply answers to this question and related questions for the particular cases $k=1$ and $k=2$. In Section 2 we settle the case $k=1$. The first

We have finally

$$F_z = \frac{(p-q)^{-1}}{1-r^z}\left[(a-z)(r^z+1) + 2a\left(\frac{r^z - r^a}{r^a - 1}\right)\right]. \tag{9}$$

3. In the symmetrical case, $p = q = 1/2$ it is not at all surprising that when $a = 2z$ or in other words when the random walk starts at a point equidistant from 0 and a, the conditional expectations E_z and F_z are equal. In this case, as results (5) and (6) show they both equal z^2. That this equality persists even when $p \neq q$ is somewhat surprising. To prove it, we set $a = 2z$ in result (8) so that

$$E_z = \frac{z}{(p-q)}\left[\frac{1+r^z}{1-r^z} - \frac{4r^z}{1-r^{2z}}\right] = \frac{z}{(p-q)}\frac{1-r^z}{1+r^z}.$$

Using $a = 2z$ in result (9) we get that the formula for F_z also reduces to the same expression. By relation (3) or direct calculation it is clear that if $a = 2z$ this expression also gives D_z, the unconditional expected duration. In short, if $a = 2z$,

$$D_z = E_z = F_z = z(p-q)^{-1}(1-r^z)(1+r^z)^{-1}$$

for $p \neq q$ and $r = q/p$.

Reference

1. W. Feller, An Introduction to Probability Theory, 3rd ed., Wiley, New York, 1968.

ON SUMS OF CONSECUTIVE kth POWERS, $k = 1, 2$

JOHN A. EWELL, Northern Illinois University

1. Introduction. For each given triple of natural numbers j, k, m, let the functional value $f(j,k,m)$ be defined by the following equation:

$$f(j,k,m) = \sum_{i=0}^{j} (m+i)^k.$$

(Here, the term "natural number" will mean a "strictly positive integer".) Further, for given k, let the set M_k be defined by:

$$M_k = \{n : n \in \mathbf{Z}^+ \quad \text{and} \quad \exists\, j, m\,(j, m \in \mathbf{Z}^+) \quad \text{and} \quad n = f(j,k,m)\}.$$

We now raise the question: "Which numbers belong to M_k?" It is the purpose of this paper to supply answers to this question and related questions for the particular cases $k = 1$ and $k = 2$. In Section 2 we settle the case $k = 1$. The first assertion in the main theorem of this section is classical (e.g., see [**1**, p. 19]); however, the second assertion, though easy to prove, seems to have escaped attention. Section 3 deals with the case $k = 2$. Here, we also characterize the primes which belong to M_2.

Throughout this paper the notation "$n = 2^a \cdot b$" will mean that a and b are nonnegative integers with b being odd.

2. The case $k = 1$. This case is completely settled by the following

THEOREM 1. *For each given natural number* $n = 2^a \cdot b$, *n belongs to* M_1 *iff* $b > 1$. *Moreover, if* $N(n)$ *denotes the number of ways of representing n as a sum of consecutive natural numbers, it follows that*

$$N(n) = d(b) - 1,$$

$d(b)$ *denoting the number of divisors of b.*

Proof. Here, $f(j, 1, m) = (j + 1)(2m + j)/2$. Hence, if $n = 2^a \cdot b$ belongs to M_1, so that for some pair $j, m \in \mathbf{Z}^+$, $2^a \cdot b = (j + 1)(2m + j)/2$, we must have $b > 1$. (Consider the parity of j.)

We pass now to the proof of the second assertion in our theorem, as this assertion entails sufficiency of "$b > 1$" for "$n = 2^a \cdot b \in M_1$". Accordingly, let $n = 2^a \cdot b$ be given and define a function g from the set of all divisors $\beta > 1$ of b into the set of all ordered pairs (j, m) such that $n = f(j, 1, m)$ as follows: For each $\beta > 1$, let

$$g(\beta) = (j, m),$$

where

$$j = 2^{a+1} \cdot (b/\beta) - 1,$$
$$m = [\beta - (2^{a+1} \cdot (b/\beta) - 1)]/2,$$

in case $2^{a+1} \cdot (b/\beta) - 1 > \beta$; and, in case $2^{a+1} \cdot (b/\beta) - 1 \geqq \beta$, so that $2^{a+1}(b/\beta) > \beta - 1$, let

$$g(\beta) = (j, m),$$

where

$$j = \beta - 1,$$
$$m = [2^{a+1} \cdot (b/\beta) - (\beta - 1)]/2.$$

In either case, $n = f(j, 1, m)$, and g is obviously single-valued. But, g is also one-to-one. For, suppose that $g(\beta) = (j, m)$, $g(\beta') = (j', m')$ and $(j, m) = (j', m')$. Then, either

$$2^{a+1} \cdot (b/\beta) - 1 = j = j' = 2^{a+1} \cdot (b/\beta') - 1,$$

where $\beta = \beta'$, or

$$\beta - 1 = j = j' = \beta' - 1,$$

where $\beta = \beta'$, as well, Finally g is onto. For, $n = 2^a \cdot b = (j+1)(2m+j)/2$ implies that either (i) j odd, and for some divisor $\beta > 1$ of b,

$$j + 1 = 2^{a+1} \cdot (b/\beta),$$
$$2m + j = \beta,$$

or (ii) j is even, and for some divisor $\beta > 1$ of b,

$$j + 1 = \beta,$$
$$2m + j = 2^{a+1} \cdot (b/\beta).$$

In either case, $g(\beta) = (j, m)$. No pair corresponds to the divisor $\beta = 1$ of b, whence

$$N(n) = d(b) - 1,$$

whenever $n = 2^a \cdot b$, with $b > 1$. However, since $d(1) = 1$, the above formula is also valid for $n = 2^a$.

3. The case $k = 2$. We characterize the natural numbers which comprise M_2 according to the dictates of

THEOREM 2. *For each given natural number $n = 2^a \cdot b$, n belongs to M_2 iff either* (i) *there exists a factorization $\alpha\beta$ of $3b$ such that $\beta > (2^{a+1}\alpha - 1)(2^{a+2}\alpha - 1)$ and $3 \cdot \{2(\beta - 2^{2a+1} \cdot \alpha^2) + 1\}$ is a square or* (ii) *there exists a factorization $\gamma\delta$ of $3b$ such that $\delta > 1$, $2^{a+1} \cdot \gamma > (\delta - 1)(2\delta - 1)$ and $3\{2^{a+2} \cdot \gamma - \delta^2 + 1\}$ is a square.*

Proof. In this case $f(j, 2, m) = (j+1)(6m^2 + 6mj - 2j^2 - j)/6$. Now, choose any $n = 2^a \cdot b \in M_2$. Then, for some pair j, m, $2^a \cdot b = (j+1)(6m^2 + 6mj + 2j^2 + j)/6$. We consider two cases: "I. j is odd"; and "II. j is even."

I. Here, $j + 1$ is even and $(6m^2 + 6mj + 2j^2 + j)$ is odd, whence for some factorization $\alpha\beta$ of $3b$, we have:

$$2^{a+1} \cdot \alpha = j + 1,$$
$$\beta = 6m^2 + 6mj + 2j^2 + j.$$

Substituting $j = 2^{a+1} \cdot \alpha - 1$ into the second equation, we have

$$6m^2 + 6(2^{a+1} \cdot \alpha - 1) \cdot m + (2^{a+1} \cdot \alpha - 1) \cdot [2(2^{a+1} \cdot \alpha - 1) + 1] - \beta = 0.$$

By the quadratic formula,

$$6m = -3 \cdot (2^{a+1} \cdot \alpha - 1) \pm \sqrt{3 \cdot \{2(\beta - 2^{2a+1} \cdot \alpha^2) + 1\}}.$$

Since m and $(2^{a+1} \cdot \alpha - 1)$ are natural numbers, we must have

$$(3 \cdot \{2(\beta - 2^{2a+1} \cdot \alpha^2) + 1\})^{\frac{1}{2}} > 3 \cdot (2^{a+1} \cdot \alpha - 1),$$

which is equivalent to $\beta > (2^{a+1} \cdot \alpha - 1)(2^{a+2} \cdot \alpha - 1)$, and $3 \cdot \{2(\beta - 2^{2a+1} \cdot \alpha^2) + 1\}$ is a square.

II. Here, $j+1$ is odd and $(6m^2+6mj+2j^2+j)$ is even, whence for some factorization $\gamma\delta$ of $3b$, we have

$$2^{a+1}\cdot\gamma = 6m^2+6mj+2j^2+j,$$
$$\delta = j+1.$$

Following the elimination procedure of case **I**, we conclude that $2^{a+1}\cdot\gamma > (\delta-1)(2\delta-1)$ and $3\{2^{a+2}\cdot\gamma-\delta^2+1\}$ is a square. Obviously, $\delta>1$.

Conversely, suppose that $n=2^a\cdot b$ is a given natural number for which either condition (i) or condition (ii) holds. Should (i) arise, we define j and m as follows:

$$j = 2^{a+1}\cdot\alpha-1,$$
$$6\cdot m = (3\cdot\{2(\beta-2^{2a+1}\cdot\alpha^2)+1\})^{\frac{1}{2}}-3\cdot(2^{a+1}\cdot\alpha-1).$$

Note that the right side of the second equation is divisible by 2 and 3, and hence divisible by 6. And, hypothesis guarantees that m is positive. It is now a purely straightforward matter to verify that $n=f(j,2,m)$, whence $n\in M_2$. In case condition (ii) is satisfied, we set

$$j=\delta-1,$$
$$6\cdot m = (3\{2^{a+2}\cdot\gamma-\delta^2+1\})^{\frac{1}{2}}-3\cdot(\delta-1).$$

Then, by hypothesis both j and m are natural numbers, and again it is easy to check that $n=f(j,2,m)$, whence $n\in M_2$ in this case, as well.

Remarks. Obviously, every odd prime belongs to M_1 and is uniquely expressible as a sum of two consecutive natural numbers. The primes which belong to M_2 are, however, much more rare; they are, in fact, described by the following

COROLLARY. *A prime p belongs to M_2 iff either* (i) $2p-1$ *is a square and the corresponding $j=1$ or* (ii) $p>55$, $(2p-35)/3$ *is a square and the corresponding $j=5$ or* (iii) $p>5$, $(p-2)/3$ *is a square and the corresponding $j=2$.*

Proof. Suppose that p is prime and $p\in M_2$. Then, $p=2^a\cdot b$, where $a=0$ and $b=p$. Now, all of the factorizations of $3p$ are accounted for by the sets $\{1,3p\}$ and $\{3,p\}$. Therefore, either (1) $3p>(2\cdot1-1)(4\cdot1-1)=3$ and $3\{2(3p-2)+1\}=3^2\cdot(2p-1)$ is a square, or simply, $(2p-1)$ is a square, or $p>(2\cdot3-1)(4\cdot3-1)=55$ and $3\{2(p-18)+1\}=3\cdot\{2p-35\}$ is a square, or (2) $2p>(3-1(2\cdot3-1)\Leftrightarrow p>5$ and $3\cdot\{4p-9+1\}=3\{4\cdot(p-2)\}$ is a square.

Now, it is an easy matter to reverse the steps of the above argument for a given prime p. But, the above conditions (1) and (2) are just a relabeling of the conditions (i), (ii) and (iii) in the statement of our corollary. Hence, (i), (ii) or (iii) implies that $p\in M_2$.

In the table of Figure 1 we have computed values of $f(j,2,m)$ for $j=1(1)9$ and $m=1(1)9$. Under condition (i) we have the primes: 5, 13, 41, 61, 113, 181; under (ii) the listed primes are: 139, 199, 271, 811, and under (ii) we have: 29, 149.

Finally, in a forthcoming paper we plan to discuss the particular cases $k = 3$ and $k = 4$.

m/j	1	2	3	4	5	6	7	8	9
1	5	14	30	55	91	140	204	285	385
2	13	29	54	90	139	203	284	384	505
3	25	50	86	135	199	280	380	501	645
4	41	77	126	190	271	371	492	636	705
5	61	110	174	255	355	476	620	689	885
6	85	149	230	330	451	595	664	860	1085
7	113	194	294	415	559	628	824	1049	1305
8	145	245	366	510	679	875	1100	1356	1645
9	181	302	446	615	811	1036	1292	1581	1905

FIG. 1.

Reference

1. I. Niven and H. S. Zuckerman, An Introduction to the Theory of Numbers, 2nd ed., Wiley, New York, 1966.

AN INTERESTING CONTINUED FRACTION

JEFFREY SHALLIT, Student, Lower Merion High School

I. Introduction. Consider the following continued fraction

(1) $$\alpha = 1 + \frac{b-2}{2}\;\frac{b+2}{-\,b+1}\;\frac{1}{-\,b}\;\frac{1}{-\,b}\;\frac{1}{-\,b}\;-\cdots \qquad (b \geqq 2)$$

This continued fraction and its convergents have many unusual properties. In fact, the numerators and denominators of the convergents to (1) form many sequences that occur in number-theoretic problems.

II. Value of α. Since the related continued fraction

(2) $$\beta = \frac{1}{b}\;\frac{1}{-\,b}\;\frac{1}{-\,b}\;-\cdots$$

is easily shown to be equal to $\frac{1}{2}(b - \sqrt{b^2-4})$, it readily follows that $\alpha = \frac{1}{2}(b + \sqrt{b^2-4})$. It is also obvious that α is the conjugate of β and that $\alpha = 1/\beta$. α and β are the roots of the quadratic $x^2 - bx + 1 = 0$.

Finally, in a forthcoming paper we plan to discuss the particular cases $k = 3$ and $k = 4$.

m/j	1	2	3	4	5	6	7	8	9
1	5	14	30	55	91	140	204	285	385
2	13	29	54	90	139	203	284	384	505
3	25	50	86	135	199	280	380	501	645
4	41	77	126	190	271	371	492	636	705
5	61	110	174	255	355	476	620	689	885
6	85	149	230	330	451	595	664	860	1085
7	113	194	294	415	559	628	824	1049	1305
8	145	245	366	510	679	875	1100	1356	1645
9	181	302	446	615	811	1036	1292	1581	1905

FIG. 1.

Reference

1. I. Niven and H. S. Zuckerman, An Introduction to the Theory of Numbers, 2nd ed., Wiley, New York, 1966.

AN INTERESTING CONTINUED FRACTION

JEFFREY SHALLIT, Student, Lower Merion High School

I. Introduction. Consider the following continued fraction

$$\alpha = 1 + \frac{b-2}{2} \; \frac{b+2}{-\, b+1} \; \frac{1}{-\, b} \; \frac{1}{-\, b} \; \frac{1}{-\, b} \; \cdots \qquad (b \geqq 2) \tag{1}$$

This continued fraction and its convergents have many unusual properties. In fact, the numerators and denominators of the convergents to (1) form many sequences that occur in number-theoretic problems.

II. Value of α. Since the related continued fraction

$$\beta = \frac{1}{b} \; \frac{1}{-\, b} \; \frac{1}{-\, b} \; \cdots \tag{2}$$

is easily shown to be equal to $\frac{1}{2}(b - \sqrt{b^2 - 4})$, it readily follows that $\alpha = \frac{1}{2}(b + \sqrt{b^2 - 4})$. It is also obvious that α is the conjugate of β and that $\alpha = 1/\beta$. α and β are the roots of the quadratic $x^2 - bx + 1 = 0$.

The numbers α and β have the property that each number plus its reciprocal equals b:

$$\alpha + 1/\alpha = \beta + 1/\beta = b.$$

The *simple* continued fraction expansions (as contrasted with (1) and (2), which are irregular) of both α and β are interesting:

$$\alpha = b - 1 \; \frac{1}{+1+} \; \frac{1}{b-2+} \; \frac{1}{1+} \; \frac{1}{b-2+} \cdots$$

$$\beta = \frac{1}{b-1+} \; \frac{1}{1+} \; \frac{1}{b-2+} \; \frac{1}{1+} \; \frac{1}{b-2+} \cdots$$

In certain cases, $\sqrt{\alpha}$ and $\sqrt{\beta}$ are also quadratic irrationals and not quartic (biquadratic) irrationals. For we have

$$\sqrt{\alpha} = (\sqrt{b} + \sqrt{b^2-4})/\sqrt{2} = \tfrac{1}{2}(\sqrt{b+2} + \sqrt{b-2}).$$

Now if $b = x^2 + 2$, then

$$\sqrt{\alpha} = \tfrac{1}{2}(x + \sqrt{x^2+4}) = x \; \frac{1}{+x} \; \frac{1}{+x} \; \frac{1}{+x+} \cdots$$

If $b = x^2 - 2$, then

$$\sqrt{\alpha} = \tfrac{1}{2}(x + \sqrt{x^2-4}) = x - 1 \; \frac{1}{+1+} \; \frac{1}{x-2+} \; \frac{1}{1+} \; \frac{1}{x-2+} \cdots$$

There are similar expansions for $\sqrt{\beta}$. We have

$$\sqrt{\beta} = (\sqrt{b} - \sqrt{b^2-4})/\sqrt{2} = \tfrac{1}{2}(\sqrt{b+2} - \sqrt{b-2}).$$

If $b = x^2 + 2$, then

$$\sqrt{\beta} = \tfrac{1}{2}(\sqrt{x^2+4} - x) = \frac{1}{x} \; \frac{1}{+x} \; \frac{1}{+x} \; \frac{1}{+x+} \cdots$$

If $b = x^2 - 2$, then

$$\sqrt{\beta} = \tfrac{1}{2}(x - \sqrt{x^2-4}) = \frac{1}{x-1+} \; \frac{1}{1+} \; \frac{1}{x-2+} \; \frac{1}{1+} \; \frac{1}{x-2+} \cdots$$

III. Convergents to (1). The first few convergents, p_n/q_n, to (1) for $b = 2, 3, 4, 5, 6$ and $n = 1, 2, 3, 4, 5, 6, 7, 8, 9$ are given in Table I.

By the rule for determining the convergents to a continued fraction, we have

$$p_1/q_1 = 1/1, \quad p_2/q_2 = b/2, \quad p_3/q_3 = (b^2-2)/b.$$

Also, for $n > 1$, we have $p_n = q_{n+1}$.

TABLE I

Convergents to (1)

$b \diagdown n$	1	2	3	4	5	6	7	8	9
2	$\frac{1}{1}$	$\frac{2}{2}$	$\frac{2}{2}$	$\frac{2}{2}$	$\frac{2}{2}$	$\frac{2}{2}$	$\frac{2}{2}$	$\frac{2}{2}$	$\frac{2}{2}$
3	$\frac{1}{1}$	$\frac{3}{2}$	$\frac{7}{3}$	$\frac{18}{7}$	$\frac{47}{18}$	$\frac{123}{47}$	$\frac{322}{123}$	$\frac{843}{322}$	$\frac{2207}{843}$
4	$\frac{1}{1}$	$\frac{4}{2}$	$\frac{14}{4}$	$\frac{52}{14}$	$\frac{194}{52}$	$\frac{724}{194}$	$\frac{2702}{724}$	$\frac{10084}{2702}$	$\frac{37634}{10084}$
5	$\frac{1}{1}$	$\frac{5}{2}$	$\frac{23}{5}$	$\frac{110}{23}$	$\frac{527}{110}$	$\frac{2525}{527}$	$\frac{12098}{2525}$	$\frac{57965}{12098}$	$\frac{277727}{57965}$
6	$\frac{1}{1}$	$\frac{6}{2}$	$\frac{34}{6}$	$\frac{198}{34}$	$\frac{1154}{198}$	$\frac{6726}{1154}$	$\frac{39202}{6726}$	$\frac{228486}{39202}$	$\frac{1331714}{228486}$

Now consider the Fibonacci-like sequence defined by the second order recurrence

$$a_n = ba_{n-1} - a_{n-2}, \qquad a_0 = 2, \qquad a_1 = b.$$

By the theory of difference equations, it can be shown that

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

But $\alpha = \frac{1}{2}(b + \sqrt{b^2-4})$, $\beta = \frac{1}{2}(b - \sqrt{b^2-4})$. Therefore, $a_n = \alpha^n + \beta^n$. By induction, it is easily demonstrated that $a_n = p_{n+1} = q_{n+2}$. From equation (3) it also easily follows that $a_{2n} = a_n^2 - 2$.

IV. The case $b = 3$. If $b = 3$, then $\alpha = \frac{1}{2}(3 + \sqrt{5}) = \phi + 1$, where ϕ is phi, the golden ratio [1], and $\beta = \frac{1}{2}(3 - \sqrt{5}) = 2 - \phi$.

The sequence of numerators, p_n, to the continued fraction in equation (1) is

$$3, 7, 18, 47, 123, 322, 843, 2207, \cdots$$

In fact, for $n > 1$, $p_n = L_{2n-2}$, where L_n is the Lucas sequence defined by $L_0 = 2$, $L_1 = 1$, $L_n = L_{n-1} + L_{n-2}$ [2].

Also, $p_{2^n+1} = r_n$ is another one of the sequences studied by Lucas [3], defined by $r_0 = 3$, $r_{n+1} = r_n^2 - 2$. This sequence

$$3, 7, 47, 2207, 4870847, \cdots$$

was employed by Lucas to test the primality of Mersenne numbers of the form $2^{4m+3} - 1$, where $4m + 3$ is prime.

Sierpinski [4] noted that

$$\beta=\tfrac{1}{2}(3-\sqrt{5})=2-\phi=\frac{1}{r_0}+\frac{1}{r_0r_1}+\frac{1}{r_0r_1r_2}+\frac{1}{r_0r_1r_2r_3}+\cdots$$

V. The case $b=4$. If $b=4$, then $\alpha=2+\sqrt{3}$, $\beta=2-\sqrt{3}$, and p_n is the sequence

$$4, 14, 52, 194, 724, 2702, 10084, 37634, \cdots$$

$p_{2^n+1}=s_n$ is another sequence discussed by Lucas, defined by $s_0=4$, $s_{n+1}=s_n^2-2$. Lucas employed this sequence to test the primality of Mersenne numbers [5]. Lehmer [6] improved the test to the following form:

If n is an odd prime, then 2^n-1 is prime if and only if it evenly divides s_{n-1}. The sequence s_n is

$$4, 14, 194, 37634, 1416317954, \cdots$$

VI. The case $b=6$. If $b=6$, then $\alpha=3+2\sqrt{2}$, $\beta=3-2\sqrt{2}$, and p_n is the sequence

$$6, 34, 198, 1154, 6726, 39202, \cdots$$

This sequence is involved in the determination of whether or not the product of three consecutive triangular numbers, $T_{n-1}T_nT_{n+1}$, is a square. In fact, $T_{n-1}T_nT_{n+1}$ is a square if $n=(3p_k-2)/4$. See Beiler [7].

$P_{2^n+1}=v_n$ is still another sequence discussed by Lucas [8]. The sequence v_n is as follows:

$$6, 34, 1154, 1331714, \cdots$$

where $v_0=6$, $v_{n+1}=v_n^2-2$. This sequence was employed by Lucas to test the primality of Fermat numbers $2^{2^n}+1$.

VII. The case $b=\sqrt{5}$. This case is rather unusual because b is not an integer, so none of the convergents except p_1/q_1 represent rational numbers. We have $\alpha=\tfrac{1}{2}(1+\sqrt{5})=\phi$ and $\beta=\tfrac{1}{2}(\sqrt{5}-1)=\phi-1$. The first 9 convergents to (1) with $b=\sqrt{5}$ are given in Table II.

TABLE II

p_n/q_n for $b=\sqrt{5}$

n	1	2	3	4	5	6	7	8	9
$p_n \mid q_n$	$\frac{1}{1}$	$\frac{\sqrt{5}}{2}$	$\frac{3}{\sqrt{5}}$	$\frac{2\sqrt{5}}{3}$	$\frac{7}{2\sqrt{5}}$	$\frac{5\sqrt{5}}{7}$	$\frac{18}{5\sqrt{5}}$	$\frac{13\sqrt{5}}{18}$	$\frac{47}{13\sqrt{5}}$

From equation (3) it is easy to show that $p_{2n}/\sqrt{5} = F_{2n-1}$, where F_n is the famous Fibonacci sequence defined by $F_0 = 0$, $F_1 = 1$, $F_n = F_{n-1} + F_{n-2}$. The first few terms of the Fibonacci sequence are

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, \cdots$$

From equation (3) it also can be shown that $p_{2n+1} = L_{2n}$, where L_n is the Lucas sequence discussed in part IV.

References

1. H. E. Huntley, The Divine Proportion, Dover, New York, 1970, pp. 25–26.
2. V. E. Hoggatt, Jr., Fibonacci and Lucas Numbers, Houghton Mifflin, Boston, 1969, p. 7.
3. Edouard Lucas, Comptes Rendus, Paris, V, 83 (1876) 1286–8.
4. W. Sierpinski, Elementary Theory of Numbers, Panstwowe Wydawnictwo Naukowe, Warsaw, 1964.
5. A. H. Beiler, Recreations in the Theory of Numbers, Dover, New York, 1966, pp. 17–18.
6. D. H. Lehmer, On Lucas's Test for the Primality of Mersenne's Numbers, J. London Math. Soc., V, 10 (1935) 162.
7. A. H. Beiler, loc. cit., p. 198.
8. Edouard Lucas, Théorie des fonctions numériques simplement périodiques, Amer. J. Math., V, 1 (1878) 184–239; 289–321.

Bibliography

A. V. Aho and N. J. A. Sloane, Some Doubly Exponential Sequences, Fibonacci Quart., V, 11 (1973) 429–437.

L. E. Dickson, History of the Theory of Numbers, V. 1 (Divisibility and Primality), Chelsea, New York, 1952.

Donald Knuth, The Art of Computer Programming, V. 2 (Seminumerical Algorithms), Addison-Wesley, Reading, Mass., 1969.

N. J. A. Sloane, A Handbook of Integer Sequences, Academic Press, New York, 1973.

———, Supplement I to A Handbook of Integer Sequences, Bell Laboratories, Murray Hill, New Jersey, 1973.

QUARTIC EQUATIONS AND TETRAHEDRAL SYMMETRIES

ROGER CHALKLEY, University of Cincinnati

1. Introduction. In Section 2, we give a short derivation of formulas for the roots of a quartic equation. A closely related representation of the symmetric group S_4 by matrices of size 3×3 is presented in Section 3. Geometric interpretations follow in Section 6.

Throughout, let F be a field in which $1 + 1 \neq 0$. For us, the matrix

$$H = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

From equation (3) it is easy to show that $p_{2n}/\sqrt{5} = F_{2n-1}$, where F_n is the famous Fibonacci sequence defined by $F_0 = 0$, $F_1 = 1$, $F_n = F_{n-1} + F_{n-2}$. The first few terms of the Fibonacci sequence are

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, \cdots$$

From equation (3) it also can be shown that $p_{2n+1} = L_{2n}$, where L_n is the Lucas sequence discussed in part IV.

References

1. H. E. Huntley, The Divine Proportion, Dover, New York, 1970, pp. 25–26.
2. V. E. Hoggatt, Jr., Fibonacci and Lucas Numbers, Houghton Mifflin, Boston, 1969, p. 7.
3. Edouard Lucas, Comptes Rendus, Paris, V, 83 (1876) 1286–8.
4. W. Sierpinski, Elementary Theory of Numbers, Panstwowe Wydawnictwo Naukowe, Warsaw, 1964.
5. A. H. Beiler, Recreations in the Theory of Numbers, Dover, New York, 1966, pp. 17–18.
6. D. H. Lehmer, On Lucas's Test for the Primality of Mersenne's Numbers, J. London Math. Soc., V, 10 (1935) 162.
7. A. H. Beiler, loc. cit., p. 198.
8. Edouard Lucas, Théorie des fonctions numériques simplement périodiques, Amer. J. Math., V, 1 (1878) 184–239; 289–321.

Bibliography

A. V. Aho and N. J. A. Sloane, Some Doubly Exponential Sequences, Fibonacci Quart., V, 11 (1973) 429–437.

L. E. Dickson, History of the Theory of Numbers, V. 1 (Divisibility and Primality), Chelsea, New York, 1952.

Donald Knuth, The Art of Computer Programming, V. 2 (Seminumerical Algorithms), Addison-Wesley, Reading, Mass., 1969.

N. J. A. Sloane, A Handbook of Integer Sequences, Academic Press, New York, 1973.

———, Supplement I to A Handbook of Integer Sequences, Bell Laboratories, Murray Hill, New Jersey, 1973.

QUARTIC EQUATIONS AND TETRAHEDRAL SYMMETRIES

ROGER CHALKLEY, University of Cincinnati

1. Introduction. In Section 2, we give a short derivation of formulas for the roots of a quartic equation. A closely related representation of the symmetric group S_4 by matrices of size 3×3 is presented in Section 3. Geometric interpretations follow in Section 6.

Throughout, let F be a field in which $1 + 1 \neq 0$. For us, the matrix

$$H = \tfrac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

Suppose 4×4 matrices B and D over F satisfy $BH = HD$. Then, it is easy to verify: D is a diagonal matrix if and only if B has the form

$$B = \begin{bmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \delta & \gamma \\ \gamma & \delta & \alpha & \beta \\ \delta & \gamma & \beta & \alpha \end{bmatrix}. \tag{1}$$

When B is given by (1), the diagonal elements of $D = HBH$ are

$$\begin{aligned} \xi_1 &= \alpha + \beta + \gamma + \delta, \\ \xi_2 &= \alpha + \beta - \gamma - \delta, \\ \xi_3 &= \alpha - \beta + \gamma - \delta, \\ \xi_4 &= \alpha - \beta - \gamma + \delta. \end{aligned} \tag{2}$$

2. Formulas for the roots of a quartic equation.

THEOREM. *Suppose a, b, c, r_1, r_2, r_3 are elements of F such that*

$$Y^3 + 2aY^2 + (a^2 - 4c)Y - b^2 = (Y - r_1^2)(Y - r_2^2)(Y - r_3^2) \tag{3}$$

and $r_1r_2r_3 = -b$. Then, the quartic equation

$$X^4 + aX^2 + bX + c = 0$$

has four roots $\xi_1, \xi_2, \xi_3, \xi_4$ in F given by

$$[\xi_1, \xi_2, \xi_3, \xi_4] = [0, r_1, r_2, r_3]H. \tag{4}$$

Proof. Set $\alpha = 0$, $\beta = \frac{1}{2}r_1$, $\gamma = \frac{1}{2}r_2$, $\delta = \frac{1}{2}r_3$; define B by (1); and, set

$$X^4 + \rho X^2 + \sigma X + \tau = \det(XI - B). \tag{5}$$

We expand (5) and use (3) with $r_1r_2r_3 = -b$ to obtain

$$\begin{aligned} \rho &= -2(\beta^2 + \gamma^2 + \delta^2) = (-r_1^2 - r_2^2 - r_3^2)/2 = (2a)/2 = a, \\ \sigma &= -8\beta\gamma\delta = -r_1r_2r_3 = b, \quad \text{and} \\ \tau &= \beta^4 + \gamma^4 + \delta^4 - 2(\beta^2\gamma^2 + \beta^2\delta^2 + \gamma^2\delta^2) \\ &= (\beta^2 + \gamma^2 + \delta^2)^2 - 4(\beta^2\gamma^2 + \beta^2\delta^2 + \gamma^2\delta^2) \\ &= (r_1^2 + r_2^2 + r_3^2)^2/16 - 4(r_1^2r_2^2 + r_1^2r_3^2 + r_2^2r_3^2)/16 \\ &= (-2a)^2/16 - 4(a^2 - 4c)/16 = c. \end{aligned}$$

For the diagonal elements of $D = HBH$, (2) yields (4). We find

$$X^4 + aX^2 + bX + c = \det(XI - B) = \det(XI - D) = \prod_{s=1}^{4} (X - \xi_s).$$

This completes the proof.

For other derivations, see [7], [5], [6], and [1]; the notation of [5] and [6] corresponds to a substitution of $-Y$ for Y in (3).

3. A representation for the symmetric group S_4. Let $Q_1, Q_2, \cdots, Q_6$ be the six permutation matrices of size 3×3. Set

$$R_1=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},\qquad R_2=\begin{bmatrix}1&0&0\\0&-1&0\\0&0&-1\end{bmatrix},$$

$$R_3=\begin{bmatrix}-1&0&0\\0&1&0\\0&0&-1\end{bmatrix},\qquad R_4=\begin{bmatrix}-1&0&0\\0&-1&0\\0&0&1\end{bmatrix}.$$

PROPOSITION. *Under multiplication, the twenty-four 3×3 matrices*

$$Q_jR_k, \quad \text{for} \quad j=1,\cdots,6 \quad \text{and} \quad k=1,\cdots,4, \tag{6}$$

form a group which is isomorphic to S_4.

Proof. We define matrices P_j and T_k of size 4×4 by

$$P_j=\begin{bmatrix}1&0\;0\;0\\0&\\0&Q_j\\0&\end{bmatrix} \quad\text{and}\quad T_k=\begin{bmatrix}1&0\;0\;0\\0&\\0&R_k\\0&\end{bmatrix},$$

for $j=1,\cdots,6$ and $k=1,\cdots,4$. We shall show that the matrices

$$P_jT_k, \quad \text{for} \quad j=1,\cdots,6 \quad \text{and} \quad k=1,\cdots,4, \tag{7}$$

form a group isomorphic to S_4; then, the same is clearly true for (6).

Let P be a 4×4 permutation matrix. Then, $2PH$ is obtained from $2H$ by a permutation of the rows of $2H$. Thus, the number of components equal to -1 in the first row of $2PH$ is either zero or two; and, each component of the first column of $2PH$ equals 1. We select k from 1, 2, 3, 4 such that each component of the first row of $2PHT_k$ equals 1. Then, each component of the first column of $2PHT_k$ also equals 1; the four columns of $2PHT_k$ are distinct; and, each of the last three columns of $2PHT_k$ has two components equal to 1 and two components equal to -1. Hence, $2H$ can be obtained from $2PHT_k$ by a permutation of the last three columns of $2PHT_k$. We select j from $1, 2, \cdots, 6$ so that

$$(2PHT_k)P_j^{-1}=2H \quad\text{and}\quad HPH=P_jT_k^{-1}=P_jT_k.$$

Consequently, HPH is one of the matrices in (7).

Under multiplication, the twenty-four permutation matrices of size 4×4 form a group which is isomorphic to S_4. As P ranges over the permutation matrices of size 4×4, the corresponding matrices HPH form a group isomorphic to S_4; but, these are the matrices of (7). This completes the proof.

4. Permutations of the roots. From Section 2, we have

$$[0, r_1, r_2, r_3] = [\xi_1, \xi_2, \xi_3, \xi_4]H.$$

Thus, we obtain $r_1 - r_2 = \xi_2 - \xi_3$, $r_1 + r_2 = \xi_1 - \xi_4, \cdots$, and

$$(r_1^2 - r_2^2)(r_1^2 - r_3^2)(r_2^2 - r_3^2) = \prod_{1 \leq j < k \leq 4} (\xi_j - \xi_k).$$

(For the discriminant of $X^4 + aX^2 + bX + c$, see [7], pp. 173–174.)

As Q_jR_k ranges over the 3×3 matrices of (6), the formula

$$[r_1', r_2', r_3'] = [r_1, r_2, r_3](Q_jR_k)$$

specifies the elements r_1', r_2', r_3' of F which can be substituted for r_1, r_2, r_3 to satisfy the hypothesis of the Theorem in Section 2. The corresponding roots $\xi_1', \xi_2', \xi_3', \xi_4'$ are given by

$$(8) \qquad \begin{aligned} [\xi_1', \xi_2', \xi_3', \xi_4'] &= [0, r_1', r_2', r_3']H = [0, r_1, r_2, r_3](P_jT_k)H \\ &= [0, r_1, r_2, r_3]H(HP_jT_kH) = [\xi_1, \xi_2, \xi_3, \xi_4]P, \end{aligned}$$

where P is the 4×4 permutation matrix $P = HP_jT_kH$.

5. Several observations. Here, let F be the field of complex numbers. Then, each finite group has a character table over F ([4], p. 306). The $n \times n$ matrix M_n of [2] specifies a character table for a cyclic group of order n ([4], ex. 10 of pp. 309, 342); and, for $n = 4$, the formula

$$[\xi_1', \xi_2', \xi_3', \xi_4'] = [0, u_0', v_0', w_0']M_4$$

relates the roots of $X^4 + aX^2 + bX + c$ to elements u_0', v_0', w_0' described in [2]. With (8), we have

$$[0, u_0', v_0', w_0'] = [0, r_1, r_2, r_3](HPH)(HM_4^{-1}).$$

In contrast, the matrix $2H$ specifies a character table for the four-group ([4], ex. 11 of pp. 309, 342). Thus, two nonisomorphic groups of order 4 account for the distinct solution procedures (based on M_4 and H) for a quartic equation.

6. Tetrahedral symmetries. Now, let F be the field of real numbers. Relative to a rectangular cartesian coordinate system, the points $(1, 1, 1)$, $(1, -1, -1)$, $(-1, 1, -1)$, $(-1, -1, 1)$ are the vertices of a regular tetrahedron. The matrices of (6) represent isometries which map this tetrahedron onto itself. Each of R_2, R_3, R_4 specifies a half-turn about a line through the midpoints of a pair of opposite edges of the tetrahedron. The isometries represented by $Q_1, Q_2, \cdots, Q_6$ form a group of three rotations and three reflections; they leave the vertex $(1, 1, 1)$ fixed and permute the other three vertices. Thus, the full group of isometries for a regular tetrahedron is conveniently represented by (6); it is isomorphic to S_4. The twelve matrices of (6) with determinant equal to 1 represent all the rotations for the tetrahedron; they correspond to the even permutations in S_4. Thus, the group of directisometries for a regular tetrahedron is isomorphic to the alternating subgroup of S_4. For other viewpoints, see [3].

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. H. S. M. Coxeter, Introduction to Geometry, Wiley, New York, 1961, pp. 271–272.
4. E. D. Nering, Linear Algebra and Matrix Theory, Wiley, New York, 1970, pp. 292–309.
5. B. L. van der Waerden, Modern Algebra, vol, 1, Ungar, New York, 1953, pp. 180–182.
6. S. Warner, Modern Algebra, vol. 2, Prentice-Hall, Englewood Cliffs, New Jersey, 1965, pp. 574–578.
7. H. Weber, Lehrbuch der Algebra, vol. 1, Chelsea, New York, 1963, pp. 135–136.

NOTES ON THE HISTORY OF GEOMETRICAL IDEAS I. HOMOGENEOUS COORDINATES

DAN PEDOE, University of Minnesota

It is agreed, even by those who disparage them (see p. 712 of [3]) that barycentric coordinates, first introduced by August Ferdinand Möbius [1] in 1827, were the first homogeneous coordinates systematically used in geometry. The Möbius idea, in plane geometry for example, is to attach masses p, q, and r respectively to three noncollinear points A, B and C in the plane under consideration, and then to consider the centroid $P = pA + qB + rC$ of the three masses. The point P necessarily lies in the plane, and varies as the ratios $p:q:r$ vary. As Möbius points out: *Und umgekehrt, ist irgend ein Punct P der Ebene gegeben, so sind damit auch die Werthe der Verhältnissee p : q : r immer und ohne Zweideutigkeit bestimmbar* (and conversely, given any point P in the plane the ratios $p:q:r$ are always and uniquely determinable).

It will be noted that Möbius was using position-vectors for his points in 1827, and reading of the text shows that he developed all the techniques of homogeneous coordinates known nowadays, changing the simplex of reference, if necessary, and so on. For a more accessible account, see Section 4.2 of [2]. Nobody has suggested that there is a better system of coordinates for projective geometry.

But new ideas are not always easily accepted. All the same, it is strange nowadays to read some of Cauchy's criticisms (XI of [1]). He says: *Ce n'est que par une étude plus approfondie qu'on peut décider si les avantages de cette méthode en compensent les difficultés* (only by deeper study can one decide whether the advantages of this method outweigh its difficulties), adding: "*Il faut être bien sûr qu'on fait faire à la science un grand pas, pour la surcharger de tant de termes nouveaux, et d'exiger des lecteurs qu'ils vous suivent dans des recherches, qui s'offrent à eux avec tant d'étrangeté.*" (One should be quite sure

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. H. S. M. Coxeter, Introduction to Geometry, Wiley, New York, 1961, pp. 271–272.
4. E. D. Nering, Linear Algebra and Matrix Theory, Wiley, New York, 1970, pp. 292–309.
5. B. L. van der Waerden, Modern Algebra, vol, 1, Ungar, New York, 1953, pp. 180–182.
6. S. Warner, Modern Algebra, vol. 2, Prentice-Hall, Englewood Cliffs, New Jersey, 1965, pp. 574–578.
7. H. Weber, Lehrbuch der Algebra, vol. 1, Chelsea, New York, 1963, pp. 135–136.

NOTES ON THE HISTORY OF GEOMETRICAL IDEAS I. HOMOGENEOUS COORDINATES

DAN PEDOE, University of Minnesota

It is agreed, even by those who disparage them (see p. 712 of [3]) that barycentric coordinates, first introduced by August Ferdinand Möbius [1] in 1827, were the first homogeneous coordinates systematically used in geometry. The Möbius idea, in plane geometry for example, is to attach masses p, q, and r respectively to three noncollinear points A, B and C in the plane under consideration, and then to consider the centroid $P = pA + qB + rC$ of the three masses. The point P necessarily lies in the plane, and varies as the ratios $p:q:r$ vary. As Möbius points out: *Und umgekehrt, ist irgend ein Punct P der Ebene gegeben, so sind damit auch die Werthe der Verhältnissee p : q : r immer und ohne Zweideutigkeit bestimmbar* (and conversely, given any point P in the plane the ratios $p:q:r$ are always and uniquely determinable).

It will be noted that Möbius was using position-vectors for his points in 1827, and reading of the text shows that he developed all the techniques of homogeneous coordinates known nowadays, changing the simplex of reference, if necessary, and so on. For a more accessible account, see Section 4.2 of [2]. Nobody has suggested that there is a better system of coordinates for projective geometry.

But new ideas are not always easily accepted. All the same, it is strange nowadays to read some of Cauchy's criticisms (XI of [1]). He says: *Ce n'est que par une étude plus approfondie qu'on peut décider si les avantages de cette méthode en compensent les difficultés* (only by deeper study can one decide whether the advantages of this method outweigh its difficulties), adding: "*Il faut être bien sûr qu'on fait faire à la science un grand pas, pour la surcharger de tant de termes nouveaux, et d'exiger des lecteurs qu'ils vous suivent dans des recherches, qui s'offrent à eux avec tant d'étrangeté.*" (One should be quite sure that he is making a considerable advance in science before introducing so many new terms and requiring readers to follow studies which confront them with so much strangeness.)

It should be said at this point that the Möbius work is beautifully lucid and unpedantic. As Möbius' editor, Richard Baltzer, points out, Cauchy's next remark shows that Cauchy, like most reviewers, had not read Möbius very carefully. Cauchy says: "*On doit penser que l'auteur du calcul barycentrique n'a point eu connaissance de la théorie générale de réciprocité entre les propriétés d'un système de points et d'un système de lignes, que M. Gergonne a établie.*" (One must suppose that the author of the barycentric calculus was unaware of the general theory of reciprocity which Gergonne has established between the properties of systems of points and systems of lines.) In fact Möbius, in Chapters IV and V of the third section of his book, gives a very clear statement, with applications, of The Principle of Reciprocity and the perhaps more general Principle of Duality, and these discoveries were made independently of Poncelet and Gergonne.

Gauss, writing to Schumacher rather later, in 1843, also confessed that he found the new ideas of Möbius difficult, and commits himself to a rather heavy philosophical statement about new ideas in general. This can also be read in the introduction to the Möbius work (XII, [1]). But the editor Baltzer, although an advocate for Möbius, makes a statement which I have been unable to verify. He says that although the Möbius coordinates are of special significance historically, they were soon overtaken by homogeneous coordinates and the introduction of homogeneous equations (IX, [1]). While the Möbius work was being printed, says Baltzer, Plücker, using ideas of Gergonne, with some points of contact with Lamé and Bobillier, was writing his Entwicklungen [4] in which the homogeneous coordinates of a point with respect to three lines and four planes occur more and more.

It is true that Möbius never writes down an equation for his lines or his conics. Everything is treated *parametrically*, as we say nowadays. I have read Plücker carefully, in an attempt to confirm Baltzer's assertion, and I find that although Plücker conceives the idea of the homogeneous coordinates $[u, v, w]$ of a line $uX + vY + w = 0$ at a very early stage in his two-volume work, there is a very definite psychological block with regard to the homogeneous coordinates of a point. These are mentioned for the first time in Section 416 of the second volume, where he says of the equation $UX + VY + WZ = 0$: "*Diese neue Form der Gleichung der geraden Linie, eine Gleichung die rücksichtlich der drei veränderlichen Grössen y, x und z homogen ist, scheint mir in allen demjenigen Entwicklungen, bei welchen keine Elimination vorkommt, in Beziehung auf Eleganz und Symmetrie den Vorrang zu behaupten.*" (This new form of the equation of a straight line, which is homogeneous in the three variables y, x, and z, seems to me to be preferable in regard to elegance and symmetry in all those situations where no elimination is involved.)

It is significant that Plücker rules out the use of homogeneous coordinates in elimination, and in fact he devotes nearly 200 pages to the conic considered as a line-locus in his second volume, and only about 140 pages to the conic considered as a point-locus in his first volume, and there is no systematic use of homogeneous point-coordinates in these two volumes.

What Baltzer may refer to is Plücker's very ingenious use of what we call *abridged notation* nowadays. The simple, but seminal idea behind this is that if $U = 0$ and $V = 0$ are the equations of two curves, then $U + kV = 0$, for all values of k, is a curve which passes through the intersections of the two given curves. When the curves are lines, Plücker uses, in essence, the trilinear coordinates of a point in proving his many theorems about triangles. Trilinear coordinates are the distances of a point, from the sides of a given triangle in two dimensions, and the sides of a given tetrahedron in three dimensions, and they are a form of homogeneous coordinates.

But there is no doubt that Plücker had an enormous influence on the development of analytical geometry, and there is little that one can add to his extensive treatment of circles, for example. He even uses inversion. Möbius avoids circles, and as we know, they are not too easy to deal with in homogeneous coordinates. Laguerre had not yet arrived on the scene, and the analytic treatment of the circular points at infinity is very helpful in this connection. But Möbius, besides solving many delightful problems in a masterly way, is the only author I have ever encountered who considers the different types of conic which can be drawn through five given points in a plane. His approach is a mixture of notions of convexity and the use of theorems already derived by his methods, and he says:

"Given five points chosen arbitrarily in a plane, the chance that the unique conic which can be drawn through these points is a hyperbola rather than an ellipse is $\sqrt{\infty}$: 1."

This is, of course, a preCantorian statement, but it embodies a definite theorem. Does it come under the heading of *elementary* or *advanced* problems? Treated algebraically, the theorem gives information about the range of a certain polynomial in 10 variables.

References

1. A. F. Möbius, Gesammelte Werke, Erster Band, S. Hirzel, Leipzig, 1885.

2. D. Pedoe, A Course of Geometry for Colleges and Universities, Cambridge University Press, New York, 1970.

3. ———, Thinking geometrically, Amer. Math. Monthly, 77 (1970) 711–721.

4. J. Plücker, Analytisch-Geometrische Entwicklungen, Baedeker, Essen, vol. I, 1828, vol. II, 1831.

CENTRALIZERS AND NORMALIZERS IN HAUSDORFF GROUPS

DOUGLASS L. GRANT, St. Francis Xavier University (Sydney Campus)

A classical, elementary theorem of topological groups states that the center of a Hausdorff group is closed. The two theorems below are natural generalizations of this fact, which the author has never seen in print. The principal interest lies, however, in the method of proof, which seems to afford more elegance than those of the classical result appearing in [**2**] and [**3**].

If G is a group and A a subset thereof, the *centralizer* of A is the set of elements of G which commute with each element of A, and the *normalizer* of A is the set of elements, conjugation with respect to which leaves A unchanged; they will be denoted by $C_G(A)$ and $N_G(A)$, respectively. The conjugation map $x \mapsto bxb^{-1}$ will be denoted by c_b.

THEOREM. *Let A be any subset of a Hausdorff topological group G. Then $C_G(A)$ is closed.*

Proof. Since conjugation maps are continuous and G is Hausdorff, the set on which any such map agrees with the identity map is closed [**1**, p. 140]. Then $C_G(A) = \{x \in G\colon xa = ax, \text{ all } a \in A\} = \cap \{x\colon c_a(x) = x\}$, the intersection being taken over all elements of A. Hence, $C_G(A)$ is closed, being an intersection of closed sets.

Letting $A = G$, we obtain as a corollary the standard result that the center of a Hausdorff group is closed.

THEOREM. *If A is a closed subgroup of a Hausdorff group G, then $N_G(A)$ is closed.*

Proof. We observe that $N_G(A) = \{x\colon c_a(x)A = xA, \text{ each } a \in A\}$. If $x \in N_G(A)$, then $c_a(x)A = axa^{-1}A = axA = xc_{x^{-1}}(a)A = xA$. Conversely, if $c_a(x)A = xA$ for all $a \in A$, then $axa^{-1}A = xA$, whence $x^{-1}axa^{-1} \in A$. Thus, $x^{-1}ax \in A$, for each $a \in A$, and so $x \in N_G(A)$.

Let q denote the (continuous) quotient map from G to its space of left cosets G/A, which is Hausdorff since A is closed. Then, $N_G(A) = \{x\colon c_a(x)A = xA, \text{ all } a \in A\} = \cap \{x\colon qc_a(x) = q(x)\}$, the intersection being taken over all elements of A. By the same theorem from [**1**], it follows that $N_G(A)$ is closed.

References

1. J. Dugundji, Topology, Allyn and Bacon, Boston, 1966.
2. T. Husain, Introduction to Topological Groups, Saunders, Philadelphia, 1966.
3. L. Pontryagin, Topological Groups, Princeton University Press, Princeton, 1946.

A NOTE ON DEMAR'S "A SIMPLE APPROACH TO ISOPERIMETRIC PROBLEMS IN THE PLANE" AND AN EPILOGUE

ALFRED D. GARVIN, University of Cincinnati

In his article on "... *Isoperimetric Problems* ..." (this MAGAZINE, Jan-Feb, 1975) DeMar presented, *inter alia*, the following problem and its solution (paraphrased for continuity):

PROBLEM. Given a triangular region T with perimeter P and with circumference of the inscribed circle equal to c, and given a number p such that $c < p < P$, among all regions of perimeter p contained in T, find the region R of maximum area.

Solution. [This] region R of maximum area has a boundary consisting of three circular arcs, all of the same radius, each tangent to two adjacent sides of the boundary of T, together with the three segments of sides of T between the endpoints of these arcs (Figure 1).

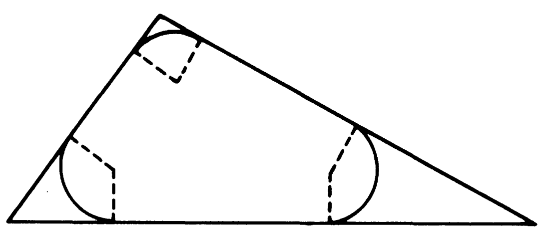


FIG. 1.

Before presenting his proof of this solution Professor DeMar was kind enough to acknowledge his gratitude to this author for "suggesting this problem and for conjecturing the correct solution which he arrived at *by physical experiments* " [emphasis supplied]. It might be of some interest to the readers of his article to know just what kinds of physical experiments led this author to conjecture this solution.

The first experiment was based on the theory that a horizontal plane liquid film under high surface tension and confined within a triangular region would assume an isoperimetrically optimum configuration. A small triangular reservoir having a flat bottom and vertical sides was formed of glass slabs laid on a horizontal glass plate. Liquid mercury was poured slowly into this reservoir. Predictably, the mercury formed a circular puddle that grew until it formed the inscribed circle of that triangle. With further pouring, this puddle flattened against the sides of this reservoir at the three points of tangency while three generally circular arcs of shrinking radii pushed toward the interior angles of the triangle. The configuration of these arcs was carefully observed and traced with a grease pencil at many intermediate stages of the puddle's growth.

The second experiment was based on the same theory as the first one. Here, however, the triangular reservoir was partially filled with water and then a suitably viscous oil was poured slowly onto this water. The same kinds of observations described in the first experiment were made again here and the same general configural phenomena were seen.

The third experiment was based on the theory that a flexible band of material bounding a region subjected to internal horizontal pressure and confined within a triangular region would also assume an isoperimetrically optimum configuration. A triangular fence resembling an asymmetrical billiard ball rack was constructed of wood. A half-inch wide strip of sturdy paper of length p intermediate between c and P, to use DeMar's notation, was stapled end-to-end to form a closed fence and this was deployed within the triangular fence. Hundreds of BB shot pellets were poured into the region enclosed by this flexible fence and pressed down until no more could be contained. Here, the flexible fence also flattened against the sides of the rigid bounding triangular fence while bulging into the internal angles of the triangle in three generally circular arcs. These arcs, too, were carefully observed and traced.

All of the observations made in all three experiments indicated that the arcs formed were essentially circular. Trial-and-error tests with a compass confirmed this indication. In the two surface tension experiments, a meniscus effect (convex for mercury, concave for oil) deformed the shape of the endpoints of these arcs. The third internal pressure experiment indicated that the endpoints of these arcs were, indeed, tangent to the sides of the triangular region.

Three hypotheses suggested themselves regarding the relationships among the three circular arcs formed in any of these physically formed regions:

1. They had equal radii.
2. They had equal chords.
3. They had equal length.

Mensurational formulas were developed for finding the area of any such confined region with a given perimeter p ($c < p < P$) according to each of these three hypotheses. Examples were calculated for several values of p within each of several configurations of triangular regions T. Hypothesis 1 yielded the maximum area for any given p and configuration of T in every case except (of course) that of an equilateral T where all three hypotheses yielded the same area.

Thus it was that this author, who is obviously not a mathematician, conjectured the solution later proved by DeMar.

Epilogue

The problem discussed here was merely part of a larger problem being pursued by this author. Using DeMar's notation, the larger problem was this:

PROBLEM. Given a triangular region T with perimeter P and with circumference of the inscribed circle equal to c, among all regions contained in T, find that region R having the maximum *ratio of area to perimeter.*

It was readily determined that the perimeter p of such a region must be such that $c < p < P$. The next step in this problem was to determine the general configuration of such a region. The procedures described above gave an answer to that problem, so far as this author was concerned. Taking this answer as an assumption, he soon solved his larger problem by equally farfetched methods. Extensions of this problem were conceived and solved. A great cathedral of conjecture was constructed.

At this point this author had the good sense to seek out Professor DeMar to verify his conjectures and, if they were verified, to provide the undergirdings of rigorous mathematical proofs. As it happened, DeMar verified them all and a collaboration was begun to transform this tissue of conjecture into a piece worthy of respectable publication.

Regrettably, Professor Richard F. DeMar died shortly after the article to which this note refers was published. So far as is known, that article will be the last publication of his prolific scholarly career. As one of the very minor consequences of his passing, the publication of this author's "larger" work will be delayed until he is able to enlist the assistance of another so generous a scholar.

ON REPRESENTING INTEGERS AS SUMS OF ODD COMPOSITE INTEGERS

A. M. VAIDYA, Gujarat University, Ahmedabad 9, India

Recently, Just and Schaumberger [1] have discovered an interesting property of 38. They have proved that 38 is the largest even integer which cannot be expressed as a sum of two composite odd integers. In this note we extend this result in two different directions.

THEOREM 1. *For each positive integer t, the integer $9t + 20$ cannot be expressed as a sum of t composite odd integers.*

Proof. It is easy to verify the theorem for $t = 1,2$ and 3. We can therefore start induction. Suppose the theorem is true for $t = k - 1$, where $k \geqq 4$.

If possible, suppose $9k + 20$ is a sum of k odd composite numbers. We claim that at least one of these odd composite numbers must be 9; for otherwise (as 9 and 15 are the two least odd composite numbers) each of the k odd composite numbers must be $\geqq 15$. But then their sum would be $\geqq 15k$, i.e., $9k + 20 \geqq 15k$ which is false because $k \geqq 4$.

PROBLEM. Given a triangular region T with perimeter P and with circumference of the inscribed circle equal to c, among all regions contained in T, find that region R having the maximum *ratio of area to perimeter.*

It was readily determined that the perimeter p of such a region must be such that $c < p < P$. The next step in this problem was to determine the general configuration of such a region. The procedures described above gave an answer to that problem, so far as this author was concerned. Taking this answer as an assumption, he soon solved his larger problem by equally farfetched methods. Extensions of this problem were conceived and solved. A great cathedral of conjecture was constructed.

At this point this author had the good sense to seek out Professor DeMar to verify his conjectures and, if they were verified, to provide the undergirdings of rigorous mathematical proofs. As it happened, DeMar verified them all and a collaboration was begun to transform this tissue of conjecture into a piece worthy of respectable publication.

Regrettably, Professor Richard F. DeMar died shortly after the article to which this note refers was published. So far as is known, that article will be the last publication of his prolific scholarly career. As one of the very minor consequences of his passing, the publication of this author's "larger" work will be delayed until he is able to enlist the assistance of another so generous a scholar.

ON REPRESENTING INTEGERS AS SUMS OF ODD COMPOSITE INTEGERS

A. M. VAIDYA, Gujarat University, Ahmedabad 9, India

Recently, Just and Schaumberger [1] have discovered an interesting property of 38. They have proved that 38 is the largest even integer which cannot be expressed as a sum of two composite odd integers. In this note we extend this result in two different directions.

THEOREM 1. *For each positive integer t, the integer $9t + 20$ cannot be expressed as a sum of t composite odd integers.*

Proof. It is easy to verify the theorem for $t = 1, 2$ and 3. We can therefore start induction. Suppose the theorem is true for $t = k - 1$, where $k \geqq 4$.

If possible, suppose $9k + 20$ is a sum of k odd composite numbers. We claim that at least one of these odd composite numbers must be 9; for otherwise (as 9 and 15 are the two least odd composite numbers) each of the k odd composite numbers must be $\geqq 15$. But then their sum would be $\geqq 15k$, i.e., $9k + 20 \geqq 15k$ which is false because $k \geqq 4$.

Thus if $9k+20$ is a sum of k odd composite numbers, then one of them must be 9. Taking this term away, we get that $9k+20-9=9(k-1)+20$ is a sum of $(k-1)$ odd composite numbers contradicting the induction hypothesis. This proves the theorem.

THEOREM 2. *If $t \geqq 2$ is an integer and if $n > 9t+20$ has the same parity as t, then n can be expressed as the sum of t composite odd integers.*

Proof. Since $n > 9t+20$ and n has the same parity as t, we must have

$$n = 9t + 20 + 2m$$

for some positive integer m.

By Just and Schaumberger's result, the integer $2m+38$ is a sum of two odd composite numbers, say u and v. Then

$$n = 9t + 20 + 2m = 9(t-2) + (2m+38)$$

$$= 9 + 9 + 9 + 9 + \cdots + 9 + u + v,$$

where the number of 9's is $t-2$. Thus n is a sum of t odd composite integers.

Theorems 1 and 2 can be combined to yield the following extension of J and S's result:

THEOREM 3. *Let $t \geqq 2$ be an integer. Then among all the integers with the same parity as t, $9t+20$ is the largest which cannot be expressed as the sum of t odd composite integers.*

If one examines J and S's proof of their result, one observes that they have actually proved more than their assertion. It is proved that every even integer > 38 is a sum of two composite odd integers one of which is a multiple of 3 and the other is a multiple of 5. This suggests the following question: Suppose distinct odd primes p and q are given, what is the largest even integer (if any) which cannot be expressed as a sum of two composite odd integers, one of which is a multiple of p and the other is a multiple of q? We answer this question.

The following result is well known (see [**2**] or [**3**]):

LEMMA 1. *If p and q are relatively prime positive integers, then the largest integer not expressible in the form $pr+qs$ with r and s nonnegative integers is $pq-p-q$.*

This lemma helps us to prove the following generalization of J and S's result:

THEOREM 4. *If p and q are distinct odd primes, the largest even integer not expressible as a sum of an odd composite multiple of p and an odd composite multiple of q is $2pq+p+q$.*

Proof. An even integer $2n$ has an expression of the specified kind iff there exist integers $r \geqq 0$ and $s \geqq 0$ such that

$$p(2r+3) + q(2s+3) = 2n,$$

i.e., iff there exist integers $r \geqq 0$ and $s \geqq 0$, such that $pr + qs = n - \frac{3}{2}(p+q)$. Hence by the above lemma, the largest even integer $2n$ which cannot be expressed in the form $p(2r+3) + q(2s+3)$ corresponds to

$$n - \tfrac{3}{2}(p+q) = pq - p - q,$$

i.e.,

$$2n = 2pq + p + q.$$

Note. Putting $p = 3$, $q = 5$, we get J and S's result.

Shah [4] has generalized Lemma 1 as follows:

LEMMA 2. *Let p and q be relatively prime positive integers and k a fixed positive integer. Then the largest integer which has less than k representations in the form $pr + qs$, $r \geqq 0$, $s \geqq 0$ is $kpq - p - q$.*

We can now prove the following extension of J and S's result in the same way as Theorem 4:

THEOREM 5. *The largest even integer which has less than k representations as a sum of an odd composite multiple of p and an odd composite multiple of q is*

$$2kpq + p + q.$$

Let us call a representation of a number as a sum of two composite odd numbers a standard representation. For a positive integer k, let $F(k)$ be the largest integer not having k standard representations, then numerical evidence suggests that

$$F(1) = 38, \quad F(2) = F(3) = 68, \quad F(4) = 94,$$

$$F(5) = 122, \quad F(6) = F(7) = 128, \quad F(8) = 136.$$

References

1. Erwin Just and Norman Schaumberger, A curious property of the integer 38, this MAGAZINE, 46 (1973) 221.
2. W. LeVeque, Topics in Number Theory, vol. I., Addison-Wesley, Reading, 1956, p. 22.
3. Problem E 1637, Amer. Math. Monthly, 70 (1963) 1005, and 71 (1964) 799.
4. A. P. Shah, Contribution to.... a problem of Frobenius, Ph.D. thesis submitted to Gujarat University, 1970.

A FIXED POINT THEOREM

B. FISHER, University of Leicester, England

In a paper by R. Kannan [1] he proved the following theorem:

THEOREM 1. *If T is a mapping of the complete metric space X into itself satisfying the condition*

$$\rho(Tx, Ty) \leqq c[\rho(x, Tx) + \rho(y, Ty)]$$

for all x, y in X, where $0 \leqq c < \frac{1}{2}$, then T has a unique fixed point.

i.e., iff there exist integers $r \geqq 0$ and $s \geqq 0$, such that $pr + qs = n - \frac{3}{2}(p+q)$. Hence by the above lemma, the largest even integer $2n$ which cannot be expressed in the form $p(2r+3) + q(2s+3)$ corresponds to

$$n - \tfrac{3}{2}(p+q) = pq - p - q,$$

i.e.,

$$2n = 2pq + p + q.$$

Note. Putting $p = 3$, $q = 5$, we get J and S's result.

Shah [4] has generalized Lemma 1 as follows:

LEMMA 2. *Let p and q be relatively prime positive integers and k a fixed positive integer. Then the largest integer which has less than k representations in the form $pr + qs$, $r \geqq 0$, $s \geqq 0$ is $kpq - p - q$.*

We can now prove the following extension of J and S's result in the same way as Theorem 4:

THEOREM 5. *The largest even integer which has less than k representations as a sum of an odd composite multiple of p and an odd composite multiple of q is*

$$2kpq + p + q.$$

Let us call a representation of a number as a sum of two composite odd numbers a standard representation. For a positive integer k, let $F(k)$ be the largest integer not having k standard representations, then numerical evidence suggests that

$$F(1) = 38, \quad F(2) = F(3) = 68, \quad F(4) = 94,$$

$$F(5) = 122, \quad F(6) = F(7) = 128, \quad F(8) = 136.$$

References

1. Erwin Just and Norman Schaumberger, A curious property of the integer 38, this MAGAZINE, 46 (1973) 221.
2. W. LeVeque, Topics in Number Theory, vol. I., Addison-Wesley, Reading, 1956, p. 22.
3. Problem E 1637, Amer. Math. Monthly, 70 (1963) 1005, and 71 (1964) 799.
4. A. P. Shah, Contribution to.... a problem of Frobenius, Ph.D. thesis submitted to Gujarat University, 1970.

A FIXED POINT THEOREM

B. FISHER, University of Leicester, England

In a paper by R. Kannan [1] he proved the following theorem:

THEOREM 1. *If T is a mapping of the complete metric space X into itself satisfying the condition*

$$\rho(Tx, Ty) \leqq c[\rho(x, Tx) + \rho(y, Ty)]$$

for all x, y in X, where $0 \leqq c < \frac{1}{2}$, then T has a unique fixed point.

We will now prove the following similar type of theorem:

THEOREM 2. *If T is a mapping of the complete metric space X into itself satisfying the condition*

$$\rho(Tx, Ty) \leqq c[\rho(x, Ty) + \rho(y, Tx)]$$

for all x, y in X, where $0 \leqq c < \frac{1}{2}$, then T has a unique fixed point.

Proof. Let x be an arbitrary point in X. Then

$$\begin{aligned} \rho(T^n x, T^{n+1}x) &\leqq c[\rho(T^{n-1}x, T^{n+1}x) + \rho(T^n x, T^n x)] \\ &= c\rho(T^{n-1}x, T^{n+1}x) \\ &\leqq c[\rho(T^{n-1}x, T^n x) + \rho(T^n x, T^{n+1}x)]. \end{aligned}$$

Thus

$$\begin{aligned} \rho(T^n x, T^{n+1}x) &\leqq \frac{c}{1-c}\rho(T^{n-1}x, T^n x) \\ &\leqq \left(\frac{c}{1-c}\right)^2 \rho(T^{n-2}x, T^{n-1}x) \\ &\leqq \left(\frac{c}{1-c}\right)^n \rho(x, Tx). \end{aligned}$$

Hence

$$\begin{aligned} \rho(T^n x, T^{n+r}x) &\leqq \rho(T^n x, T^{n+1}x) + \cdots + \rho(T^{n+r-1}x, T^{n+r}x) \\ &\leqq \left[\left(\frac{c}{1-c}\right)^n + \cdots + \left(\frac{c}{1-c}\right)^{n+r-1}\right]\rho(x, Tx) \\ &\leqq \left(\frac{c}{1-c}\right)^n \frac{1-c}{1-2c}\rho(x, Tx). \end{aligned}$$

Since $c(1-c)^{-1} < 1$ it follows that $\{T^n x\}$ is a Cauchy sequence in X and so has a limit z in X, since X is complete.

We now have

$$\begin{aligned} \rho(z, Tz) &\leqq \rho(z, T^n x) + \rho(T^n x, Tz) \\ &\leqq \rho(z, T^n x) + c[\rho(T^{n-1}x, Tz) + \rho(T^n x, z)]. \end{aligned}$$

Letting n tend to infinity we see that

$$\rho(z, Tz) \leqq c\rho(z, Tz)$$

and since $c < \frac{1}{2}$ it follows that

$$Tz = z.$$

Hence z is a fixed point.

Now suppose T has a second fixed point z'. Then

$$\begin{aligned}\rho(z,z') &= \rho(Tz,Tz') \\ &\leqq c[\rho(z,Tz')+\rho(z',Tz)] \\ &= 2c\rho(z,z').\end{aligned}$$

Since $c<\frac{1}{2}$ it follows that $z=z'$ and so the fixed point is unique.

Reference

1. R. Kannan, Some results on fixed points, Bull. Calcutta Math. Soc., 60 (1968) 71–6.

ON THE SUBSEMIGROUPS OF N

WILLIAM Y. SIT, City College of New York and
MAN-KEUNG SIU, University of Miami

By a subsemigroup M of $\mathbf{N}$, the additive semigroup of natural numbers, we mean a subset M of $\mathbf{N}$ such that $x+y\in M$ whenever $x,y\in M$. The subsemigroup $\{0\}$ is trivial and shall be excluded from discussion henceforth. In this note, we characterize all (nontrivial) subsemigroups of $\mathbf{N}$ by the divisibility properties of their elements. More precisely, for any $d\in\mathbf{N}$ with $d\geqq 1$ we say a subset M is *ultimately a d-segment* if there is some $N\in\mathbf{N}$ such that for $x\geqq N$, we have $x\in M$ if and only if $d\mid x$. We note that M is a 1-segment means M contains all integers $x\geqq N$ for some positive integer N. Our main result may now be stated.

THEOREM. *Every subsemigroup M of $\mathbf{N}$ is ultimately a d-segment with $d=\min_{x,y\in M}\{(x,y)\}$, where (x,y) denotes the highest common factor of x and y.*

Proof. Let $d=\min_{x,y\in M}\{(x,y)\}$. Then $d=(x_0,y_0)$ for some $x_0,\ y_0\in M$ and hence $d=ax_0+by_0$ for some integers a,b. Choose integers $a'\geqq 0$ and $b'\geqq 0$ such that $a'+a\geqq 0$ and $b'+b\geqq 0$. Then $s=a'x_0+b'y_0\in M$ and $s+d=(a'+a)x_0+(b'+b)y_0\in M$. Notice that $s=\alpha d$ for some $\alpha\in\mathbf{N}$. We put $N=\alpha^2 d$ and let $x\geqq N$. If $d\mid x$, then $x=\beta d$ where $\beta\geqq\alpha^2$. We now claim that $x\in M$. We first observe that for $k\geqq 1$, $(k\alpha+r)d\in M$ for $0\leqq r\leqq k-1$. We show this by induction on k. The case $k=1$ is just $s=\alpha d\in M$. Suppose $k>1$, then for $0\leqq r\leqq k$, we have

$$[(k+1)\alpha+r]d=(k\alpha+r-1)d+(\alpha+1)d\in M$$

by our induction assumption. Now let $\beta-\alpha^2=q\alpha+r$ where $0\leqq r<\alpha$; then $\beta=(\alpha+q)\alpha+r$ and $0\leqq r<\alpha+q$ so that $x=\beta d\in M$ by putting $k=\alpha+q$. This proves our claim.

Now suppose T has a second fixed point z'. Then

$$\begin{aligned}\rho(z,z') &= \rho(Tz,Tz') \\ &\leqq c[\rho(z,Tz')+\rho(z',Tz)] \\ &= 2c\rho(z,z').\end{aligned}$$

Since $c<\frac{1}{2}$ it follows that $z=z'$ and so the fixed point is unique.

Reference

1. R. Kannan, Some results on fixed points, Bull. Calcutta Math. Soc., 60 (1968) 71–6.

ON THE SUBSEMIGROUPS OF N

WILLIAM Y. SIT, City College of New York and
MAN-KEUNG SIU, University of Miami

By a subsemigroup M of **N**, the additive semigroup of natural numbers, we mean a subset M of **N** such that $x+y\in M$ whenever $x,y\in M$. The subsemigroup $\{0\}$ is trivial and shall be excluded from discussion henceforth. In this note, we characterize all (nontrivial) subsemigroups of **N** by the divisibility properties of their elements. More precisely, for any $d\in\mathbf{N}$ with $d\geqq 1$ we say a subset M is *ultimately a d-segment* if there is some $N\in\mathbf{N}$ such that for $x\geqq N$, we have $x\in M$ if and only if $d\mid x$. We note that M is a 1-segment means M contains all integers $x\geqq N$ for some positive integer N. Our main result may now be stated.

THEOREM. *Every subsemigroup M of* **N** *is ultimately a d-segment with* $d=\min_{x,y\in M}\{(x,y)\}$, *where (x,y) denotes the highest common factor of x and y.*

Proof. Let $d=\min_{x,y\in M}\{(x,y)\}$. Then $d=(x_0,y_0)$ for some x_0, $y_0\in M$ and hence $d=ax_0+by_0$ for some integers a,b. Choose integers $a'\geqq 0$ and $b'\geqq 0$ such that $a'+a\geqq 0$ and $b'+b\geqq 0$. Then $s=a'x_0+b'y_0\in M$ and $s+d=(a'+a)x_0+(b'+b)y_0\in M$. Notice that $s=\alpha d$ for some $\alpha\in\mathbf{N}$. We put $N=\alpha^2d$ and let $x\geqq N$. If $d\mid x$, then $x=\beta d$ where $\beta\geqq\alpha^2$. We now claim that $x\in M$. We first observe that for $k\geqq 1$, $(k\alpha+r)d\in M$ for $0\leqq r\leqq k-1$. We show this by induction on k. The case $k=1$ is just $s=\alpha d\in M$. Suppose $k>1$, then for $0\leqq r\leqq k$, we have

$$[(k+1)\alpha+r]d=(k\alpha+r-1)d+(\alpha+1)d\in M$$

by our induction assumption. Now let $\beta-\alpha^2=q\alpha+r$ where $0\leqq r<\alpha$; then $\beta=(\alpha+q)\alpha+r$ and $0\leqq r<\alpha+q$ so that $x=\beta d\in M$ by putting $k=\alpha+q$. This proves our claim.

Conversely, if $x \in M$, we shall show that $d \mid x$. If not, then $\beta d < x < (\beta + 1)d$ for some $\beta \geqq \alpha^2$. Let $(\beta d, x) = d'$ and $x = \beta d + x'$. Then $0 < x' < d$ and $d' \mid x'$; in particular $d' \leqq x' < d$. This contradicts the minimality of d and completes the proof that M is ultimately a d-segment.

Thus every subsemigroup M of $\mathbf{N}$ consists of almost all nonnegative multiples of a certain number d. This number d can also be characterized in another nice way (as pointed out to the authors by the referee of this paper): let $M^* = \{x - y \mid x, y \in M \cup \{0\}\}$, then it is immediate that M^* is an ideal in $\mathbf{Z}$ so that $M^* = (d^*)$ for some $d^* \in \mathbf{N}$; since M is ultimately a d-segment and $M \subset M^*$, $d = d^*$, the smallest positive number which can be written as the difference of two elements in M.

Given a subsemigroup M, we say a subset $X \subset M$ is a set of generators of M if every number in M can be expressed as the sum of nonnegative multiples of elements of X. We say M is *finitely generated* if such a subset X exists and is finite. An important corollary of the theorem is

COROLLARY 1. *Any subsemigroup M of $\mathbf{N}$ is finitely generated.*

Proof. This corollary comes from the proof of the theorem rather than the theorem itself. We see from the proof that M is generated by αd, $(\alpha + 1)d$ and those $x \in M$ with $x \leqq \alpha^2 d$.

The next two corollaries are of mild interest.

COROLLARY 2. *A subsemigroup M of $\mathbf{N}$ is ultimately a 1-segment if and only if there are x, $y \in M$ with $(x, y) = 1$.*

Proof. Clear.

COROLLARY 3. *Given positive integers m and n, a necessary and sufficient condition that there exists a positive integer N such that every integer $x \geqq N$ can be written as $x = am + bn$ with a, $b \in \mathbf{N}$ is that m and n be relatively prime.*

Proof. Consider the subsemigroup $M = \{am + bn \mid a, b \in \mathbf{N}\}$ generated by m and n. If $(m, n) = 1$, then by Corollary 2, M is ultimately a 1-segment. Conversely if $(m, n) = d \neq 1$, then $\min_{x,y \in M} \{(x, y)\} = d > 1$ and M cannot be ultimately a 1-segment.

In closing, we would like to discuss two related questions, both of which were brought to our attention by Nan-Shan Shou.

QUESTION 1. Let $x_1, \cdots, x_k$ be k natural numbers and let $M = M(x_1, \cdots, x_k)$ be the subsemigroup generated by $x_1, \cdots, x_k$. Since M is ultimately a d-segment for some $d \in \mathbf{N}$, it would be interesting to find the smallest number $N = N(x_1, \cdots, x_k) \in \mathbf{N}$ such that for all $s \geqq N$, $sd \in M$. When $x_1, \cdots, x_k$ are relatively prime, N. S. Mendelsohn [Problem E2247, p. 677, AMERICAN MATHEMATICAL MONTHLY, June-July issue, 1971] proved certain properties of N and in particular when $k = 2$, he showed that $N(m, n) = (m - 1)(n - 1)$. Both the referee and Shou have found other proofs in this special case and the reader may well discover his own; the referee also noted that $N(m,n)=(m_1-1)(m_2-1)d$ where $m=dm_1$, $n=dn_1$ and $(m,n)=d$. Our Corollary 3 is then an easy consequence of this result. The general question seems unsolved.

QUESTION 2. Given $(m,n)=1$, consider the subring $k[X^m,X^n]$ of the polynomial ring $k[X]$ over a field k. If $N=N(m,n)$ as defined in Question 1, it can be shown that the principal ideal (X^N) is the conductor, that is, the largest common ideal in $k[X^m,X^n]$ and $k[X]$. Since $(m,n)=1$, $k[X^m,X^n]$ is isomorphic to the coordinate ring $k[X,Y]/(X^m-Y^n)$ of the irreducible curve $f(X,Y)=X^m-Y^n$, while $k[X]$ is the coordinate ring of the affine line. The conductor (X^N) then has certain geometric interpretation. Our question is: Is there any geometric significance attached to the number $N=(m-1)(n-1)$?

A BINOMIAL IDENTITY DERIVED FROM A MATHEMATICAL MODEL OF THE WORLD SERIES

PEGGY TANG STRAIT, Queens College, CUNY

Let $\binom{n}{k}$ denote the binomial coefficient

$$\binom{n}{k}=n!/k!(n-k)!$$

for nonnegative integers k and n with $k\leqq n$, and let $b(k;n,p)$ denote the kth term of the binomial distribution

$$b(k;n,p)=\binom{n}{k}p^k(1-p)^{n-k}$$

for $k=0,1,\cdots,n$ and $0\leqq p\leqq 1$. As we know from elementary probability, the numerical value of the binomial term $b(k;n,p)$ is the probability of exactly k successes in n independent trials with the assumption that the probability of success in each trial is p. We shall show that the following well-known [1] binomial identity (Theorem 1) may be derived from the mathematics of a game of chance.

THEOREM 1.

$$\sum_{r=0}^{k} b(k;k+r,1/2)=1$$

or equivalently,

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

for positive integers k.

own; the referee also noted that $N(m,n) = (m_1 - 1)(m_2 - 1)d$ where $m = dm_1$, $n = dn_1$ and $(m,n) = d$. Our Corollary 3 is then an easy consequence of this result. The general question seems unsolved.

QUESTION 2. Given $(m,n) = 1$, consider the subring $k[X^m, X^n]$ of the polynomial ring $k[X]$ over a field k. If $N = N(m,n)$ as defined in Question 1, it can be shown that the principal ideal (X^N) is the conductor, that is, the largest common ideal in $k[X^m, X^n]$ and $k[X]$. Since $(m,n) = 1$, $k[X^m, X^n]$ is isomorphic to the coordinate ring $k[X, Y]/(X^m - Y^n)$ of the irreducible curve $f(X, Y) = X^m - Y^n$, while $k[X]$ is the coordinate ring of the affine line. The conductor (X^N) then has certain geometric interpretation. Our question is: Is there any geometric significance attached to the number $N = (m-1)(n-1)$?

A BINOMIAL IDENTITY DERIVED FROM A MATHEMATICAL MODEL OF THE WORLD SERIES

PEGGY TANG STRAIT, Queens College, CUNY

Let $\binom{n}{k}$ denote the binomial coefficient

$$\binom{n}{k} = n!/k!(n-k)!$$

for nonnegative integers k and n with $k \leqq n$, and let $b(k; n, p)$ denote the kth term of the binomial distribution

$$b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k = 0, 1, \cdots, n$ and $0 \leqq p \leqq 1$. As we know from elementary probability, the numerical value of the binomial term $b(k; n, p)$ is the probability of exactly k successes in n independent trials with the assumption that the probability of success in each trial is p. We shall show that the following well-known [1] binomial identity (Theorem 1) may be derived from the mathematics of a game of chance.

THEOREM 1.

$$\sum_{r=0}^{k} b(k; k+r, 1/2) = 1$$

or equivalently,

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

for positive integers k.

Proof. Consider the following game of chance. Team A plays a series of games with Team B. The first team to win $k+1$ games is declared the winner of the series. (This is a generalization of the World Series in baseball.) If p is the probability that Team A wins a single game of the series and the games are assumed to be independent trials, then the probability that Team A wins the series may be computed as follows.

Prob. (team A wins the series)

$$= \sum_{r=0}^{k} [\text{Prob. \{team A wins exactly } k \text{ of the first } (k+r) \text{ games\}}$$

$$\cdot \text{Prob. \{team A wins the } (k+r+1)\text{st game\}]}$$

$$= \sum_{r=0}^{k} b(k; k+r, p)p.$$

Observe that the probabilities added above to obtain the probability that team A wins the series are of disjoint events so that it is appropriate to add them. In the case where $p=\frac{1}{2}$, it is clear that

$$\text{Prob. (team A wins the series)} = \text{Prob. (team B wins the series)} = \tfrac{1}{2}.$$

Thus, we have

$$\sum_{r=0}^{k} b(k; k+r, \tfrac{1}{2})\tfrac{1}{2} = \tfrac{1}{2}$$

or equivalently,

$$\sum_{r=0}^{k} b(k; k+r, \tfrac{1}{2}) = 1,$$

which is the statement of the theorem.

Reference

1. H. W. Gould, Combinatorial identities, a standardized set of tables listing 500 binomial coefficient summations, West Virginia University.

THE ELLIPSE AS AN HYPOTROCHOID

DAN PEDOE, University of Minnesota

If a triangular cutout ABC moves with the vertex A on a given line l, and the vertex B on a given line m, the vertex C describes an ellipse. As a special case, if the lines l and m are at right angles, and C is a fixed point on the segment AB or AB produced, and A moves on l and B moves on m, then the lines l and m are the principal axes of the ellipse. The first fact was certainly known to Leonardo

Proof. Consider the following game of chance. Team A plays a series of games with Team B. The first team to win $k+1$ games is declared the winner of the series. (This is a generalization of the World Series in baseball.) If p is the probability that Team A wins a single game of the series and the games are assumed to be independent trials, then the probability that Team A wins the series may be computed as follows.

Prob. (team A wins the series)

$$= \sum_{r=0}^{k} [\text{Prob. \{team A wins exactly } k \text{ of the first } (k+r) \text{ games\}}$$

$$\cdot \text{Prob. \{team A wins the } (k+r+1)\text{st game\}}]$$

$$= \sum_{r=0}^{k} b(k; k+r, p)p.$$

Observe that the probabilities added above to obtain the probability that team A wins the series are of disjoint events so that it is appropriate to add them. In the case where $p = \frac{1}{2}$, it is clear that

$$\text{Prob. (team A wins the series)} = \text{Prob. (team B wins the series)} = \tfrac{1}{2}.$$

Thus, we have

$$\sum_{r=0}^{k} b(k; k+r, \tfrac{1}{2})\tfrac{1}{2} = \tfrac{1}{2}$$

or equivalently,

$$\sum_{r=0}^{k} b(k; k+r, \tfrac{1}{2}) = 1,$$

which is the statement of the theorem.

Reference

1. H. W. Gould, Combinatorial identities, a standardized set of tables listing 500 binomial coefficient summations, West Virginia University.

THE ELLIPSE AS AN HYPOTROCHOID

DAN PEDOE, University of Minnesota

If a triangular cutout ABC moves with the vertex A on a given line l, and the vertex B on a given line m, the vertex C describes an ellipse. As a special case, if the lines l and m are at right angles, and C is a fixed point on the segment AB or AB produced, and A moves on l and B moves on m, then the lines l and m are the principal axes of the ellipse. The first fact was certainly known to Leonardo da Vinci [1]. Even liberal arts students, when asked to verify these facts by drawing, spend hours trying to discover how the ellipse which emerges as a result of their efforts depends on the triangle ABC and the given lines. Some discover that the locus of C is sometimes a line. The following solution comes into the domain of roulettes and glissettes, absent from modern textbooks on geometry, but probably to be found in texts on engineering mathematics.

The key to the da Vinci locus problem is the fact that if circle $\mathscr{C}$ rolls on the inside of a fixed circle $\mathscr{D}$ of twice its radius, then a given point on the circumference of $\mathscr{C}$ moves on a straight line which passes through the center O of $\mathscr{D}$. This follows immediately from the notion of rolling, and the theorem that the angle subtended at the center of a circle is twice that subtended at a point on the circumference. Given a triangle ABC [Figure 1] we merely have to find the

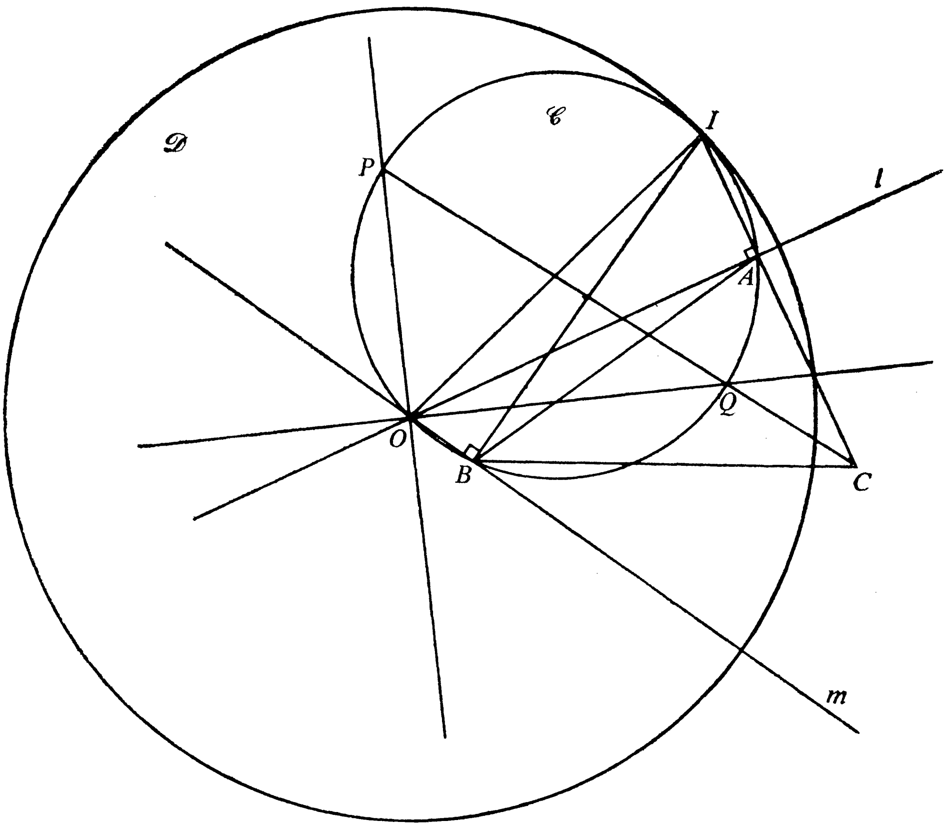


FIG. 1.

circles $\mathscr{C}$ and $\mathscr{D}$. Draw perpendiculars, to l at A and to m at B, and let I be their intersection. The circle $\mathscr{C}$ on OI as diameter passes through A and B. Draw the circle $\mathscr{D}$ center O which passes through I. Then the respective motions of A on l and of B on m are produced by the rolling of circle $\mathscr{C}$ inside the fixed circle $\mathscr{D}$. Since the vertex C is rigidly attached to both A and B, the motion of the point C is given by this rolling.

Let the join of C to the center of circle $\mathscr{C}$ intersect this circle in P and Q. Then C is a fixed point on the fixed segment PQ, or PQ produced, where P moves on OP and Q moves on the perpendicular line OQ. Hence C describes an ellipse with principal axes OP and OQ. If triangle ABC is such that C lies either on OP or OQ, then C moves on that line.

Reference

1. Dan Pedoe, Geometry and the Liberal Arts, Penguin Books, London, in the press.

AN EXPLICIT FORMULA FOR THE *k*th PRIME NUMBER

STEPHEN REGIMBAL, Billings, Montana

This paper will develop an explicit formula for the kth prime, where k is any positive integer. The formula will be elementary, finite, and dependent only upon the choice of k.

First consider the function

$$g(n) = \sum_{i=1}^{n-1} \left[\frac{\left[\frac{n}{i}\right]}{\frac{n}{i}} \right]$$

where $[x]$ in all cases denotes the greatest integer $\leqq x$, and $n \geqq 2$. Now the terms in this finite sum are equal either to 0 or to 1 since for any $x > 0$,

$$0 < \frac{[x]}{x} \leqq 1 \Rightarrow \left[\frac{[x]}{x}\right] = 0 \quad \text{or} \quad 1.$$

Therefore, $g(n) = 1$ if and only if there exists exactly one $i \in \{1, 2, 3, \cdots, n-1\}$ such that

$$\left[\frac{\left[\frac{n}{i}\right]}{\frac{n}{i}} \right] = 1 \Leftrightarrow \left[\frac{n}{i}\right] = \frac{n}{i} \Leftrightarrow i \mid n;$$

but since $i = 1$ divides n for all positive integers n, then $g(n) = 1$ if and only if $i = 1$ is the only value of $i \in \{1, 2, 3, \cdots, n-1\}$ such that i divides n. But this means that n is prime. Thus, $g(n) = 1$ if and only if n is prime. Now if n is composite, then there exists some $i \in \{2, 3, 4, \cdots, n-1\}$ such that i divides n, and this implies that $g(n) > 1$. Therefore,

$$g(n) \begin{cases} = 1, & n \text{ prime} \\ > 1, & n, \text{ composite.} \end{cases}$$

Let the join of C to the center of circle $\mathscr{C}$ intersect this circle in P and Q. Then C is a fixed point on the fixed segment PQ, or PQ produced, where P moves on OP and Q moves on the perpendicular line OQ. Hence C describes an ellipse with principal axes OP and OQ. If triangle ABC is such that C lies either on OP or OQ, then C moves on that line.

Reference

1. Dan Pedoe, Geometry and the Liberal Arts, Penguin Books, London, in the press.

AN EXPLICIT FORMULA FOR THE kth PRIME NUMBER

STEPHEN REGIMBAL, Billings, Montana

This paper will develop an explicit formula for the kth prime, where k is any positive integer. The formula will be elementary, finite, and dependent only upon the choice of k.

First consider the function

$$g(n) = \sum_{i=1}^{n-1} \left[\frac{\left[\frac{n}{i}\right]}{\frac{n}{i}} \right]$$

where $[x]$ in all cases denotes the greatest integer $\leqq x$, and $n \geqq 2$. Now the terms in this finite sum are equal either to 0 or to 1 since for any $x > 0$,

$$0 < \frac{[x]}{x} \leqq 1 \Rightarrow \left[\frac{[x]}{x}\right] = 0 \quad \text{or} \quad 1.$$

Therefore, $g(n) = 1$ if and only if there exists exactly one $i \in \{1, 2, 3, \cdots, n-1\}$ such that

$$\left[\frac{\left[\frac{n}{i}\right]}{\frac{n}{i}} \right] = 1 \Leftrightarrow \left[\frac{n}{i}\right] = \frac{n}{i} \Leftrightarrow i \mid n;$$

but since $i = 1$ divides n for all positive integers n, then $g(n) = 1$ if and only if $i = 1$ is the only value of $i \in \{1, 2, 3, \cdots, n-1\}$ such that i divides n. But this means that n is prime. Thus, $g(n) = 1$ if and only if n is prime. Now if n is composite, then there exists some $i \in \{2, 3, 4, \cdots, n-1\}$ such that i divides n, and this implies that $g(n) > 1$. Therefore,

$$g(n) \begin{cases} = 1, & n \text{ prime} \\ > 1, & n, \text{ composite.} \end{cases}$$

Now let $f(n) = [1/(g(n))]$, then it is clear that for $n \geqq 2$,

$$f(n) = \begin{cases} 1, & n \text{ prime} \\ 0, & n \text{ composite.} \end{cases}$$

Note that $f(n)$ is the characteristic function of the set of primes.

Consider for $m \geqq 2$,

$$\pi(m) = \sum_{n=2}^{m} f(n);$$

since $f(n)$ is the characteristic function of the set of primes, then $\pi(m)$ is the number of primes less than or equal to m, and thus for $m \geqq 2$,

$$f(m)\pi(m) = \begin{cases} \text{number of primes } \leqq m, \ m \text{ prime,} \\ 0, \ m \text{ composite.} \end{cases}$$

Let k be any positive integer, and the following equality holds:

$$\left[\frac{1}{1+|k - f(m)\pi(m)|}\right] m = \begin{cases} m, \ m \text{ the } k\text{th prime,} \\ 0, \text{ otherwise.} \end{cases}$$

This is true since the above function is equal to m if and only if $k = f(m)\pi(m)$, which implies that m is prime and there are k primes less than or equal to m; but this means that m is the kth prime. If m is not the kth prime of if m is composite, then $|k - f(m)\pi(m)| > 0$, and the above function equals 0.

Now by Bertrand's postulate as proved by Hardy and Wright [**1**], there exists at least one prime p such that for all positive integers n, $n < p \leqq 2n$. Therefore, it follows that if p_k denotes the kth prime, then for all positive integers k, $p_k < p_{k+1} \leqq 2p_k$ and this supplies the induction step to prove that for all positive integers k, $2^k \geqq p_k$. Now with this result it follows that for all $k \geqq 1$, the kth prime is given by the formula

$$p_k = \sum_{m=2}^{2^k} \left[\frac{1}{1+|k - f(m)\pi(m)|}\right] m$$

since summing to 2^k guarantees that we have summed past the kth prime. Now rewriting in terms of the original expressions, we have:

$$p_k = \sum_{m=2}^{2^k} \left[\frac{1}{1 + \left| k - \left[\frac{1}{\sum_{i=1}^{m-1} \left[\frac{\left[\frac{m}{i}\right]}{\frac{m}{i}} \right]} \right] \sum_{n=2}^{m} \left[\frac{1}{\sum_{i=1}^{n-1} \left[\frac{\left[\frac{n}{i}\right]}{\frac{n}{i}} \right]} \right] \right|} \right] m.$$

Note that this method can be generalized to yield an expression for the kth smallest element of an arbitrary infinite subset A of the positive integers. The finiteness of this extended formula will depend not only upon the possibility of writing a finite expression for the characteristic function of the set A, but also upon the existence of a finitely expressible function $h(k)$ such that $h(k) \geqq$ the kth smallest element of A for all $k \geqq 1$.

Reference

1. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Oxford University Press, London, 1960, pp. 343–4.

THE GENERAL CAYLEY-HAMILTON THEOREM VIA THE EASIEST REAL CASE

J. DENMEAD SMITH, College of the Resurrection, Mirfield, Yorkshire, England.

Let A be a square matrix with elements in a field, and let $G(\lambda)$ denote the characteristic polynomial $\det(\lambda I - A)$. When A can be reduced to diagonal form by a similarity transformation, the proof that A satisfies its characteristic equation is particularly transparent, since each eigenspace is annihilated by a factor $A - \lambda_0$ in $G(A)$, where λ_0 is the characteristic value of the eigenspace. As it stands, this method is capable of giving the result only when a complete set of eigenspaces for A can be found, and general proofs usually depend either on canonical forms or on the use of adjoint matrices.

The proof of the Cayley-Hamilton theorem that we give here demonstrates how the simplest and most transparent method of diagonalization can be used to provide a proof in the general case of a square matrix A with elements that belong to an arbitrary commutative ring R. Apart from some elementary algebra, the only result that is required is the easy proof of the theorem for real matrices which have real and distinct characteristic roots.

Suppose that A is the $n \times n$ matrix $[a_{ij}]$, where $a_{ij} \in R$ for $1 \leqq i, j \leqq n$. Now let $X = [x_{ij}]$, $1 \leqq i, j \leqq n$, where each x_{ij} is an *indeterminate*, and let P denote the domain of polynomials which are generated over the integers by $\{x_{ij}\}$ and contain no 'constant' terms. A ring homomorphism $\phi\colon P \to R$ may be defined which 'evaluates' each polynomial at the point $x_{11} = a_{11}, x_{12} = a_{12}, \cdots, x_{nn} = a_{nn}$. Thus if $f(x_{11}, x_{12}, \cdots, x_{nn}) \in P$, this is mapped to $f(a_{11}, a_{12}, \cdots, a_{nn}) \in R$. The characteristic polynomials of X and A are defined over P and R as follows:

$$F(\lambda) = \det(\lambda I - X) = \lambda^n - p_1\lambda^{n-1} + \cdots + (-)^n p_n \quad \text{and}$$

$$G(\lambda) = \det(\lambda I - A) = \lambda^n - r_1\lambda^{n-1} + \cdots + (-1)^n r_n,$$

where each $p_j \in P$ and $r_j \in R$. If $f_{kl} \in P$ and $g_{kl} \in R$ denote the entries in the kl-position of the matrices $F(X)$ and $G(A)$, it follows from the relations $r_j = \phi(p_j)$ that $g_{kl} = \phi(f_{kl})$, and hence that $G(A)$ is zero if $F(X)$ is zero. Thus, to finiteness of this extended formula will depend not only upon the possibility of writing a finite expression for the characteristic function of the set A, but also upon the existence of a finitely expressible function $h(k)$ such that $h(k) \geqq$ the kth smallest element of A for all $k \geqq 1$.

Reference

1. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Oxford University Press, London, 1960, pp. 343–4.

THE GENERAL CAYLEY-HAMILTON THEOREM VIA THE EASIEST REAL CASE

J. DENMEAD SMITH, College of the Resurrection, Mirfield, Yorkshire, England.

Let A be a square matrix with elements in a field, and let $G(\lambda)$ denote the characteristic polynomial $\det(\lambda I - A)$. When A can be reduced to diagonal form by a similarity transformation, the proof that A satisfies its characteristic equation is particularly transparent, since each eigenspace is annihilated by a factor $A - \lambda_0$ in $G(A)$, where λ_0 is the characteristic value of the eigenspace. As it stands, this method is capable of giving the result only when a complete set of eigenspaces for A can be found, and general proofs usually depend either on canonical forms or on the use of adjoint matrices.

The proof of the Cayley-Hamilton theorem that we give here demonstrates how the simplest and most transparent method of diagonalization can be used to provide a proof in the general case of a square matrix A with elements that belong to an arbitrary commutative ring R. Apart from some elementary algebra, the only result that is required is the easy proof of the theorem for real matrices which have real and distinct characteristic roots.

Suppose that A is the $n \times n$ matrix $[a_{ij}]$, where $a_{ij} \in R$ for $1 \leqq i, j \leqq n$. Now let $X = [x_{ij}]$, $1 \leqq i, j \leqq n$, where each x_{ij} is an *indeterminate*, and let P denote the domain of polynomials which are generated over the integers by $\{x_{ij}\}$ and contain no 'constant' terms. A ring homomorphism $\phi: P \rightarrow R$ may be defined which 'evaluates' each polynomial at the point $x_{11} = a_{11}, x_{12} = a_{12}, \cdots, x_{nn} = a_{nn}$. Thus if $f(x_{11}, x_{12}, \cdots, x_{nn}) \in P$, this is mapped to $f(a_{11}, a_{12}, \cdots, a_{nn}) \in R$. The characteristic polynomials of X and A are defined over P and R as follows:

$$F(\lambda) = \det(\lambda I - X) = \lambda^n - p_1\lambda^{n-1} + \cdots + (-)^n p_n \quad \text{and}$$

$$G(\lambda) = \det(\lambda I - A) = \lambda^n - r_1\lambda^{n-1} + \cdots + (-1)^n r_n,$$

where each $p_j \in P$ and $r_j \in R$. If $f_{kl} \in P$ and $g_{kl} \in R$ denote the entries in the kl-position of the matrices $F(X)$ and $G(A)$, it follows from the relations $r_j = \phi(p_j)$ that $g_{kl} = \phi(f_{kl})$, and hence that $G(A)$ is zero if $F(X)$ is zero. Thus, to establish that A satisfies its characteristic equation over R it is sufficient to prove that X satisfies its characteristic equation over P, or that the n^2 polynomials $f_{kl}(x_{11}, x_{12}, \cdots, x_{nn})$ are all zero.

Now a polynomial which is defined over the integers is identically zero if and only if it is equal to zero when its arguments assume all real values. A proof of the Cayley-Hamilton theorem for every real $n \times n$ matrix, such as the one that uses reduction to complex triangular form, therefore implies that all components of the matrix $F(X)$ are identically zero, so that X and hence A satisfy their characteristic equations. However, a separate proof of the Cayley-Hamilton theorem even at this reduced level of generality can be avoided by the following device, which enables us to stay wholly within the real field and to consider only real matrices which admit a real diagonal form.

If $f(x_{11}, x_{12}, \cdots, x_{nn}) \in P$ and $S_{11}, S_{12}, \cdots, S_{nn}$ are sets of real numbers, each containing more elements than the maximum possible degree of f, then f is identically zero if it is equal to zero whenever each $x_{ij} \in S_{ij}$. Since the equation $F(X) = 0$ is equivalent to the n^2 identities $f_{kl}(x_{11}, x_{12}, \cdots, x_{nn}) = 0$, it follows that if the Cayley-Hamilton theorem is satisfied over the reals when $x_{11}, x_{12}, \cdots, x_{nn}$ take enough (in the above sense) special values, then the theorem is proved for X over the ring P. It will now be demonstrated how these special cases can be chosen to permit the easiest possible treatment.

First of all, let each x_{ij} take the value c_{ij}, where $[c_{ij}]$ is an $n \times n$ real diagonal matrix with distinct diagonal elements. From continuity and by considering the alternation of the sign of the function $F(\lambda)$ it can be seen that when each x_{ij} is real and sufficiently near to c_{ij}, i.e., $|x_{ij} - c_{ij}| < \varepsilon$ for some $\varepsilon > 0$, $F(\lambda)$ has n distinct real zeros, and hence X has a complete set of real eigenvectors. The elementary proof of the Cayley-Hamilton theorem by diagonalization is therefore applicable, and by identifying each set S_{ij} with a real ε-neighborhood of c_{ij} it can be seen that enough special cases have been included for the proofs of the theorem for X over the ring P, and hence for A over the ring R, to follow from the argument given above.

It may be remarked that the real field is not essential to this method, and a real algebraic extension of the rationals will do just as well.

A NOTE ON THE k-FREE INTEGERS

J. E. NYMANN, University of Texas at El Paso

In this paper S_k will denote the set of k-free integers (positive integers whose prime factors are all of multiplicity less than k) and $S_k(x)$ will denote the number of elements of S_k which are less than or equal to x. Gegenbauer [**2**, p. 47] has proved that $S_k(x) = x/\zeta(k) + O(x^{1/k})$. Recently [**3**] the author indicated (with details for the case $k = 2$) another proof of this result using a generalized Möbius inversion formula. (Theorem 3 in [**3**] should read $\cdots + O(x^{1/k})$ instead of $\cdots + O(x^{1-1/k})$.) The purpose of this note is to give still another proof using Möbius functions of order k which were first introduced by T. M. Apostol [**1**].

establish that A satisfies its characteristic equation over R it is sufficient to prove that X satisfies its characteristic equation over P, or that the n^2 polynomials $f_{kl}(x_{11}, x_{12}, \cdots, x_{nn})$ are all zero.

Now a polynomial which is defined over the integers is identically zero if and only if it is equal to zero when its arguments assume all real values. A proof of the Cayley-Hamilton theorem for every real $n \times n$ matrix, such as the one that uses reduction to complex triangular form, therefore implies that all components of the matrix $F(X)$ are identically zero, so that X and hence A satisfy their characteristic equations. However, a separate proof of the Cayley-Hamilton theorem even at this reduced level of generality can be avoided by the following device, which enables us to stay wholly within the real field and to consider only real matrices which admit a real diagonal form.

If $f(x_{11}, x_{12}, \cdots, x_{nn}) \in P$ and $S_{11}, S_{12}, \cdots, S_{nn}$ are sets of real numbers, each containing more elements than the maximum possible degree of f, then f is identically zero if it is equal to zero whenever each $x_{ij} \in S_{ij}$. Since the equation $F(X) = 0$ is equivalent to the n^2 identities $f_{kl}(x_{11}, x_{12}, \cdots, x_{nn}) = 0$, it follows that if the Cayley-Hamilton theorem is satisfied over the reals when $x_{11}, x_{12}, \cdots, x_{nn}$ take enough (in the above sense) special values, then the theorem is proved for X over the ring P. It will now be demonstrated how these special cases can be chosen to permit the easiest possible treatment.

First of all, let each x_{ij} take the value c_{ij}, where $[c_{ij}]$ is an $n \times n$ real diagonal matrix with distinct diagonal elements. From continuity and by considering the alternation of the sign of the function $F(\lambda)$ it can be seen that when each x_{ij} is real and sufficiently near to c_{ij}, i.e., $|x_{ij} - c_{ij}| < \varepsilon$ for some $\varepsilon > 0$, $F(\lambda)$ has n distinct real zeros, and hence X has a complete set of real eigenvectors. The elementary proof of the Cayley-Hamilton theorem by diagonalization is therefore applicable, and by identifying each set S_{ij} with a real ε-neighborhood of c_{ij} it can be seen that enough special cases have been included for the proofs of the theorem for X over the ring P, and hence for A over the ring R, to follow from the argument given above.

It may be remarked that the real field is not essential to this method, and a real algebraic extension of the rationals will do just as well.

A NOTE ON THE k-FREE INTEGERS

J. E. NYMANN, University of Texas at El Paso

In this paper S_k will denote the set of k-free integers (positive integers whose prime factors are all of multiplicity less than k) and $S_k(x)$ will denote the number of elements of S_k which are less than or equal to x. Gegenbauer [**2**, p. 47] has proved that $S_k(x) = x/\zeta(k) + O(x^{1/k})$. Recently [**3**] the author indicated (with details for the case $k = 2$) another proof of this result using a generalized Möbius inversion formula. (Theorem 3 in [**3**] should read $\cdots + O(x^{1/k})$ instead of $\cdots + O(x^{1-1/k})$.) The purpose of this note is to give still another proof using Möbius functions of order k which were first introduced by T. M. Apostol [**1**].

For completeness the definitions and the necessary elementary properties of these functions will be included here.

For k a positive integer the Möbius function of order k, denoted by μ_k, is defined as follows: $\mu_k(n)=0$ if n is divisible by the $(k+1)$st power of some prime and otherwise $\mu_k(n)=(-1)^r$, where r is the number of distinct primes whose kth powers divide n. Note that when $k=1$, μ_k is the ordinary Möbius function. The first two lemmas follow immediately from the definitions.

LEMMA 1. *$S_k(x)=\sum_{n\leqq x}|\mu_{k-1}(n)|$ for $k\geqq 2$.*

LEMMA 2. *μ_k is multiplicative for $k\geqq 1$.*

LEMMA 3. *$|\mu_k(n)|=\sum_{d^{k+1}|n}\mu(d)$ for $k\geqq 1$.*

Proof. It is easy to verify that both $|\mu_k(n)|$ and $\sum_{d^{k+1}|n}\mu(d)$ are multiplicative functions of n which agree when n is a prime power.

THEOREM. *$S_k(x)=x/\zeta(k)+O(x^{1/k})$ for $k\geqq 2$.*

Proof. Combining Lemmas 1 and 3 we have $S_k(x)=\sum_{n\leqq x}\sum_{d^k|n}\mu(d)$. Now changing the order of summation we have

$$S_k(x)=\sum_{d\leqq x^{1/k}}\sum_{n\leqq x/d^k}\mu(d)=\sum_{d\leqq x^{1/k}}\mu(d)\sum_{n\leqq x/d^k}1$$

$$=\sum_{d\leqq x^{1/k}}\mu(d)(x/d^k+O(1))=x\sum_{d\leqq x^{1/k}}\frac{\mu(d)}{d^k}+O(x^{1/k}).$$

From here on the proof is the same as the latter part of the proof alluded to in [2]. It is well known that $\sum_{d=1}^{\infty}[\mu(d)]/d^k=1/\zeta(k)$. Hence,

$$\sum_{d\leqq x^{1/k}}\frac{\mu(d)}{d^k}=1/\zeta(k)-\sum_{d=[x^{1/k}]+1}^{\infty}\frac{\mu(d)}{d^k}$$

and

$$\left|\sum_{d=[x^{1/k}]+1}^{\infty}\frac{\mu(d)}{d^k}\right|<\sum_{d=[x^{1/k}]}\frac{1}{d^k}<\int_{x^{1/k}}^{\infty}\frac{dt}{t^k}=(1/(k-1))(1/x^{1-1/k}).$$

Therefore,

$$x\sum_{d\leqq x^{1/k}}\frac{\mu(d)}{d^k}=x/\zeta(k)+O(x^{1/k})$$

and hence,

$$S_k(x)=x/\zeta(k)+O(x^{1/k}).$$

References

1. T. M. Apostol, Möbius functions of order k, Pacific J. Math., 32 (1970) 21–27.

2. L. Gegenbauer, Asymptotische Gesetze Zahlentheorie, Denkschriften der Akademie der Wissenschaften zu Wien, 49 (1885) 37–80.

3. J. E. Nymann, A note concerning the square-free integers, Amer. Math. Monthly, 79 (1972) 63–65.

NOTES AND COMMENTS

Concerning *An extension of Brocard geometry* by Paul Sidenblad in the May 1974 issue, O. Bottema writes that it is a nice paper but unfortunately all three theorems are known. They are mentioned for instance on p. 285 of R. A. Johnson, *Advanced Euclidean Geometry*, Dover, N.Y. 1960.

Regarding *Trigonometric identities* by Andy R. Magid in the September 1974 issue, Harry W. Hickey notes that it is not necessary to use commutative ring theory to prove the theorem. One needs only to observe that the polynomial equation $f(\sin x, \cos x) = 0$ can be written in the form

$$\sum_{n=0}^{N} (a_n \sin x + b_n) \cos^n x = 0$$

from which by replacing $\cos x$ by u and $\sin x$ by $\sqrt{1-u^2}$ one sees that $f(\sin x, \cos x) = 0$ is an identity iff $a_n = b_n = 0$ for all n, since otherwise $\sqrt{1-u^2}$ is the quotient of two polynomials in u.

W. R. Utz notes that the theorems of the paper *Continuous exactly k-to-one functions on R*, this MAGAZINE, 45 (1972), 224–225, are among those of his paper *Functions with uniform inverses*, Amer. Math. Monthly, 70 (1963) 405–407.

ANNOUNCEMENT OF LESTER R. FORD AWARDS

At its meeting on January 27, 1965, Denver, Colorado, the Board of Governors authorized a number of awards, to be named after Lester R. Ford, Sr., to authors of expository articles published in the MONTHLY and the MATHEMATICS MAGAZINE. A maximum of six awards will be made annually; each award is in the amount of $100. The articles are to be selected by a subcommittee of the Committee on Publications appointed for this purpose.

The 1975 recipients of these Awards, selected by a committee consisting of E. F. Beckenbach, Chairman, Emil Grosswald and I. J. Schoenberg, were announced by President Henry O. Pollak at the business meeting of the Association on August 19, 1975, at Western Michigan University. The recipients of the Ford Awards for articles published in 1974 were the following:

R. Ayoub, Euler and the Zeta Function, MONTHLY, 81 (1974) 1067–1086.

J. Callahan, Singularities and Plane Maps, MONTHLY, 81 (1974) 211–240.

D. E. Knuth, Computer Science and its Relation to Mathematics, MONTHLY, 81 (1974) 323–343.

J. C. C. Nitsche, Plateau's Problems and Their Modern Ramifications, MONTHLY, 81 (1974) 945–968.

S. K. Stein, Algebraic Tiling, MONTHLY, 81 (1974) 445–462.

L. Zalcman, Real Proofs of Complex Theorems, MONTHLY, 81 (1974) 115–137.

DAVID P. ROSELLE, *Secretary*

EDITED BY ADA PELUSO AND WILLIAM WOOTON

Materials intended for review should be sent to: Professor Ada Peluso, Department of Mathematics, Hunter College of CUNY, 695 Park Avenue, New York, New York 10021, or to Professor William Wooton, 1495 La Linda Drive, Lake San Marcos, California 92069. A boldface capital C *in the margin indicates that a review is based in part on classroom use.*

Africa Counts. By Claudia Zaslavsky. Prindle, Weber and Schmidt, Boston, Mass., 1973. x + 328 pp. $12.50.

Claudia Zaslavsky's book has made accessible a vast store of information about mathematics in Africa. Along with the accessibility comes enlightenment about the nature of mathematics in everday human endeavors. The emphasis is upon the kinds of mathematical activities that have been a part of African life for centuries, many of which are still apparent today.

A first approach to this review might have been to consider how the book could be used at various levels of education, but it soon became apparent that its utility would be a very personal matter. To put it another way, the book has something for everybody.

For those who are concerned with Africa's place in the development of quantitative and geometric aspects in society, Zaslavsky presents a critical review of the published history of African mathematics. She refers to reports about such events as: toolmaking hominids of prehistory (about 2 million years ago) in East Africa: a carved bone (about 10,000 years old) from Ishango, Zaire, which indicates a calendrical or numerical system; Egyptian mathematics which included not only numeration, calculation, and engineering, but also the use of mundane measuring vessels — standardized. It seems as though the literature gives more recognition to ancient achievements than it does to more recent use of mathematics. For example, she quotes from L. L. Conant's (1896) *The Number Concept*, wherein he interprets the use of very large numbers as "remarkable exceptions" to the "law (that) the growth of the number sense keeps pace with the growth of the intelligence in other respects." Zaslavsky effectively tears down these and other false interpretations about "primitive" minds that were passed around among early writings in anthropology. It is refreshing to read her selections from and opinions about a wide range of scholarly papers. Her bibliography is extensive and provides a good basis for going beyond this "preliminary survey of a vast field awaiting investigation."

For those interested in a lighter touch, there are descriptions of games which have direct relationship to logic, strategy, geometry, probability, or numeration. Some have familiar counterparts (in Western culture) such as three-in-a-row games (Tic-Tac-Toe type), counting (nursery rhymes), networks (bridges of Koenigsberg), magic squares, and riddles. The African setting provides special interest for classroom use, and points out the basic role of such mathematical recreations as intellectual pursuits. There is an African board game which is becoming familiar to us ("soro," "omweso," "wari," or "mancala" are names of some of its dozens of variations). It may well be the world's oldest game: I have seen "boards" which have been found carved in the ground at sites of prehistorical man. In its simplest versions it has fewer possible arrangements than checkers or chess, but as Zaslavsky points out, if it "is played with just thirty-six counters distributed in two rows of six holes, there are about 10^{24} total possibilities." Her discussion of games includes contemporary anecdotes (e. g., about the "soro" expertise of Tanzania's President Nyerere), historical and sociological perspective, pictorial illustration, as well as related mathematical background. This eclectic approach makes for pleasant and useful reading; so the book can be appreciated as an encyclopedia or as a collection of short stories.

The encyclopedic aspect is perhaps best described by a selected short list of subheadings for some of the topics in the first eight of twenty-five chapters: African Glory; How Do We Count?; Unusual Number Bases; Taboo on Counting; The Gold Trade; Cowrie Currency; Bride Wealth; Weights and Measures in Africa; Government Records. Later sections provide details by means of Regional Studies of Southwest Nigeria and East Africa.

Another section entitled "Patterns and Shape" would make a notable fifty-page pamphlet by itself. It includes a discussion of "Geometric Form in Architecture," "Geometric Form and Pattern in Art," and a treatise "Geometric Symmetries in African Art" by D. W. Crowe, University of Wisconsin. This part of the book justifies Zaslavsky's claim, "If one wanted to survey the whole field of geometric design in Africa, one would have to catalogue almost every aspect of life, from commerce to courtship."

The book does catalogue some vital aspects of ordinary life in Africa. It has provided me with information about African mathematics that goes far byond that which I learned while teaching there for two years. More important, by its humanistic emphasis it has helped me gain new insights into the way mathematics cuts across national, ethnic, and time boundaries. It is an enlightening book and merits a place on anyone's bookshelf.

LEONARD FELDMAN, San Jose State University

THE GREATER METROPOLITAN NEW YORK MATH FAIR

The Greater Metropolitan New York Math Fair will be held on March 28, 1976, at Pace University in Manhattan. It is a mathematical competition for students who have completed, or are currently taking, mathematics at the eleventh year level or higher in the public, private and parochial high schools in the New York City, Westchester, Putnam, Dutchess, or Rockland Counties. Exceptional students who do not meet these requirements but wish to submit papers for consideration by the Fair Committee are welcome to do so.

Further details and application forms may be obtained from: Dr. Theresa J. Barz, Secretary, Math Fair Committee, Department of Mathematics and Computer Science, St. John's University, Jamaica, NY 11439.

becoming familiar to us ("soro," "omweso," "wari," or "mancala" are names of some of its dozens of variations). It may well be the world's oldest game: I have seen "boards" which have been found carved in the ground at sites of prehistorical man. In its simplest versions it has fewer possible arrangements than checkers or chess, but as Zaslavsky points out, if it "is played with just thirty-six counters distributed in two rows of six holes, there are about 10^{24} total possibilities." Her discussion of games includes contemporary anecdotes (e. g., about the "soro" expertise of Tanzania's President Nyerere), historical and sociological perspective, pictorial illustration, as well as related mathematical background. This eclectic approach makes for pleasant and useful reading; so the book can be appreciated as an encyclopedia or as a collection of short stories.

The encyclopedic aspect is perhaps best described by a selected short list of subheadings for some of the topics in the first eight of twenty-five chapters: African Glory; How Do We Count?; Unusual Number Bases; Taboo on Counting; The Gold Trade; Cowrie Currency; Bride Wealth; Weights and Measures in Africa; Government Records. Later sections provide details by means of Regional Studies of Southwest Nigeria and East Africa.

Another section entitled "Patterns and Shape" would make a notable fifty-page pamphlet by itself. It includes a discussion of "Geometric Form in Architecture," "Geometric Form and Pattern in Art," and a treatise "Geometric Symmetries in African Art" by D. W. Crowe, University of Wisconsin. This part of the book justifies Zaslavsky's claim, "If one wanted to survey the whole field of geometric design in Africa, one would have to catalogue almost every aspect of life, from commerce to courtship."

The book does catalogue some vital aspects of ordinary life in Africa. It has provided me with information about African mathematics that goes far byond that which I learned while teaching there for two years. More important, by its humanistic emphasis it has helped me gain new insights into the way mathematics cuts across national, ethnic, and time boundaries. It is an enlightening book and merits a place on anyone's bookshelf.

LEONARD FELDMAN, San Jose State University

THE GREATER METROPOLITAN NEW YORK MATH FAIR

The Greater Metropolitan New York Math Fair will be held on March 28, 1976, at Pace University in Manhattan. It is a mathematical competition for students who have completed, or are currently taking, mathematics at the eleventh year level or higher in the public, private and parochial high schools in the New York City, Westchester, Putnam, Dutchess, or Rockland Counties. Exceptional students who do not meet these requirements but wish to submit papers for consideration by the Fair Committee are welcome to do so.

Further details and application forms may be obtained from: Dr. Theresa J. Barz, Secretary, Math Fair Committee, Department of Mathematics and Computer Science, St. John's University, Jamaica, NY 11439.

PROBLEMS AND SOLUTIONS

Readers of this department are invited to submit for solution problems believed to be new that may arise in study, in research, or in extra-academic situations. Problems may be submitted from any branch of mathematics and ranging in subject content from that accessible to the talented high school student to problems challenging to the professional mathematician. Proposals should be accompanied by solutions, when available, and by any information that will assist the editor. Ordinarily, problems in well-known textbooks should not be submitted.

The asterisk (*) will be placed by the problem number to indicate that the proposer did not supply a solution. Readers' solutions are solicited for all problems. Proposers' solutions may not be "best possible" and solutions by others will be given preference.

Solutions should be submitted on separate, signed sheets. Figures should be drawn in India ink, and exactly the size desired for reproduction.

Send all communications for this department to Dan Eustice, the Ohio State University, 231 W. 18th Ave., Columbus, Ohio 43210.

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

PROPOSALS

945. *Proposed by Alan Wayne, Pasco-Hernando Community College, Florida.*

Find the smallest Pythagorean triangle in which a square with integer sides can be inscribed so that an angle of the square coincides with the right angle of the triangle.

946. *Proposed by M. H. Hoehn, Santa Rosa, California.*

Two points are selected at random on the boundary of a unit square. What is the expected value of the length of the line segment joining the points?

947. *Proposed by Steve Moore and Mike Chamberlain, University of Santa Clara.*

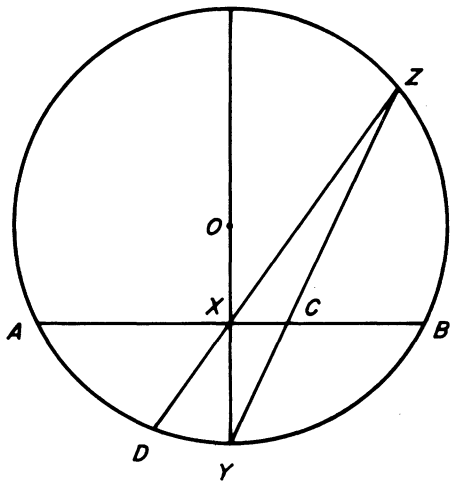
A line through the point (a, b) which is in the first quadrant forms a right triangle with the positive coordinate axes. Find the equation of the line which forms the triangle with minimum perimeter.

948.* *Proposed by Bob Prielipp and N. J. Kuenzi, Oshkosh, Wisconsin.*

Let Z_n be the ring of integers modulo n. For what values of n different from 2 do there exist permutations f and g on Z_n such that the pointwise product fg is also a permutation on Z_n?

949. *Proposed by P. Erdös, Hungarian Academy of Science, and M. S. Klamkin, University of Waterloo.*

In a circle with center O, OXY is perpendicular to chord AB (as shown).



Prove $DX \leqq CY$.

950. *Proposed by Erwin Just, Bronx Community College.*

Show that there is a unique real number c such that for every differentiable function f on $[0, 1]$ with $f(0)=0$ and $f(1)=1$, the equation $f'(x)=cx$ has a solution in $(0, 1)$.

951. *Proposed by G. A. Heuer, Concordia College.*

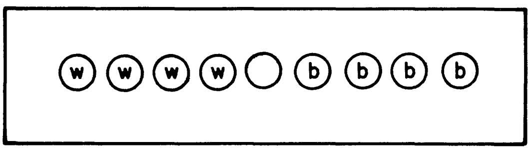
Let A be a square matrix, some scalar multiple of which differs from the identity matrix by a matrix of rank one. Give a simple necessary and sufficient condition that A be nonsingular, and find A^{-1} in this case.

952. *Proposed by F. D. Hammer, Stockton State College.*

The object of a familar puzzle is to interchange the positions of n white and n black pegs. One is allowed to jump pegs of opposite color, but never of the same color. A white (black) peg may move to the right (left) to an adjacent empty position.

Show that the transfer is always possible and establish a lower bound on the number of moves which is less than $2n(n+1)$.

$n=4$



953. *Proposed by Allan W. Johnson, Jr., Washington, D. C.*

An absolute prime is a prime number all of whose decimal digit permutations are also prime numbers. T. N. Bhargava and P. H. Doyle, *On the existence of* *absolute primes*, this MAGAZINE, (47) 233–234, noted that all absolute primes of two or more digits are composed from the digits 1, 3, 7, and 9. They also proved that no absolute prime exists which uses all four of these digits.

Problem. Show that no absolute prime number exists which contains three of the four digits 1, 3, 7, and 9.

Editor. Are there any absolute primes of more than three digits which contain two of the digits 1, 3, 7, and 9? [See Martin Gardner, *Mathematical Games*, Scientific American, June, 1964, page 118, where the result that decimal numbers composed entirely of 19 and 23 repetitions of the number 1 are prime is discussed.]

QUICKIES

This department will publish problems which may be solved by laborious methods, but which with the proper insight may be disposed of with dispatch. Readers are urged to submit their favorite problems of this type, together with the elegant solutions and the source if known.

Q625. If (a, b, c) is a Pythagorean triple, then Q605 shows that $M(a, b) < c/\sqrt{2}$, where M is the geometric mean. Prove that $A(a, b) < c/\sqrt{2}$, where A is the arithmetic mean.

[*Submitted by A. Wilansky*]

Q626. If a, b, and c are real and $b^2 < 2ac$, prove that the cubic $x^3 + ax^2 + bx + c$ has only one distinct real root.

[*Submitted by Philip Tracy.*]

Q627. It is a curious fact that $80/81 = .9876543210...$ is accurate to ten decimal places. Show that if $b \geqq 4$ is an integer, then, in the base b, $(\overline{b-2}\ 0)/(\overline{b-2}\ 1) = .\overline{b-1}\,\overline{b-2}\ldots 210\ldots$ is accurate to b b-places with error less than $(b^{-b})/2$.

[*Submitted by Michael Golomb.*]

(Answers on page 248)

SOLUTIONS

Interchanged Digits

908. [September, 1974] *Proposed by J. A. H. Hunter, Toronto, Canada.*

We define N as an integer with $(2n + 1)$ digits, the first digit not a zero. Then say N is represented as $(A)(B)$, (A) having n digits, (B) having $(n + 1)$ digits.

When (A) and (B) are interchanged the result is equal to $8N$. What is the smallest value of N that meets this requirement?

Solution by Václav Konečný, Gottwaldov, Czechoslovakia.

$N = 10^{n+1}A + B$ and $8N = 10^nB + A$. Thus $B/A = (8\cdot 10^{n+1} - 1)/(10^n - 8)$. Put $M = 8\cdot 10^{n+1} - 1$ and $D = 10^n - 8$. Then $A = kD$ and $B = kM$. We shall find k which meets the requirements of the problem. Notice first, that $10^{n-1}/(10^n - 8) \leqq k < 10^{n+1}/(8\cdot 10^{n+1} - 1)$, $(n > 1)$. $M = D80 + 639$; $D = 639q + r$, where q is an integer and $0 \leqq r < 639$. As $639 = 3^2 \times 71$ and D is not divisible by 3, the only possible common divisor $(\neq 1)$ of M and D is 71. Performing the division of $10^n - 8$ (or M) by 71 we get the remainder zero if $n = 17 + 35m$, where $m = 0, 1, 2, \cdots$. Thus, $k = 8/71$ and

$$N_n = (8/71)(10^n - 8)10^{n+1} + (8/71)(8\cdot 10^{n+1} - 1).$$

To get the smallest value of N put $n = 17$ $(m = 0)$. Explicitly:

$$N_{17} = 1126760563380281690140845070422535 2.$$

Notice the structure of $N_n = N_n(m)$:

$N_n(0) = (A)(B)$,

$N_n(1) = (\ (A)(B)(A)\ )(\ (B)(A)(B)\ )$,

$N_n(2) = (\ (A)(B)(A)(B)(A)\ )(\ (B)(A)(B)(A)(B)\ )$, etc.

Also solved by Walter Bluger (Canada), Alfred Brousseau, D. P. Choudhury (India), Stephen C. Currier, Clayton W. Dodge, Karl Heuer, Chandra M. R. Kintala, Harry D. Ruderman, Kenneth M. Wilke, and the proposer.

Relatively Prime

909. [September, 1974] *Proposed by Dennis R. Lichtman and James L. Murphy, California State College, San Bernardino.*

Let ϕ denote Euler's phi-function. Find the positive integers n such that n and $\phi(n)$ are relative primes.

Solution by Ken Rebman, California State University, Hayward.

Let $n = \Pi p_i^{a_i}$ be the factorization of n into powers of distinct primes. Then $\phi(n) = \Pi p_i^{a_i - 1}(p_i - 1)$, and n and $\phi(n)$ will be relatively prime if and only if none of the primes p_i divide $\phi(n)$. This will happen if and only if $a_i - 1 = 0$ for each i and p_j does not divide $p_i - 1$ for each pair (i, j). Put another way, n and $\phi(n)$ are relatively prime if and only if: (i) $n = p_1p_2\cdots p_k$ where the $p_i (i = 1, 2, \cdots, k)$ are distinct primes (i.e., n is square-free) and

(ii) for every pair i and j, $(i = 1, 2, \cdots, k$ and $j = 1, 2, \cdots, k)$, $p_i \not\equiv 1 \pmod{p_j}$.

Note that the only even n satisfying these two conditions is $n = 2$.

Editor's comment. Arthur Marshall found the result in D. H. Lehmer's paper, *On Euler's totient function*, Bull. Amer. Math. Soc., 38 (1932), 345–351.

Also solved by Joe Altree, Merrill Barnebey, Alfred Brousseau, D. P. Choudhury, Stephen C. Currier, Hugo D'Alarcao, Clayton W. Dodge, Thomas E. Elsner, David Farnsworth, Michael Goldberg, M. G. Greening (Australia), Lee O. Hagglund, Vaclav Konecny (Czechoslovakia), Lew Kowarski, Arthur Marshall, George A. Novacky, Jr., Leonard L. Palmer, Bob Prielipp, James V. Rauff, Daniel Mark Rosenblum, John Samoylo, Erwin Schmid, David Singmaster (England), Kenneth M. Wilke, and the proposers.

Inequalities for a Triangle

910. [September, 1974]. *Proposed by L. Carlitz, Duke University.*

Let P be a point in the interior of the triangle ABC and let r_1, r_2, r_3 denote the distances from P to the sides of ABC. Let a, b, c denote the sides and r the radius of the incircle of ABC. Show that

$$\frac{a}{r_1}+\frac{b}{r_2}+\frac{c}{r_3} \geqq \frac{2s}{r}, \tag{1}$$

$$ar_1^2+br_2^2+cr_3^2 \geqq 2r^2s, \tag{2}$$

$$(s-a)r_2r_3+(s-b)r_3r_1+(s-c)r_1r_2 \leqq r^2s, \tag{3}$$

$$ar_1^2+br_2^2+cr_3^2+(s-a)r_2r_3+(s-b)r_3r_1+(s-c)r_1r_2 \geqq 3r^2s, \tag{4}$$

where $2s = a+b+c$. In each case there is equality if and only if P is the incenter of ABC.

Solution by M. S. Klamkin, University of Waterloo.

Since $ar_1+br_2+cr_3 = 2rs = 2\Delta$ ($\Delta =$ area of ABC), it follows from Cauchy's inequality that

$$\left(\frac{x}{r_1}+\frac{y}{r_2}+\frac{z}{r_3}\right)(ar_1+br_2+cr_3) \geqq (\sqrt{ax}+\sqrt{by}+\sqrt{cz})^2, \tag{5}$$

$$(xr_1^2+yr_2^2+zr_3^2)\left(\frac{a^2}{x}+\frac{b^2}{y}+\frac{c^2}{z}\right) \geqq (ar_1+br_2+cr_3)^2, \tag{6}$$

for all $x, y, z \geqq 0$. Thus,

$$\frac{x}{r_1}+\frac{y}{r_2}+\frac{z}{r_3} \geqq (\sqrt{ax}+\sqrt{by}+\sqrt{cz})^2/2\Delta, \tag{7}$$

$$xr_1^2+yr_2^2+zr_3^2 \geqq 4\Delta^2 \Big/ \left\{\frac{a^2}{x}+\frac{b^2}{y}+\frac{c^2}{z}\right\}, \tag{8}$$

with equality in (7) and (8), respectively, *iff*

$$\frac{ar_1^2}{x}=\frac{br_2^2}{y}=\frac{cr_3^2}{z},$$

$$\frac{xr_1}{a}=\frac{yr_2}{b}=\frac{zr_3}{c}.$$

It is to be noted that (6) is valid for all real r_1, r_2, r_3. For the special case $(x, y, z) = (a, b, c)$, (7) and (8) reduce to (1) and (2). Incidentally, (2) will also follow immediately from (3) and (4).

We now show that (3) and (4) are special cases corresponding to $n = 1$ of the known master triangle inequality

$$(9) \qquad u^2 + v^2 + w^2 \geqq (-1)^{n+1}\{2vw \cos nA + 2wu \cos nB + 2uv \cos nC\}$$

where u, v, w are arbitary real numbers; A, B, C are angles of an arbitrary triangle. There is equality *iff* $u/\sin nA = v/\sin nB = w/\sin nC$ (M. S. Klamkin, *Asymmetric triangle inequalities*, Univ. Beograd Publ. Elektrotehn. Fak. Ser. Mat. Fiz., No. 357 — No. 380 (1971), 33–44). Letting $u = ax$, $v = by$, $w = cz$, (9) for $n = 1$, is also equivalent to

$$(10) \qquad a^2x^2 + b^2y^2 + c^2z^2 \geqq (b^2 + c^2 - a^2)yz + (c^2 + a^2 - b^2)zx + (a^2 + b^2 - c^2)xy.$$

The latter inequality can be traced back at least to Wolstenholme (ibid.).

By multiplying (3) and (4) by $4s$ and using $2rs = \Sigma\, ar_1$, they can be rewritten, respectively, as

$$(11) \qquad \sum_{\text{cyclic}} \{a^2r_1^2 - (b^2 + c^2 - a^2)r_2r_3\} \geqq 0,$$

$$(12) \qquad \sum_{\text{cyclic}} \{a(2b + 2c - a)r_1^2 + [(b + c)^2 - a^2 - 6bc]r_2r_3\} \geqq 0.$$

Now noting that if $a_1^2 = a(2b + 2c - a)$, $b_1^2 = b(2c + 2a - b)$, and $c_1^2 = c(2a + 2b - c)$, then a_1, b_1, c_1 are sides of a triangle, it follows that (11) and (12) are valid for all real r_1, r_2, r_3 with equality *iff* $r_1 = r_2 = r_3$.

One can also obtain (8) as a special case of (10).

A further generalization of (2) is given by

$$(xr_1^m + yr_2^m + zr_3^m)^{1/m}\left\{\left(\frac{a^m}{x}\right)^{1/(m-1)} + \left(\frac{b^m}{y}\right)^{1/(m-1)} + \left(\frac{c^m}{z}\right)^{1/(m-1)}\right\}^{(m-1)/m} \geqq 2\Delta,$$

where $x, y, z, m - 1 > 0$; this follows from Hölder's inequality.

Also solved by Alfred Brousseau, Alex G. Ferrer (Mexico), M. G. Greening (Australia), S. L. Haven & J. M. Stark, M. K. King, Vaclav Konecny (Czechoslovakia), U. V. M. Rao (India), and the proposer.

Vertices of a Die

911. [September, 1974] *Proposed by Charles W. Trigg, San Diego, California.*

On a standard cubical die, the three faces around one vertex are numbered 1, 2, 3 in order counterclockwise. The digits 4, 5, 6 are distributed on the other three faces so that the sum of the digits on each pair of opposite faces is 7.

When the numbers on the three faces around each vertex are added, the minimum sum is 6 and the maximum sum is 15. This is a span of ten integers, but there are only eight vertices. Without actually adding up the digits around the other six vertices, determine what the two numbers missing among the sums are.

Solution by Julius Vogel, Boston, Massachusetts.

We establish first that all eight vertices have different sums. Denote by (a, b, c) the vertex associated with faces a, b and c. Any of the three vertices adjacent to (a, b, c) is of the form $(a, b, 7-c)$. Therefore, no two adjacent vertices can have the same sum, because this would imply $c = 3/2$. Any of the three vertices at the other end of a face diagonal from (a, b, c) is of the form $(a, 7-b, 7-c)$. But if $a+b+c=a+7-b+7-c$, then $b+c=7$. This is impossible since b and c are adjacent, rather than opposite, sides. The vertex at the other end of the main diagonal from (a, b, c) is $(7-a, 7-b, 7-c)$. But if $a+b+c=7-a+7-b+7-c$, then $a+b+c=21/2$. Therefore, all eight vertices have different sums, so that only two integers between 6 and 15 are not represented.

Now, we note that for a vertex to have the sum 8 it must be of the form $(1, 2, 5)$ or $(1, 3, 4)$, since these are the only partitions of 8 into three distinct integers between 1 and 6. But neither of these is a possibility for a vertex because faces 2 and 5 are on opposite sides of the die, as are faces 3 and 4. Therefore, no vertex has sum 8. Sum 13 is impossible as well, since the total of the sums of the vertices at the opposite ends of a main diagonal is 21.

Also solved by Walter Bluger, D. P. Choudhury (India), Clayton W. Dodge, Thomas E. Elsner, Richard A. Gibbs, Michael Goldberg, E. Sherman Grable, Vaclav Konecny (Czechoslovakia), Sidney Kravitz, Tung-Po Lin, Graham Lord, Joseph V. Michalowicz, Thomas O'Loughlin, F. D. Parker, Ken Rebman, M. Rodeen, Gillian W. Valk, Edward T. H. Wang, and the proposer.

Binomial Identities

912. [September, 1974] *Proposed by Michael O'Rourke, University of Wisconsin — Parkside.*

Defining three functions of the natural numbers

$$f_1(n) = \sum_{i=0}^{[n/2]} (-1)^i \, 2^{n-2i} \binom{n-i}{i},$$

$$f_2(n) = \sum_{i=0}^{(n-1)/2} (-1)^i \, 2^{n-(2i+1)} \binom{n-i}{i} \binom{n-2i}{1},$$

$$f_3(n) = \sum_{i=0}^{(n-2)/2} (-1)^i \, 2^{n-(2i+2)} \binom{n-i}{i} \binom{n-2i}{2},$$

show that

$$f_1(n) = n + 1,$$

$$f_2(n) = \binom{n+2}{3},$$

$$f_3(n) = \binom{n+3}{5}.$$

Solution by L. Carlitz, Duke University.

For k an arbitrary positive integer, put

$$f_k(n) = \sum_{2i \leq n-k+1} (-1)^i \, 2^{n-2i-k+1} \binom{n-i}{i} \binom{n-2i}{k-1}$$
$$= \sum_{2i \leq n-k+1} (-1)^i \, 2^{n-2i-k+1} \binom{n-i}{i+k-1} \binom{i+k-1}{i}.$$

Then

$$\sum_{n=k-1}^{\infty} f_k(n) x^n = \sum_{n=k-1}^{\infty} x^n \sum_{2i \leq n-k+1} (-1)^i \, 2^{n-2i-k+1} \binom{n-i}{i+k-1} \binom{i+k-1}{i}$$
$$= x^{k-1} \sum_{i=0}^{\infty} (-1)^i \binom{i+k-1}{i} x^{2i} \sum_{m=0}^{\infty} \binom{m+i+k-1}{m} (2x)^m$$
$$= x^{k-1} \sum_{i=0}^{\infty} (-1)^i \binom{i+k-1}{i} x^{2i} (1-2x)^{-i-k}$$
$$= \frac{x^{k-1}}{(1-2x)^k} \sum_{i=0}^{\infty} (-1)^i \binom{i+k-1}{i} \frac{x^{2i}}{(1-2x)^i}$$
$$= \frac{x^{k-1}}{(1-2x)^k} \left(1 + \frac{x^2}{1-2x}\right)^{-k} = x^{k-1}(1-2x+x^2)^{-k}$$
$$= x^{k-1}(1-x)^{-2k} = x^{k-1} \sum_{m=0}^{\infty} \binom{m+2k-1}{2k-1} x^m$$
$$= \sum_{n=k-1}^{\infty} \binom{n+k}{2k-1} x^n.$$

Therefore

$$f_k(n) = \binom{n+k}{2k-1}.$$

In particular, for $k = 1, 2, 3$, this gives the stated results.

Editor's comment. Graham Lord references J. Riordan, *Combinatorial Identities*, Wiley, 1968, Section 2.4. Henry W. Gould references his book, *Combinatorial Identities*, rev. 2nd ed., published by the author, 1972.

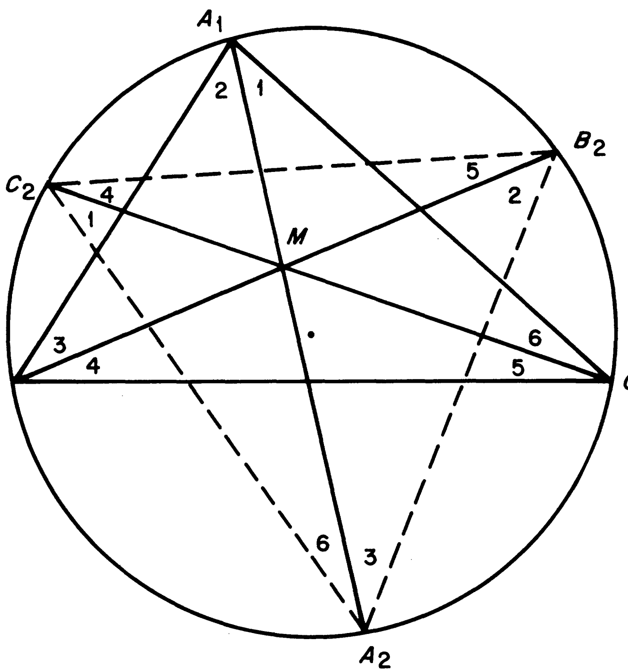
Also solved by M. T. Bird, J. C. Binz (Switzerland), Henry W. Gould, M. G. Greening (Australia), Graham Lord, John Oman, U. V. M. Rao (India), and the proposer.

Convergent Angles

913. [September, 1974] *Proposed by J. Garfunkel, Forest Hills High School, Flushing, New York.*

Triangle $A_1B_1C_1$ is inscribed in a circle. The medians are drawn and extended to the circle meeting the circle at points $A_2B_2C_2$. The medians of triangle $A_2B_2C_2$ are likewise drawn and extended to the circle to points $A_3B_3C_3$ and so on. Prove that triangle $A_nB_nC_n$ becomes equilateral as $n \to \infty$ (and very rapidly).

Amalgam of solutions by Brother Alfred Brousseau, St. Mary's College; and Edward Itors, State University.



We assume that $B_1C_1 > A_1C_1 > A_1B_1$, as in the diagram.

Since M is on the same side of the perpendicular bisector of B_1C_1 as A_1, but A_2 is on the opposite side (thus dividing arc $B_1A_2C_1$ unequally), we have
$B_1M < C_1M$, so that $\angle 5 < \angle 4$, as well as $\angle 1 < \angle 2$. Similarly,
$A_1M < B_1M$, so that $\angle 3 < \angle 2$, as well as $\angle 5 < \angle 6$; and
$A_1M < C_1M$, so that $\angle 6 < \angle 1$, as well as $\angle 4 < \angle 3$.

By considering subtended arcs, we find that

$$A_2 = \angle 6 + \angle 3 < \angle 1 + \angle 2 = A_1;$$

$$B_2 = \angle 5 + \angle 2 < \angle 1 + \angle 2 = A_1, \text{ since } \angle 5 < \angle 6 < \angle 1; \text{ and}$$

$$C_2 = \angle 1 + \angle 4 < \angle 1 + \angle 2 = A_1, \text{ since } \angle 4 < \angle 3 < \angle 2.$$

Thus, every angle of triangle $A_2B_2C_2$ is less than the largest angle of triangle $A_1B_1C_1$. The same conclusion follows if $A_1B_1C_1$ is isosceles but not equilateral. If $A_1B_1C_1$ is equilateral, then so is $A_2B_2C_2$. A similar argument shows that every angle of triangle $A_2B_2C_2$ is greater than the smallest angle of triangle $A_1B_1C_1$, except in the equilateral case.

Let T be the transformation described by the problem, with notation slightly altered so that $T(A, B, C) = (A', B', C')$, where the angle-triples (A, B, C) and (A', B', C') are in nonincreasing order. If (A_1, B_1, C_1) is given, and $(A_n, B_n, C_n) = T^{n-1}(A_1, B_1, C_1)$, then the preceding argument shows that $A_n \geqq A_{n+1} \geqq \pi/3$ and $C_n \leqq C_{n+1} \leqq \pi/3$. Hence the sequences $(A_n)_{n=1}^{\infty}$ and $(C_n)_{n=1}^{\infty}$ have limits, say A and C, and $(B_n)_{n=1}^{\infty}$ has limit $B = \pi - A - C$. From the obvious continuity of T it follows that

$$(A, B, C) = \lim_n T^n(A_1, B_1, C_1) = T(\lim_n T^{n-1}(A_1, B_1, C_1)) = T(A, B, C),$$

which, by the first part of the argument, is possible only if $A = B = C = \pi/3$.

We were unable to quantify the proposer's parenthetical remark concerning the rapidity of convergence.

Also solved by J. M. Stark and the proposer.

Balancing Weights

914. [September, 1974]. *Proposed by Murray S. Klamkin, Ford Motor Company.*

If for any n of a given $n + 1$ integral weights, there exists a balance of them on a two pan balance where a fixed number of weights are placed on one pan and the remainder on the other pan, prove that the weights are all equal.

Solution by Thomas E. Elsner, General Motors Institute.

Let $w_1, w_2, \cdots, w_{n+1}$ be the $n + 1$ integral weights. Since any n of the weights balance, the sum of any n of the weights must be even. This implies further that all the weights have the same parity (congruent (mod 2)). Now the balancing properties of the initial weights must be shared by the integers $w_i/2$ or $(w_i - 1)/2$ (depending on whether the w_i are all even or odd).

Hence, the w_i must be congruent (mod 4). Continuing in the same way, the w_i are congruent (mod 4). Continuing in the same way, the w_i are congruent (mod 2^k) for every k and this implies that the weights are equal and further, that n is even.

Editor's comment. Several solvers noted that this problem is a generalization of problem B-1 on the 1973 William Lowell Putnam Exam. The proposer referenced the USSR Olympiad Problem Book, W. H. Freeman and Co., 1962, page 8.

Also solved by Walter Bluger, Clayton W. Dodge, M. G. Greening (Australia), Graham Lord, John Oman, David L. Wright, and the proposer.

ANSWERS

A625. $4[(c/\sqrt{2})^2-((a+b)/2)^2]=(a-b)^2>0.$

A626. A cubic with real coefficients has either one or three real roots. Suppose that the cubic has three real roots $-p_1$, $-p_2$, and $-p_3$. Then, $a=p_1+p_2+p_3$, $b=p_1p_2+p_2p_3+p_3p_1$, and $c=p_1p_2p_3$. Thus $b^2-2ac=(p_1p_2)^2+(p_2p_3)^2+(p_3p_1)^2\geqq 0$, contradicting hypothesis.

A627. Let A and B denote the left and right hand terms respectively. Then

$$A=\frac{b^2-2b}{(b-1)^2},\quad B=\sum_{n=1}^{b-1}(b-n)b^{-n}=\sum_{n=0}^{b-2}b^{-n}-\sum_{n=1}^{b-1}nb^{-n}=\frac{b^2-2b+b^{-b+1}}{(b-1)^2}.$$

Thus,

$$0\leqq B-A=\frac{b^{-b+1}}{(b-1)^2}<\frac{b^{-b}}{b-2}\leqq\frac{1}{2}b^{-b}.$$

(Quickies on page 240)

The World-Renowned

NEW MATHEMATICAL LIBRARY

is now published by

THE MATHEMATICAL ASSOCIATION OF AMERICA

This continuing series of inexpensive, paperbound books by mathematical scholars is designed for the high school or college student who wants a new challenge in understanding and appreciating important mathematical concepts. Internationally acclaimed as the most distinguished series of its kind, the NEW MATHEMATICAL LIBRARY features eminent expositors Ross Honsberger, Ivan Niven, H.S.M. Coxeter, and others of equal stature. Another distinctive feature of NML is its collection of problem books from high school mathematical competitions in the U.S. and abroad. Teachers at all levels will find many stimulating ideas in these books suitable for the classroom or independent study. The books are ideal for supplementary reading, school libraries, advanced or honor students, mathematics clubs, or the teacher's professional library. *Watch for new titles that will appear regularly.*

NUMBERS: RATIONAL AND IRRATIONAL (NML-01)
by Ivan Niven.
A clear interesting exposition of number systems, beginning with the natural numbers and extending to the rational and real numbers.

WHAT IS CALCULUS ABOUT? (NML-02)
by W. W. Sawyer
Develops the basic concepts of calculus intuitively, exploiting such familiar concepts as speed, acceleration, and volume.

AN INTRODUCTION TO INEQUALITIES (NML-03)
by E. F. Beckenbach and R. Bellman
Includes an axiomatic treatment of inequalities, proofs of the classical inequalities and powerful applications to optimum problems.

GEOMETRIC INEQUALITIES (NML-04)
by N. D. Kazarinoff
An informal presentation of the arithmetic mean-geometric mean inequality and discussions of the famous isoperimetric theorems, the reflection principle, and Steiner's symmetrization—with problem solving emphasis.

THE CONTEST PROBLEM BOOK (NML-05)
Problems from the Annual High School Mathematics Contests sponsored by the MAA and four other organizations. Covers the period 1950-1960. Compiled and with solutions by Charles T. Salkind.
A complete collection of examination questions and solutions from the first decade of this national high school mathematics competition. No mathematics beyond intermediate algebra required. Elementary procedures are consistently used in solutions, but more sophisticated alternatives are included where appropriate.

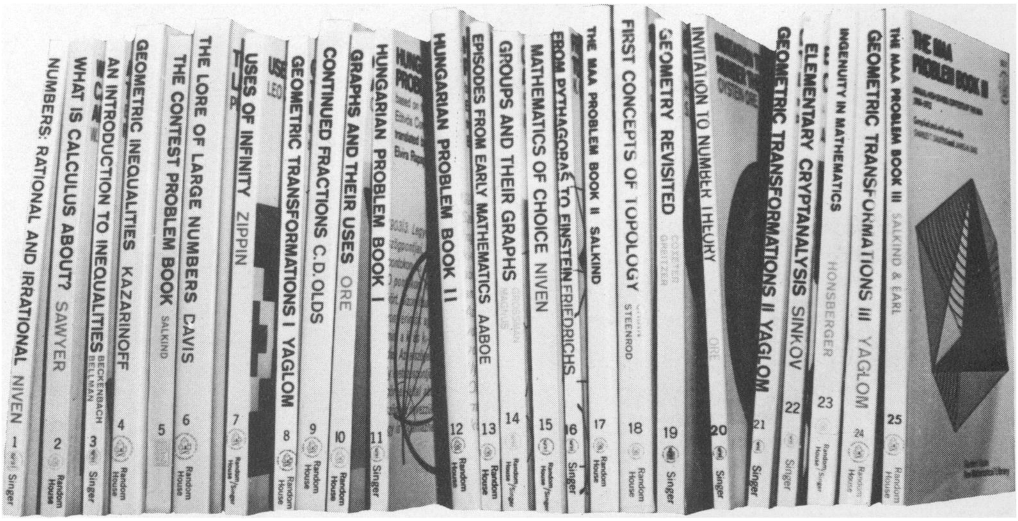
THE LORE OF LARGE NUMBERS (NML-06)
by P. J. Davis
How to work with numbers, big and small, and understand some of their less obvious properties.

USES OF INFINITY (NML-07)
by Leo Zippin
How mathematics have transformed the almost mystic concept of infinity into a precise tool essential in all branches of mathematics.

GEOMETRIC TRANSFORMATIONS (NML-08)
by I. M. Yaglom translated by Allen Shields
Concerned with isometries (distance-preserving transformations), a completely different way of looking at familiar geometrical facts that supplies the students with methods by which powerful problem solving techniques may be developed.

CONTINUED FRACTIONS (NML-09)
by Carl D. Olds
Shows how rational fractions can be expanded into continued fractions and introduces such diverse topics as the solutions of linear Diophantine equations, expansion of irrational numbers into infinite continued fractions and rational approximation to irrational numbers.

GRAPHS AND THEIR USES (NML-10)
by Oystein Ore
Develops enough graph theory to approach such problems as scheduling the games of a baseball league or a chess tournament, solving some ancient puzzles or analyzing winning (or losing) positions in certain games.



HUNGARIAN PROBLEM BOOK I (NML-11)
HUNGARIAN PROBLEM BOOK II (NML-12)
Based on the Eötvös Competitions 1894-1928
translated by E. Rapaport
The challenging problems that have become famous for the simplicity of the concepts employed, the mathematical depth reached, and the diversity of elementary mathematical fields touched.

EPISODES FROM THE EARLY HISTORY OF MATHEMATICS (NML-13)
by A. Aaboe
The contributions and amazing progress made by ancient mathematicians are revealed as the author describes Babylonian arithmetic and topics reconstructed from *Eculid's Elements;* from the writings of Archimedes and from Ptolemy's *Almagest.*

GROUPS AND THEIR GRAPHS (NML-14)
by I. Grossman and W. Magnus
In this introduction to group theory, abstract groups are made concrete in visual patterns that correspond to group structure. Suitable for students at a relatively early stage of mathematical growth.

THE MATHEMATICS OF CHOICE (NML-15)
by Ivan Niven
Stresses combinatorial mathematics and a variety of ingenious methods for solving questions about counting. Offers preparation for the study of probability.

FROM PYTHAGORAS TO EINSTEIN (NML-16)
by K. O. Friedrichs
The Pythagorean theorem and the basic facts of vector geometry are discussed in a variety of mathematical and physical contexts leading to the famous $E = mc^2$.

THE MAA PROBLEM BOOK II (NML-17)
A continuation of NML 5 containing problems and solutions from the Annual High School Mathematics Contests for the period 1961-1965.

FIRST CONCEPTS OF TOPOLOGY (NML-18)
by W. G. Chinn and N. E. Steenrod
The development of topology and some of its simple applications. The power and adaptability of topology are demonstrated in proving so-called existence theorems.

GEOMETRY REVISITED (NML-19)
by H.S.M. Coxeter and S. L. Greitzer
The purpose of this book is to revisit elementary geometry, using modern techniques such as transformations, inversive geometry, and projective geometry to facilitate geometric understanding and link the subject with other branches of mathematics.

INVITATION TO NUMBER THEORY (NML-20)
by Oystein Ore
A highly readable introduction to number theory that explains why the properties of numbers have always held such fascination for man. Problems, with solutions, allow students to find number relationships of their own.

GEOMETRIC TRANSFORMATIONS II (NML-21)
by I. M. Yaglom translated by Allen Shields
Similarity (shape-preserving) transformations are dealt with in the same manner in which distance-preserving transformations were treated in the first volume in *Geometric Transformations* (NML-8). Includes numerous problems with detailed solutions.

ELEMENTARY CRYPTANALYSIS — A Mathematical Approach (NML-22)
by Abraham Sinkov
In this systematic introduction to the mathematical aspects of cryptography, the author discusses monoalphabetic and polyalphabetic substitutions, digraphic ciphers and transpositions. The necessary mathematical tools include modular arithmetic, linear algebra of two dimensions with matrices, combinatorics and statistics. Each topic is developed as needed to solve decoding problems.

INGENUITY IN MATHEMATICS (NML-23)
by Ross Honsberger
Nineteen independent essays that reveal elegant and ingenious approaches used in thinking about such topics in elementary mathematics as number theory, geometry, combinatorics, logic and probability.

GEOMETRIC TRANSFORMATIONS III (NML-24)
by I. M. Yaglom translated by Abe Shenitzer
This part of Yaglom's work (sequel to NML-08 and NML-21) treats affine and projective transformations, and introduces the reader to non-Euclidean geometry in a supplement to hyperbolic geometry. As in the previously published parts of his work, Yaglom focuses on problems and their detailed solutions and keeps the text brief and simple.

THE MAA PROBLEM BOOK III (NML-25)
A continuation of NML-05 and NML-17, containing problems and solutions from the Annual High School Mathematics Contests for the period 1966-1972.

PLEASE DO NOT TEAR—SEND ENTIRE PAGE—THANK YOU

ORDER FORM

MAIL TO: THE MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, D. C. 20036

Please send the following NEW MATHEMATICAL LIBRARY books:

___ copies of ______________ @ $ ________ $ ________

___ copies of ______________ @ $ ________ $ ________

___ copies of ______________ @ $ ________ $ ________

___ copies of ______________ @ $ ________ $ ________

TOTAL $ ________

□ I enclose $ _____ and understand that the books will be sent postage and handling free. ($4.00 per copy. Individual members of the MAA may purchase *one* copy of each title for $3.00. Sign below for member's rate.)

□ Please bill me (For orders totalling $5.00 or more. Postage and handling fee will be added.)

I am an individual member of the MAA in good standing. The books ordered above are for my personal use.

____
Signature

Mail the books to: Name ______________________

Address ______________________

______________________ Zip Code ________