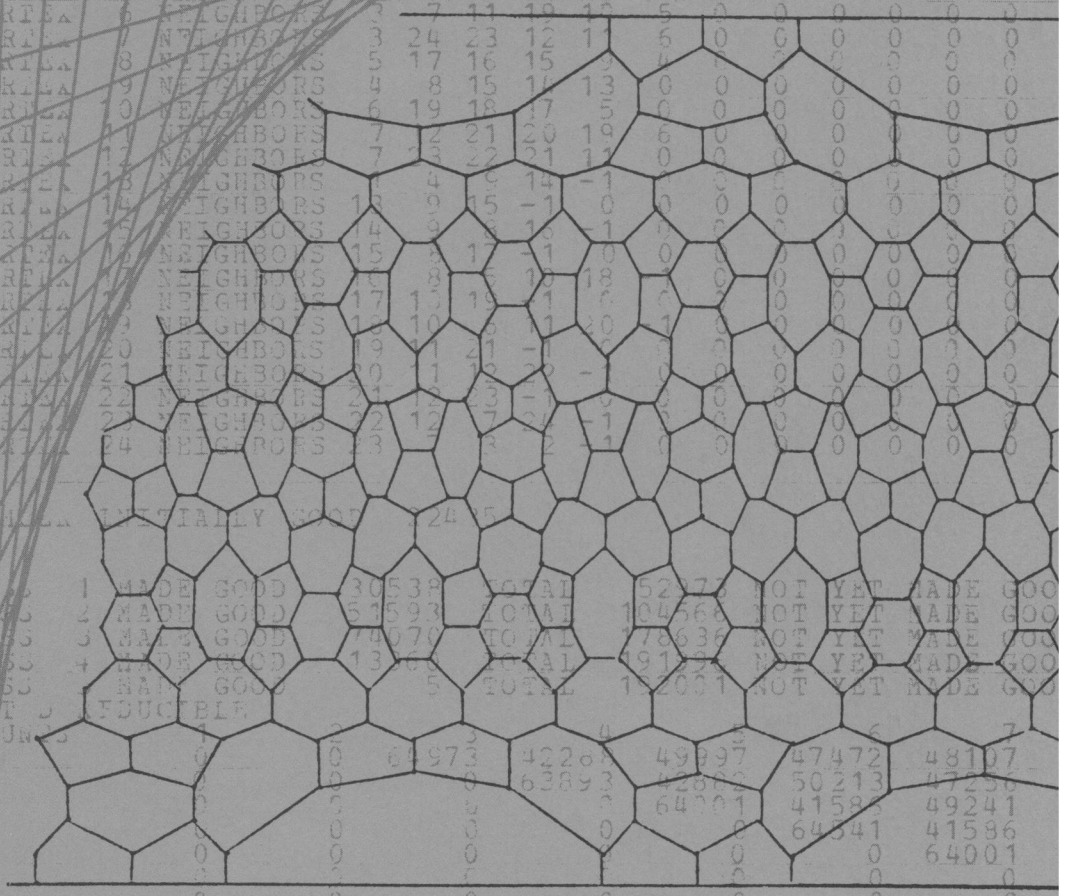
MATHEMATICS MAGAZINE



SOLUTION OF THE FOUR COLOR PROBLEM

SIMPLE GROUPS • TRIHEXAFLEXAGONS

SINGULAR MATRICES • CONWAY STONES

Vol. 49, No. 4
September, 1976
CODEN: MAMGA8

Eminent Mathematicians and Mathematical Expositors speak to STUDENTS and TEACHERS in. . .

The NEW MATHEMATICAL LIBRARY

An internationally acclaimed paperback series providing •

- *stimulating excursions for students beyond traditional school mathematics*
- *supplementary reading for school and college classrooms*
- *valuable background reading for teachers*
- *challenging problems for solvers of all ages from high school competitions in the US and abroad.*

Founded and raised to maturity by the SCHOOL MATHEMATICS STUDY GROUP. Adopted in 1975 by the MATHEMATICAL ASSOCIATION OF AMERICA with a pledge to continue and expand the respected NML tradition.

NML belongs on YOUR bookshelf. *Fill out your collection today! Watch for coming new titles!*

LIST PRICE FOR EACH TITLE: $4
PRICE TO MAA MEMBERS AND HIGH SCHOOL STUDENTS (Prepaid only): $3

NUMBERS: RATIONAL AND IRRATIONAL by Ivan Niven, NML-01

WHAT IS CALCULUS ABOUT? by W. W. Sawyer, NML-02

AN INTRODUCTION TO INEQUALITIES, by E. F. Beckenbach, and R. Bellman, NML-03

GEOMETRIC INEQUALITIES, by N. D. Kazarinoff, NML-04

THE CONTEST PROBLEM BOOK. Problems from the Annual High School Mathematics Contests sponsored by the MAA, NCTM, Mu Alpha Theta, The Society of Actuaries, and the Casualty Actuarial Society. Covers the period 1950-1960. Compiled and with solutions by C. T. Salkind. NML-05.

THE LORE OF LARGE NUMBERS, by P. J. Davis, NML-06

USES OF INFINITY, by Leo Zippin, NML-07

GEOMETRIC TRANSFORMATIONS, by I. M. Yaglom, translated by Allen Shields, NML-08

CONTINUED FRACTIONS, by C. D. Olds, NML-09

GRAPHS AND THEIR USES, by Oystein Ore, NML-10

HUNGARIAN PROBLEM BOOKS I and II, based on the Eötvös Competitions 1894-1905 and 1906-1928. Translated by E. Rapaport, NML-11 and NML-12.

EPISODES FROM THE EARLY HISTORY OF MATHEMATICS, by A. Aaboe, NML-13

GROUPS AND THEIR GRAPHS, by I. Grossman and W. Magnus, NML-14

THE MATHEMATICS OF CHOICE, by Ivan Niven, NML-15

FROM PYTHAGORAS TO EINSTEIN, by K. O. Friedrichs, NML-16

THE MAA PROBLEM BOOK II. A continuation of NML-05 containing problems and solutions from the Annual High-School Mathematics Contests for the period 1961-1965.

FIRST CONCEPTS OF TOPOLOGY, by W. G. Chinn and N.E. Steenrood, NML-18

GEOMETRY REVISITED, by H.S.M. Coxeter, and S. L. Greitzer, NML-19

INVITATION TO NUMBER THEORY, by Oystein Ore, NML-20

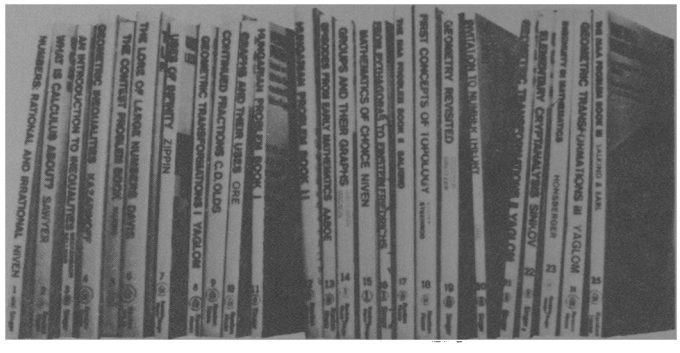
GEOMETRIC TRANSFORMATIONS II, by I. M. Yaglom, translated by Allen Shields, NML-21

ELEMENTARY CRYPTANALYSIS—A Mathematical Approach, by Abraham Sinkov, NML-22

INGENUITY IN MATHEMATICS, by Ross Honsberger, NML-23

GEOMETRIC TRANSFORMATIONS III, by I. M. Yaglom, translated by Abe Shenitzer, NML-24

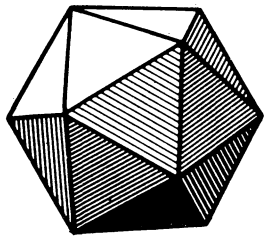
THE MAA PROBLEM BOOK III. A continuation of NML-05 and NML-17, containing problems and solutions from the Annual High School Mathematics Contests for the period 1966-1972.



Send orders to: **The Mathematical Association of America**
1225 Connecticut Ave., NW, Washington, D.C. 20036

MATHEMATICS MAGAZINE

Volume 49, Number 4, September 1976



COVER: Part of a map with 341 regions from the collection of Edward F. Moore. This map (reproduced in full on p. 219) illustrates that, in some sense, no simple "proof" of the four color conjecture can work. A report on the recent computer-based verification of this famous conjecture is given in the News and Letters section.

ARTICLES

NOTES

PROBLEMS

NEWS AND LETTERS

EDITORIAL POLICY

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

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

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

ABOUT OUR AUTHORS

Joseph A. Gallian ("The Search for Finite Simple Groups") holds a Ph.D. from Notre Dame and now teaches at the University of Minnesota at Duluth. His investigation of the range problem for simple groups began in 1972 when he selected as a problem suitable for investigation by an undergraduate research team the question of determining which integers between 1 and 500 could be orders of simple groups. An extensive literature search together with personal conversations, especially with participants of the Park City Conference on Finite Groups in 1975, provided sufficient information to complete the history of the range problem through order 1,000,000. Professor Gallian's interest in bringing group theory to a college audience has also produced "Computers in Group Theory" (this *Magazine*, March 1976) and "Group Theory and the Design of a Letter Facing Machine" which will appear in a forthcoming issue of the *Amer. Math. Monthly*.

Romae Cormier and Roger Eggleton ("Counting by Correspondence") whose common research interests include combinatorics, number theory and graph theory, collaborated on the present paper as the result of a seminar discussion on why the answer to a certain combinatorial problem was of so simple a form. Eggleton comes from Australia where he studied at the University of Melbourne before earning his doctorate from the University of Calgary in 1973. He was a visiting fellow at the Weizmann Institute of Science in 1973–74 and has taught both in Australia and the U.S. Cormier studied at the Universities of Tennessee and Missouri while consulting for Vitro Corp and General Electric. Presently both Cormier and Eggleton are at Northern Illinois University.

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

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

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

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

The Search for Finite Simple Groups

The eighty year quest for the building blocks of group theory reflects sporadic growth spurts whenever new basic techniques were discovered.

JOSEPH A. GALLIAN
University of Minnesota, Duluth

At present, simple group theory is the most active and glamorous area of research in the theory of groups and it seems certain that this will remain the case for many years to come. Roughly speaking, the central problem is to find some reasonable description of all finite simple groups. A number of expository papers [36], [42], [45], [47], [49], [79] and books [21], [46], [67] detailing progress on this problem have been written for professional group theorists, but very little has appeared which is accessible to undergraduates. (Only Goldschmidt's proof of the Brauer-Suzuki-Wall theorem [44] comes to mind.) This paper is intended as a historical account of the search for simple groups for readers who are not experts in the subject. It is the hope of the author that the paper may profitably be read by one who is conversant with the contents of Herstein's algebra book [55]. A complete discussion of all important contributions to simple group theory is beyond the scope of this paper.

What are simple groups and why are they important? Évariste Galois (1811–1832) called a group simple if its only normal subgroups were the identity subgroup and the group itself. The Abelian simple groups are the group of order 1 and the cyclic groups of prime order, while the nonabelian simple groups generally have very complicated structures. These groups are important because they play a role in group theory somewhat analogous to that which the primes play in number theory or the elements do in chemistry; that is, they serve as the "building blocks" for all groups. These "building blocks" are called the composition factors of the group and may be determined in the following way. Given a finite group G, choose a maximal normal subgroup G_1 of $G = G_0$. Then the factor group G_0/G_1 is simple, and we next choose a maximal normal subgroup G_2 of G_1. Then G_1/G_2 is also simple, and we continue in this fashion until we arrive at $G_n = \{e\}$. The simple groups $G_0/G_1, G_1/G_2, \ldots, G_{n-1}/G_n$ are the composition factors of G and by the Jordan-Hölder theorem these groups are independent of the choices of the normal subgroups made in the process described. In a certain sense, a group can be reconstructed from its composition factors and many of the properties of a group are determined by the nature of its composition factors. This, and the fact that many questions about finite groups can be reduced (by induction) to questions about simple groups, make clear the importance of determining all nonabelian finite simple groups.

The narrative which follows is divided into 16 sections which appear in more or less chronological

order according to theme. Within a particular section however, we usually include a number of results which are related to the theme without regard to time. Thus, for example, the section on the odd order problem appears early in the paper but includes results ranging from 1895 to 1963. In this way we hope to emphasize two points: (1) the problems of one generation very often have deep roots in the work of previous generations and (2) there is frequently a large temporal gap between certain results and their subsequent improvements.

Throughout the remainder of this paper we use the term simple group to mean a finite nonabelian simple group.

1. The alternating groups and the classical linear groups

Although Galois had formulated the definition of a simple group and had observed that the alternating group (of even permutations) on 5 symbols was simple, the first major results in the theory were due to Camille Jordan (1838–1922). In 1870, Jordan published *Traité des Substitutions*, the first book ever written on group theory [**58**]. In this book he established the existence of five infinite families of finite simple groups. One of these families, which we denote by A_n, consists of the alternating permutation groups on $n > 4$ symbols. Jordan formed the other four families by using matrices with entries from finite fields. One of these may be described as follows. For $m > 1$, the special linear group $SL(m, p^n)$ is the multiplicative group of $m \times m$ matrices of determinant 1 with entries from the field with p^n elements and the projective special linear group $PSL(m, p^n)$ is the factor group $SL(m, p^n)/Z(SL(m, p^n))$ where $Z(SL(m, p^n))$, the center of $SL(m, p^n)$, is the subgroup of $SL(m, p^n)$ consisting of all scalar matrices with determinant 1. Jordan proved that $PSL(m, p)$ is simple when (m, p) is not $(2, 2)$ or $(2, 3)$. The other three families have been given the names orthogonal, unitary and symplectic groups and, following Hermann Weyl, mathematicians refer to these four families collectively as the classical simple groups.

The last three types mentioned above are most easily defined as certain groups of invertible linear transformations of a finite dimensional vector space V over a finite field modulo the center of the group and in each case the group is obtained by considering those transformations T which leave a nondegenerate form f of V invariant (i.e., f is a certain function from $V \times V$ into the field and $f(Tx, Ty) = f(x, y)$ for all x, y in V). A symmetric bilinear form (i.e., a dot product) gives an orthogonal group; hermitian, a unitary group; and skew symmetric bilinear, a symplectic group. The precise definitions of these groups are not needed here but the reader can find them in [**2**] and [**6**]. Jordan introduced these three families as groups of matrices instead of groups of linear transformations and proved they are simple when the field has prime order (except for a few trivial cases).

2. Range problem 1–660

A different approach was taken by Otto Hölder (1859–1937) when in 1892 he initiated what we will call the range problem; namely the complete determination of all simple groups whose orders are in a given range. Here both the existence and the uniqueness questions must be considered; that is, it must be determined which integers in the range are the orders of simple groups and, for each such integer, all possible simple groups of that order must also be determined (up to isomorphism). Hölder [**56**] proved that the only two simple groups whose orders lie between 1 and 200 are A_5 of order 60 and $PSL(2, 7)$ of order 168. F. N. Cole (1861–1927), the first American-born mathematician to publish in group theory, followed Hölder's lead in 1892 [**23**] when he examined the integers between 201 and 500 for simple groups. He was not totally successful for he was unable to prove that A_6 was the unique simple group of order 360; nor was he able to show 432 was not the order of a simple group. He overcame these difficulties [**24**] a year later, however, when he completed the determination of all the simple groups with orders in the range 1 to 660. In addition to the ones in this range already found by Jordan, Cole discovered one more, $PSL(2, 8)$, having order 504. This provided the first example of a simple group not known to Jordan and the first proof of the simplicity of one of the groups $PSL(m, q)$ with q not prime.

Date	Integers	Individual
1892	1–200	Hölder [**56**]
1892–93	201–660	Cole [**23**, **24**]
1895	661–1092	Burnside [**15**]
1900	1093–2000	Ling and Miller [**60**]
1912	2001–3640 (except 2520)	Siceloff [**71**]
1922	2520	Miller [**65**]
1924	3641–6232 (except 5616 and 6048)	Cole [**26**]
1942	5616 and 6048	Brauer [**7**]
1963	6233–20,000	Michaels [**62**]
1972	1–1,000,000 (21 exceptions)	Hall [**49**, **50**]
1975	Hall's exceptions	Beisiegel and Stingl [**3**]

CHRONOLOGY OF THE RANGE PROBLEM: The search for all simple groups of specified orders reveals sporadic progress as new methods made possible sudden bursts of successful analysis on groups of increasingly large order. At present the range problem is completed through groups of order 1,000,000: all simple groups of order less than 1,000,000 are known, but only some of those beyond that order have been discovered.

The methods of Hölder and Cole are of interest. The three Sylow theorems rule out 596 of the first 660 integers as possible orders of simple groups. (In fact, they rule out 9431 of the first 10,000 [**74**].) If G is assumed to be simple and H is a proper subgroup of G, then G is isomorphic to a subgroup of the symmetric group on the cosets of H in G (compare with the proof of Theorem 2.92 in [**55**]). Thus G with order $|G|$ is represented as a group of permutations on $|G|/|H|$ symbols and it follows that $(|G|/|H|)!$ is a multiple of $|G|$. This last fact is called the index theorem and it further reduces the list of integers to be examined to 47. Finally, a combination of the Sylow theorems, the index theorem, and other elementary techniques such as counting elements reduces to 33 the list of those integers from 1 to 660 which require *ad hoc* arguments. Since the theory of permutation groups was much further developed than the theory of abstract groups at that time, these remaining 33 integers were handled with permutation group techniques. An example of a permutation-type argument will be given later.

It is noteworthy that while the proofs of the non-simplicity of a group of order 144 or 180 occupied more than 10 pages of Hölder's paper, the author has had undergraduates [**59**] who have done this in less than 2 pages using only the results found in Herstein [**55**]. Similarly, using a bit more machinery, three undergraduates from the University of Wisconsin [**27**] covered all the integers up to 1000 with the exceptions of 720 and the uniqueness question. Their proofs for the cases 144 and 180 require only 12 lines.

3. $PSL(m, p^n)$

Cole's discovery of the simplicity of $PSL(2, 8)$ had far-reaching consequences because that same year E. H. Moore (1862–1932), the first mathematics department chairman of the University of Chicago, used it for the starting point of his investigations which resulted [**66**] in a proof that the family of groups $PSL(2, p^n)$ are all simple except when $p^n = 2$ or 3. William Burnside (1852–1927) also obtained this result [**13**] shortly after Moore. Moore's paper, in turn, led his first Ph.D. student, Leonard E. Dickson (1874–1954), to the complete generalization of Jordan's original result when in 1897 he proved [**29**] that the family of groups $PSL(m, p^n)$ $(m > 1)$ consisted of simple groups except when $p^n = 2$ or 3. Dickson called this family a triply infinite system since each of p, n, and m may take on infinitely many values. Moore's paper also contains many interesting results on finite fields, the most important of which is that for each prime power p^k there exists a unique field of order p^k (Galois had proved such fields exist in 1830 [**41**]). In the opinion of E.T. Bell [**4**, p.10] these results on finite fields clearly mark the beginning of abstract algebra in America.

4. Range problem to 1092

In 1895 Burnside [**14**, **15**] obtained several powerful arithmetic tests for simple groups. By far the most important of these is the fact that a simple group of even order must be divisible by one of 12, 16 or 56. In proving this result, he showed that an even order simple group cannot have a cyclic Sylow 2-subgroup. This theorem appears to be the first nonsimplicity criterion which is based on the structure of the Sylow 2-subgroups. In the past two decades much of the research in simple group theory has dealt with the problem of classifying all simple groups whose Sylow 2-subgroups have a specified structure. (All the Sylow 2-subgroups of a group are isomorphic.) For example, John Walter [**77**], in a long (110 pages!) and difficult proof, obtained a broad generalization of Burnside's result when he determined all simple groups with Abelian Sylow 2-subgroups. Similarly, all simple groups whose Sylow 2-subgroups are dihedral (that is, are groups of symmetries of some regular n-sided polygon) have been determined by Daniel Gorenstein and Walter [**48**]. The proof of this important result appears in three papers and runs 160 pages! Commenting on this proof in *Mathematical Reviews* (32 #7634), John Thompson wrote "The techniques of these papers cover the spectrum of finite group theory more thoroughly than any single paper known to the reviewer."

By 1893 the range problem had been completed as far as 660. Since 1092 was the next integer known to be the order of a simple group, Burnside [**15**] decided to examine the integers between these two. The arithmetical tests of his previous paper disposed of all but 17 of the 432 integers in this range and the Sylow theorems ruled out six more. The remaining 11 integers were considered individually although his proof for the hardest integer in the range, 720, was erroneous and he inadvertently omitted 1008. The efficiency of Burnside's nonsimplicity tests is further evidenced by the fact that they dispose of all odd integers up to 2025 and all but 14 of the odd integers less than 9000. As a rule, even integers are much harder to eliminate than odd integers but Burnside's "12, 16, 56 theorem" alone rules out 3691 of the first 5000 even integers [**74**].

5. Permutation representations and character theory

In obtaining their results Burnside, Cole, and Hölder utilized permutation representations of groups. Certain permutation groups—transitive, doubly transitive, and primitive—play an especially important role in simple group theory. (A permutation group on a set S is called transitive on S if for each pair a, b of letters of S there is an element in G which sends a to b; G is called doubly transitive on S if for each two ordered pairs of distinct letters of S, (a, a') and (b, b'), there is an element in G which sends a to b and a' to b'; see [**78**, p. 15] for the definition of a primitive group.) The reasons for the importance of these groups are that the representation of a group as permutations of the cosets of a subgroup is transitive and that many of the known simple groups can be represented as a doubly transitive (and therefore primitive) permutation group. Thus, a common technique when dealing with a simple group G is to represent it as a transitive, doubly transitive or primitive group and then utilize the theory of these groups to obtain important information about G.

Much effort was devoted in the late 1800's and early 1900's to classifying the transitive and primitive permutation groups of low degree. These results often prove useful in simple group theory. To illustrate, let us consider Siceloff's proof [**71**] that there is no simple group of order $1188 = 2^2 \cdot 3^3 \cdot 11$. If G were a simple group of order 1188, Sylow's theorem implies G has 12 subgroups of order 11 which are conjugate in G. If for each element g in G we define T_g to be the mapping which sends the Sylow 11-subgroup S to the Sylow 11-subgroup $g^{-1}Sg$, we see that G may be viewed as a transitive permutation group on the set of Sylow 11-subgroups of G (cf. proof of 2.92 in [**55**]). Then letting H denote a subgroup of G which consists of all permutations which have some Sylow 11-subgroup fixed, it follows that $|H| = 3^2 \cdot 11$ (see [**78**, p. 51]) and H is a permutation group on the other 11 Sylow 11-subgroups. By Sylow's theorem H has an element of order 11 and so this element is an 11-cycle. Thus H is a transitive permutation group of degree 11 and order 99. But Cole [**25**] has shown no such group exists.

A homomorphism from a group into a group of matrices with entries from some field is called a

representation of the group. If T is a representation of G, the character of this representation is the function X from G to the field defined by $X(g) = \text{trace}(T(g))$ for all g in G. There exist numerous arithmetical relations on the characters of a group G which are intimately related to the structure of G. Thus a knowledge of the characters of a group reveals much information about the group itself. The theory of group characters has profoundly influenced the search for simple groups. This theory was developed by Georg Frobenius (1849–1917) in a series of papers beginning in 1896. (The historical background to Frobenius' creation of group characters is detailed in [**51**, **52**].) Around the turn of the century Issai Schur (1875–1941) and especially Burnside simplified the theory and found many important applications of it. In recent times, character theory has been further developed and refined by Brauer, Suzuki, and Feit.

6. Odd order problem

During the period 1895–1901 much attention was focused, particularly by Burnside, on the possibility of the existence of a simple group of odd order. In his 1895 paper [**15**], Burnside had shown that there is no simple group of odd order less than 2025. He later extended this to 9000 [**16**] and then to 40,000 [**18**]. Numerous arithmetical theorems obtained by Burnside in this period reduced the list of possible odd orders less than 40,000 to 7; these were then eliminated by elementary considerations.

In 1901 Burnside [**17**] used character theory to prove that a nonsolvable transitive permutation group of prime degree is doubly transitive. Since a simple group which has a subgroup of prime index can be represented as a transitive permutation group on the cosets of this subgroup it must be doubly transitive. But the order of a doubly transitive group of degree n is divisible by $n(n-1)$ [**78**, p. 20] so Burnside's result shows there is no odd order simple group which has a subgroup of prime index. Burnside [**18**] also proved in 1901 that if a simple group has odd order n and p is the smallest prime divisor of n then n is divisible either by p^4 or by both p^3 and a prime factor of p^2+p+1.

Burnside's efforts convinced him that there were no simple groups of odd order and that the eventual proof of this would involve the use of character theory. In fact, he wrote [**20**, p. 503] "The contrast that these results shew between groups of odd and of even order suggests inevitably that simple groups of odd order do not exist." He further wrote [**17**] "The results obtained in this paper, partial as they necessarily are, appear to me to indicate that an answer to the interesting question as to the existence or non-existence of simple groups of odd order may be arrived at by further study of the theory of group characters."

The next important step in this direction however, did not come for more than 50 years. In 1957, Michio Suzuki [**72**] used character theory to prove that a simple group in which the centralizer of any nonidentity element is Abelian must have even order. (The centralizer of an element x in a group G is the subgroup $C(x) = \{g \in G \mid gx = xg\}$.) Three years later, in a major work [**37**], Walter Feit, Marshall Hall, Jr., and John Thompson obtained a broad generalization of Suzuki's result by showing that "Abelian" could be replaced by the much weaker condition "nilpotent." (A group is nilpotent if all of its Sylow subgroups are normal.) Their proof was similar to Suzuki's and character theory played an important role in it.

Burnside's prophecy was at last fulfilled in 1963 when Feit and Thompson expanded on the ideas of the two papers mentioned above and proved [**38**] that groups of odd order are solvable. (A finite group is solvable if all of its composition factors have prime order; thus, solvable groups are not simple.) The difficulty of this proof and the significance of both the theorem and the methods employed cannot be exaggerated. Concerning one portion of the proof, Suzuki wrote in *Mathematical Reviews* (29 #3538) "... [This 50 page portion] represents one of the highest points ever achieved in the theory of finite groups."

The proof of the "Odd Order Theorem" occupies an entire 255 page issue of the *Pacific Journal of Mathematics*. It proceeds by assuming that there is a group G of minimal odd order which is not solvable. Then every proper subgroup of G is solvable and therefore Philip Hall's extensive work on solvable groups could be brought to bear on the subgroups. Ultimately, they were able to derive a

contradiction. For their achievement, Feit and Thompson were awarded the Frank Nelson Cole Prize in Algebra by the American Mathematical Society in 1965. (The Cole Prize is named after the same Cole who had determined the simple groups with orders between 201 and 660 and was established in his honor in recognition of his many years of service to the Society.)

7. Dickson's simple groups

In the period from 1897 to 1905 Dickson made many fundamental contributions to the theory of simple groups. In a series of papers appearing from 1897 to 1899 he extended Jordan's results on the simplicity of the orthogonal, unitary and symplectic groups over fields of prime order to arbitrary finite fields. Much of this work emanated from his Ph.D. dissertation, the first one ever done in mathematics at the University of Chicago. Whether there exist two nonisomorphic simple groups of the same order had been a long-standing problem by 1899. But Dickson's proof in 1897 that $PSL(3,4)$ is simple provided a possible answer to this question since it and the simple group A_8 both have order 20,160. It was quickly suspected that these two were not isomorphic since A_8 contains elements of orders 6 and 15 while no such elements were known to be in $PSL(3,4)$. At Moore's suggestion, Ida Schottenfels investigated these two groups and proved [**70**] they were not isomorphic. Shortly thereafter, Dickson showed [**30**] that there are infinitely many such examples. Since these examples were given by Dickson no others have been found and there is no known triple of nonisomorphic simple groups of equal order. After 20,160 the next known integer for which there is a pair of nonisomorphic simple groups of equal order is 4,585,351,680, and it wasn't until the mid 1960's that 20,160 was shown to be the smallest possible integer for which this can happen.

In his classic book *Linear Groups* Dickson listed all the isomorphisms between the simple groups he knew. For example, A_5, $PSL(2,4)$, and $PSL(2,5)$ are defined differently but are isomorphic. The question of whether Dickson's list of isomorphisms contained all which were possible among the simple groups known to him was not answered until 50 years later when Jean Dieudonné proved [**35**]

A Chronological Collection of ...

1870	Jordan	Established simplicity of alternating groups and linear groups over fields of prime order.
1892	Hölder	Began range problem.
1895–1900	Cole, Miller	Proved simplicity of Mathieu groups.
1896–1901	Frobenius-Burnside	Developed character theory.
1897–1905	Dickson	Established simplicity of linear groups over arbitrary finite fields. Discovered a family of simple groups of Lie type.
1904	Burnside	Proved p^aq^b theorem.
1954	Brauer	Began the program of characterizing simple groups in terms of centralizers of involutions.
1955	Chevalley	Discovered new approach to simple groups. Discovered new families of simple groups of Lie type.
1958–1961	Steinberg, Tits, Hertzig, Ree	Extended Chevalley's methods and discovered new infinite families of simple groups of Lie type.

that Dickson's list was complete. Dickson also listed all the coincidences in the orders of the simple groups known to him but whether this list was complete was not determined until 1955 when Emil Artin (1898–1962) proved [1] with an elegant number-theoretical study that it was.

Without going into detail we mention that the classical linear groups over the field of complex numbers are Lie groups (because, roughly, they possess a smooth geometric structure) and Wilhelm Killing (1847–1923) and Elie Cartan (1869–1951) proved (1888–1894) that besides the simple Lie groups corresponding to the classical groups there are only five additional simple ones called exceptional Lie groups. In two papers in 1901 and 1905 Dickson discovered a new infinite family of finite simple groups by defining analogs over finite fields of one of these exceptional Lie groups [**31**, **32**]. It is remarkable that no additional new finite simple groups were found until Claude Chevalley and others, 50 years later, were able to show (among other results) that the remaining four exceptional Lie groups also had finite analogs.

8. The Mathieu groups

In 1861 E. Mathieu discovered a family of five transitive permutation groups. This remarkable family has become very important in both the theory of simple groups and coding theory as well as in permutation group theory. In 1895, while determining all transitive permutation groups on 10 or 11 symbols, Cole observed [**25**] that the smallest Mathieu group (order 7920) is simple and by 1900 G. A. Miller (1863–1951) had shown [**63**, **64**] the other four are also simple. Three of these groups have order less than 1,000,000 and this brought to 53 the number of such simple groups known in 1900. This number would not be enlarged until 1960.

Among all the simple groups known by 1905 the Mathieu groups had the peculiar distinction of being the only ones which were not part of an infinite family of simple groups (such as A_n or $PSL(m, p^n)$). To this date they (and 21 or so other simple groups) still have not been shown to be members of any infinite family of simple groups in a natural way.

... Highlights in the Theory of Simple Groups

1960	Suzuki	Discovered new infinite family of simple groups (only simple groups with orders not divisible by 3).
1963	Feit-Thompson	Proved simple groups have even order.
1965	Gorenstein-Walter	Classified all simple groups with dihedral Sylow 2-subgroups.
1966–1975	Janko, Hall, Higman, Sims, McKay, McLaughlin, Suzuki, Held, Conway, Thompson, Fischer, Lyons, Rudvalis, Wales, O'Nan, Smith	Discovered new sporadic simple groups.
1968	Thompson	Proved N-theorem. Classified all minimal simple groups.
1969	Walter	Classified all simple groups with Abelian Sylow 2-subgroups.
1971	Thompson	Proved Suzuki groups are the only simple groups with orders not divisible by 3.
1972	Wales	Classified all simple groups with orders of the form p^aq^br.
1975	Hall, Beisiegel-Stingl	Completed range problem to 1,000,000.

9. Range problem to 6232

The determination of all simple groups of order up to 2000 was completed in 1900 when G. Ling and Miller proved [**60**] that there is no simple group whose order is between 1093 and 2000. It is interesting to note that although there are 908 integers in this range only 28 required any special treatment. For example, Burnside's result on odd order simple groups and his "even, but not divisible by 12, 16 or 56" theorem eliminated all but 118 possibilities. Then the Sylow theorems, the index theorem, and several results of Frobenius on groups with orders of the form $p^a q^b$ reduced the possibilities to 28.

In 1902 Frobenius determined [**40**] all transitive permutation groups of degree $p+1$ and order $p(p^2-1)/2$ where p is prime. Since any simple group of order $p(p^2-1)/2$ can be represented as a transitive permutation group on the $p+1$ conjugate Sylow p-subgroups, it follows from Frobenius' theorem that $PSL(2,p)$ is the only simple group of order $p(p^2-1)/2$. This appears to be the first arithmetical characterization of an infinite family of simple groups. Fifty-six years later this uniqueness theorem was generalized in answer to a question of Artin. Artin had observed that the simple groups $PSL(2,p)$ $(p>3)$ and $PSL(2,2^n)$, where 2^n+1 is prime, have orders which are divisible by a prime whose cube exceeds the order of the group. He conjectured that these were the only such simple groups and Brauer and Reynolds [**11**] used modular character theory (the character values lie in a field of prime characteristic) to prove this conjecture.

Twelve years after Ling and Miller completed the range problem up to 2000, L. P. Siceloff, at the suggestion of Cole, proved [**71**] that the only integers between 2001 and 3640 which are orders of simple groups are 2448, 2520, and 3420. All of these were on Dickson's list in 1901. Since 2448 and 3420 have the form $p(p^2-1)/2$ the uniqueness question concerning these integers had been answered affirmatively by Frobenius ten years earlier and it was Miller who ten years afterward showed [**65**] that A_7 is the unique simple group of order 2520. Thus by 1922 all simple groups of order up to 3640 had been determined.

Again there was a twelve year hiatus before the exhaustive enumeration of integers in a certain range was continued. In 1924, Cole returned to the problem again and showed [**26**] that the only integers between 3641 and 6232 which are orders of simple groups were the four on Dickson's 1901 list. Unfortunately Cole's paper was so lacking in detail that its value was diminished. The uniqueness question for $PSL(2,23)$ (order 6072) had previously been settled affirmatively by Frobenius, and Cole did the same for $PSL(2,16)$ (order 4080). Eighteen years after Cole's paper Brauer [**7**] showed, with the use of character theory, that the remaining two integers also corresponded to unique simple groups. So by 1942 all simple groups of order as far as 6232 had been determined.

10. Burnside's $p^a q^b$ theorem

In 1904, Burnside used character theory to prove [**19**] that every group of order $p^a q^b$ where p and q are primes is solvable. This theorem represented the final generalization of a large number of special cases which had been established by Sylow, Frobenius, Burnside, Jordan, and Cole; it has become the classic example of the power of character theory. With character theory a simple proof was possible 70 years ago, but a character-free proof of Burnside's theorem, although long sought, has appeared only in the past few years. Thompson had indicated in his fundamental paper [**75**] on minimal simple groups (see section 14) that a character-free proof of the Burnside theorem could be extracted from that paper and the "odd order paper" [**38**]. David Goldschmidt [**43**], in 1970, gave a short character-free proof of the theorem when p and q are odd and Helmut Bender [**5**] two years later proved the general result without character theory. By combining the arguments of Bender in the odd order case and of Hiroshi Matsuyama [**61**] in the even order case it is now possible to obtain a short and attractive character-free proof of the $p^a q^b$ theorem.

Despite the early outstanding achievements with character theory by Burnside and Frobenius, others seemed to ignore it as a tool in simple group theory until Brauer brought it to the forefront in

the 1940's and 50's. This is partly explained by the fact that interest in simple groups began to wane around 1905.

11. The Chevalley groups

In 1948 Dieudonné [**34**] classified all the known simple groups according to their method of construction. This classification scheme modernized the one Dickson had devised in 1901. Whereas Dickson obtained his results by means of complicated matrix calculations which often obscured the underlying ideas, Dieudonné utilized the geometric properties of linear transformations and vector space theory to simplify and clarify Dickson's work. Although Dieudonné's approach to the classical linear groups is much more elegant than Dickson's, his methods still required that each family of simple groups be treated individually and he found no new simple groups. Thus, simple group theory was revitalized in 1955 when Chevalley, in his celebrated paper [**22**], introduced a new approach which provided a uniform method for investigating three kinds of the classical linear groups and Dickson's simple groups of Lie type as well. In addition to encompassing most of the then-known simple groups, his method also yielded several new infinite families of finite simple groups.

This was accomplished in the following way. With every pair (L, K) where L is a Lie algebra over the complex numbers (i.e., an algebra whose associative law is replaced by the Jacobi identity and also satisfies the condition $x^2 = 0$ for all x) and K is a field, a new Lie algebra L_K is constructed. Chevalley was able to associate with every such pair a certain subgroup of the automorphism group of L_K which is simple. With the appropriate choice of L and K these simple groups are those investigated by Jordan, Dickson, and Dieudonné. With other choices of L and K Chevalley obtained his new simple groups (the smallest of these has order $2^{24}3^65^27^213 \cdot 17$). These groups were the first new simple groups found in more than 50 years. That they were indeed new was established by comparing their orders with the orders of the simple groups which had been known. The formulas for the orders of the Chevalley groups over finite fields were derived by using topological properties of the Lie group of the same type. Artin [**1**] developed a new classification scheme for the known simple groups which included the Chevalley groups. He used fewer classes than did Dickson and Dieudonné and his method considerably improved theirs (see also [**2**]).

12. Groups of Lie type

During the period 1958–1959 Chevalley's methods were extended and modified by Robert Steinberg, Jacques Tits, and D. Hertzig (see [**21**]) to obtain additional new infinite families of simple groups and the classical groups not handled by Chevalley. Shortly thereafter, Suzuki [**73**], while in the process of classifying a certain type of doubly transitive permutation groups, also discovered another new infinite family. Analyzing the Suzuki groups, Rimhak Ree noticed that when interpreted from a Lie-theoretical point of view, they were closely related to a certain family of Chevalley groups. He then showed that the method of Steinberg could be used to construct the Suzuki groups. This in turn led him to investigate two other similar situations and eventually discover his two families of simple groups [**68**, **69**]. The Suzuki and Ree groups together with those of Chevalley and Steinberg are collectively referred to as the simple groups of Lie type. These, together with the alternating groups A_n $(n \geqq 5)$ account for all but 26 or so of the finite simple groups known to date.

The Suzuki groups are noteworthy for another reason. They provided the first examples of simple groups whose orders are not divisible by 3, and Thompson, in a major classification theorem, has recently shown that these are the only possible such groups. The elements of the Suzuki groups are certain 4×4 matrices with entries from the Galois fields of order 2^{2n+1}, again illustrating the extremely important role that matrix groups over finite fields play in simple group theory.

13. Sporadic simple groups

A simple group which no one has yet been able to fit into an infinite class of simple groups in a natural way is called a sporadic simple group. For example, A_n and $PSL(n, q)$ are infinite families of

The Known Finite Simple Groups ...

Date and Discoverer	Type: *Name*	Group Notation	Order (p is a prime, $q = p^n$)
1870 Jordan	— *alternating group of degree m*	$A_m\ (m > 4)$	$m!/2$
1870 Jordan	classical linear *projective special linear*	$L_m(p)$ or $PSL(m, p)\ (m > 1)$	$d^{-1}p^{m(m-1)/2}\prod_{i=2}^{m}(p^i - 1)$; $d = (m, p-1)$
1870 Jordan	classical linear *symplectic*	$S_{2m}(p)\ (m > 1)$	$d^{-1}p^{m^2}\prod_{i=1}^{m}(p^{2i} - 1)$; $d = (2, p-1)$
1870 Jordan	classical linear *orthogonal*	$O_{2m}(\varepsilon, p),\ \varepsilon = \pm 1,\ (m > 3)$	$d^{-1}p^{m(m-1)}(p^m - \varepsilon)\cdot\prod_{i=1}^{m-1}(p^{2i} - 1)$; $d = (4, p - \varepsilon)$
1870 Jordan	classical linear *unitary*	$U_m(p)\ (m > 2)$	$d^{-1}p^{m(m-1)/2}\prod_{i=2}^{m}(p^i - (-1)^i)$ $d = (m, p+1)$
1893 Cole	classical linear $L_2(8)$	$L_2(8)$ or $PSL(2, 8)$	$504 = 2^3 \cdot 3^2 \cdot 7$
1893 Moore	classical linear *projective special linear*	$L_2(q)$ or $PSL(2, q)\ (q > 3)$	$d^{-1}q(q^2 - 1);\ d = (2, q-1)$
1895 Mathieu-Cole	sporadic *Mathieu* 11	M_{11}	$7920 = 2^4 \cdot 3^2 \cdot 5 \cdot 11$
1897 Dickson	classical linear *projective special linear*	$L_m(q)$ or $PSL(m, q)(m > 1)$	$d^{-1}q^{m(m-1)/2}\prod_{i=2}^{m}(q^i - 1)$ $d = (m, q-1)$
1897 Dickson	classical linear *symplectic*	$S_{2m}(q)\ (m > 1)$	$d^{-1}q^{m^2}\prod_{i=1}^{m}(q^{2i} - 1)$; $d = (2, q-1)$
1898 Dickson	classical linear *orthogonal*	$O_{2m}(\varepsilon, q),\ \varepsilon = \pm 1,\ (m > 3)$	$d^{-1}q^{m(m-1)}(q^m - \varepsilon)\cdot\prod_{i=1}^{m-1}(q^{2i} - 1)$; $d = (4, q^m - \varepsilon)$
1898 Dickson	classical linear *unitary*	$U_m(q)\ (m > 2)$	$d^{-1}q^{m(m-1)/2}\cdot\prod_{i=2}^{m}(q^i - (-1)^i)$ $d = (m, q+1)$
1899 Mathieu-Miller	sporadic *Mathieu* 12	M_{12}	$95{,}040 = 2^6 \cdot 3^3 \cdot 5 \cdot 11$
1900 Mathieu-Miller	sporadic *Mathieu* 22	M_{22}	$443{,}520 = 2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11$
1900 Mathieu-Miller	sporadic *Mathieu* 23	M_{23}	$10{,}200{,}960 = 2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 23$

simple groups while the five Mathieu groups are sporadic. In the classification schemes of Dickson, Dieudonné, and Artin only the five Mathieu groups are sporadic and no new ones were discovered until Zvonimir Janko found one of order 175,560 in 1966 [**57**]. Remarkably, the discoveries of Mathieu and Janko are separated by more than one hundred years. Using other results of Janko, M. Hall proved [**49**] the existence of a sporadic simple group of order 604, 800 in 1967. This brought to 56 the number of known simple groups of order less than 1,000,000 and to 7 the number of sporadic simple groups. Paraphrasing Gorenstein [**47**, p. 14] we recount the following anecdote in connection with Hall's simple group:

> Shortly after Hall constructed his group he gave a lecture on it at Oxford. Donald Higman and Charles Sims were in the audience and they were both struck by the fact that one of the Mathieu groups had certain permutation properties analogous to those of a group which Hall had used in his construction. That very same night this observation led them to the construction of a new sporadic simple group!

Hall's method has also led to the discovery of two others by analogous methods. By now there are 26 or so sporadic simple groups and it has been proved that there are exactly 56 simple groups of order less than 1,000,000. (Of late, new sporadic simple groups are being discovered so frequently that it is difficult for one to be sure of their precise number.)

Some of the sporadic simple groups have been discovered in the course of solving certain problems in permutation group theory while others have turned up as the automorphism group of a distance-transitive graph (see Chapter 4 and the Appendix in [**6**]). Quite often, two or more sporadic groups are related in some way. Indeed, the Conway .1 simple group contains 12 sporadic simple groups as subgroups! The existence of many of these recently-discovered groups has been verified by means of a permutation representation of the group and extensive use of computers.

An important technique which has led to the discovery of a number of sporadic simple groups involves the notion of the centralizer of an involution (i.e., of an element of order 2). This method is employed in the following way. Choose H to be the centralizer of an involution from some known simple group G. Next, assume G^* is any simple group which contains an involution x such that $C(x)$ is isomorphic to H. (By a theorem of Brauer and Fowler [**10**] only a finite number of such groups can exist, so H "almost" determines G.) Then a great deal of information about G^* can be obtained. With this information it is often possible to show that G^* is isomorphic to G or to some other known simple group. For example Dieter Held [**53**] has proved such a theorem when $G = A_8$ or A_9. If it cannot be shown that G^* must be isomorphic to some known simple group, the information may be adequate to suggest a method of constructing a new simple group. Held [**54**] has also been instrumental in accomplishing this. He began with the observation that $PSL(5,2)$ and the largest Mathieu group possess involutions with isomorphic centralizers and no other known simple group has this property. Choosing this for H, he was led to three possible configurations for G^*. Ultimately, enough properties of this third group were derived so that Graham Higman and John McKay were able to construct it with the use of a computer.

Similarly, one may proceed by assuming G is a simple group which contains an involution whose centralizer closely resembles the form of a centralizer of an involution from a known simple group. In this case, if the information about G is not self-contradictory it suggests the possible existence of a new simple group and may be sufficient to lead to an actual construction of the group. This is how Janko discovered his simple group of order 175,560. Each member of a family of simple groups discovered by Ree has a centralizer isomorphic to $Z_2 \times PSL(2, 3^n)$ and has its Sylow 2-subgroups isomorphic to $Z_2 \times Z_2 \times Z_2$. Janko set out to determine all simple groups which have a centralizer isomorphic to $Z_2 \times PSL(2, p^n)$, p odd, and with Sylow 2-subgroups isomorphic to $Z_2 \times Z_2 \times Z_2$. Eventually he was able to show that either $p = 3$ and the group is of Ree type or $p^n = 5$. The information he obtained about this latter case led him to write down a pair of 7×7 matrices with entries in the field of order 11 which generated a new simple group.

... Their Types, Notations and Orders ...

Date and Discoverer	Type: *Name*	Group Notation	Order p is a prime, $q = p^n$
1900 Mathieu-Miller	sporadic *Mathieu 24*	M_{24}	$244{,}823{,}040 = 2^{10}\cdot 3^3\cdot 5\cdot 7\cdot 11\cdot 23$
1901 and 1905 Dickson	Lie *groups of type* G_2	$G_2(q)$	$q^6(q^6-1)(q^2-1)$
1955 Chevalley	Lie *Chevalley groups of type* E_4	$E_4(q)$	$q^{24}(q^{12}-1)(q^8-1)\cdot (q^6-1)(q^2-1)$
1955 Chevalley	Lie *Chevalley groups of type* E_6	$E_6(q)$	$d^{-1}q^{36}(q^{12}-1)(q^9-1)\cdot (q^8-1)(q^6-1)(q^5-1)\cdot (q^2-1);\ d=(3,q-1)$
1955 Chevalley	Lie *Chevalley groups of type* E_7	$E_7(q)$	$d^{-1}q^{63}(q^{18}-1)(q^{14}-1)\cdot (q^{12}-1)(q^{10}-1)(q^8-1)\cdot (q^6-1)(q^2-1);\ d=(2,q-1)$
1955 Chevalley	Lie *Chevalley groups of type* E_8	$E_8(q)$	$q^{120}(q^{30}-1)(q^{24}-1)\cdot (q^{20}-1)(q^{18}-1)(q^{14}-1)\cdot (q^{12}-1)(q^8-1)(q^2-1)$
1959 Steinberg-Tits-Hertzig	Lie *twisted groups of type* E_6	${}^2E_6(q^2)$	$d^{-1}q^{36}(q^2-1)(q^5+1)\cdot (q^6-1)(q^8-1)(q^9+1)\cdot (q^{12}-1);\ d=(3,q+1)$
1959 Steinberg-Tits-Hertzig	Lie *twisted groups of type* D_4	${}^3D_4(q^3)$	$q^{12}(q^2-1)(q^6-1)\cdot (q^8+q^4+1)$
1960 Suzuki	Lie *Suzuki groups*	$Sz(q)$ or ${}^2B_2(q)$, $q=2^{2m+1}$	$q^2(q^2+1)(q-1)$
1961 Ree	Lie *Ree groups of type* G_2	${}^2G_2(q)$, $q=3^{2m+1}$	$q^3(q^3+1)(q-1)$
1961 Ree	Lie *Ree groups of type* F_4	${}^2F_4(q)$, $q=2^{2m+1}$	$q^{12}(q^6+1)(q^4-1)\cdot (q^3+1)(q-1)$
1966 Janko	sporadic *Janko*	*Ja*	$175{,}560 = 2^3\cdot 3\cdot 5\cdot 7\cdot 11\cdot 19$
1967 Hall-Janko	sporadic *Hall-Janko*	*HaJ*	$604{,}800 = 2^7\cdot 3^3\cdot 5^2\cdot 7$
1968 Higman-Sims	sporadic *Higman-Sims*	*HiS*	$44{,}352{,}000 = 2^9\cdot 3^2\cdot 5^3\cdot 7\cdot 11$
1969 Hall-Janko-McKay	sporadic *Hall-Janko-McKay*	*HJM*	$50{,}232{,}960 = 2^7\cdot 3^5\cdot 5\cdot 17\cdot 19$
1969 McLaughlin	sporadic *McLaughlin*	*McL*	$898{,}128{,}000 = 2^7\cdot 3^6\cdot 5^3\cdot 7\cdot 11$

This method was also recently used independently by Bernd Fischer and Thompson and by Robert Griess to discover a possible new simple group. This object, called the "Monster" M, is not defined in terms of generators and relations. In fact, it has not been defined at all! What Fischer, Thompson and Griess did was to assume that there exists a finite simple group M that satisfies certain hypotheses. They showed, under this assumption, that M would be in fact a new simple group (i.e., not isomorphic to any existing simple group). They also obtained a great deal of other information about M, such as its order, properties of certain subgroups, and a portion of its character table. Thompson has computed the order of M to be

$$808{,}017{,}424{,}794{,}512{,}875{,}886{,}459{,}904{,}961{,}710{,}757{,}005{,}754{,}368{,}000{,}000{,}000$$
$$= 2^{46}3^{20}5^{9}7^{6}11^{2}13^{3}17\cdot 19\cdot 23\cdot 29\cdot 31\cdot 41\cdot 59\cdot 71$$

(hence the name). If there is a simple group satisfying the stipulated hypothesis we should be able to deduce what it must look like and once we know what it looks like, we should be able to define it. For the Monster M this last step has not yet been accomplished.

A certain section of M (i.e., a group of the form H/K where H and K are subgroups of M and K is normal in H), the "Baby Monster" B, is also a possible new simple group. Fischer has computed the order of B to be $4{,}154{,}781{,}481{,}226{,}426{,}191{,}177{,}580{,}544{,}000{,}000 = 2^{41}3^{13}5^{6}7^{2}11\cdot 13\cdot 17\cdot 19\cdot 23\cdot 41\cdot 47$. Of course, neither of the groups M or B has yet been shown to exist.

In addition to the role the centralizers of involutions play in the discovery of new simple groups, these subgroups are important for another reason. There is presently underway a systematic attempt to use the centralizers of involutions as a means for classifying all the known finite simple groups. This program began in 1954 with Brauer's characterization of $PSL(3, q)$, q odd, and is now almost complete. According to Gorenstein [**47**, p. 21], "Probably more individuals have been involved in this effort than in any other single area of simple group theory." We refer the reader to section 4.4 of [**36**] for a survey of results of this type.

14. Thompson's N-paper

Most simple groups contain other simple groups as subgroups. For example, $A_5 \subset A_6 \subset A_7 \ldots$. On the other hand, a minimal simple group is one, all of whose proper subgroups are solvable. It follows then that every simple group has a minimal simple group as a section. Minimal simple groups are therefore basic and the complete determination of all such groups would clearly be of great value. In the early 1960's Thompson set out to do just this. Such an endeavor was a natural successor to the Odd Order Theorem since the minimal counterexample G in that proof was a minimal simple group. Actually, Thompson decided to tackle a more general classification problem. The normalizer of a nonidentity solvable subgroup of a group G is called a local subgroup of G and an N-group is one in which all local subgroups are solvable. Evidently, every minimal simple group is also an N-group.

As early as 1963, Thompson had concluded that with only finitely many exceptions the simple N-groups were $PSL(2, q)$ $(q > 3)$ and the Suzuki groups. The complete classification of all nonsolvable N-groups however, did not come until several years later. The 407 page proof (!!) [**75**] of this remarkable theorem is spread out over six journal issues during the seven year period 1968–1974. Describing his approach, Thompson writes [**75**, p. 383]:

> In a broad way, this paper may be thought of as a successful transformation of the theory of solvable groups to the theory of simple groups. By this is meant that a substantial structure is constructed which makes it possible to exploit properties of solvable groups to obtain delicate information about the structure and embedding of many solvable subgroups of the simple group under consideration. In this way, routine results about solvable groups acquire great power.

(An essay which outlines the organization of the proof and discusses some of the arguments used is given in [**46**, pp. 473–480].)

Date and Discoverer	Type: *Name*	Group Notation	Order p is a prime, $q = p^n$
1969 Suzuki	sporadic *Suzuki*	*Suz*	448,345,497,600 = $2^{13}\cdot3^7\cdot5^2\cdot7\cdot11\cdot$
1969 Held-Higman-McKay	sporadic *Held-Higman-McKay*	*HHM*	4,030,387,200 = $2^{10}\cdot3^3\cdot5^2\cdot7^3\cdot17$
1969 Conway-Thompson	sporadic *Conway's .1 group*	Co_1	4,157,776,806,543,360,000 = $2^{21}\cdot3^9\cdot5^4\cdot7^2\cdot11\cdot13\cdot23$
1969 Conway-Thompson	sporadic *Conway's .2 group*	Co_2	42,305,421,312,000 = $2^{18}\cdot3^6\cdot5^3\cdot7\cdot11\cdot23$
1969 Conway-Thompson	sporadic *Conway's .3 group*	Co_3	495,766,656,000 = $2^{10}\cdot3^7\cdot 5^3\cdot7\cdot11\cdot23$
1969 Fischer	sporadic *Fischer 22*	Fi_{22}	64,561,751,654,400 = $2^{17}\cdot3^9\cdot5^2\cdot7\cdot11\cdot13$
1969 Fischer	sporadic *Fischer 23*	Fi_{23}	4,089,460,473,293,004,800 = $2^{18}\cdot3^{13}\cdot5^2\cdot7\cdot11\cdot13\cdot17\cdot23$
1969 Fischer	sporadic *Fischer 24*	Fi'_{24}	1,255,205,709,190,661,721,292, 800 = $2^{21}\cdot3^{16}\cdot5^2\cdot7^3\cdot11\cdot13\cdot17\cdot23\cdot2$
1971 Lyons-Sims	sporadic *Lyons-Sims*	*LyS*	51,765,179,004,000,000 = $2^8\cdot3^7\cdot5^6\cdot7\cdot11\cdot31\cdot37\cdot67$
1972 Rudvalis-Conway-Wales	sporadic *Rudvalis*	*Rud*	145,926,144,000 = $2^{14}\cdot3^3\cdot5^3\cdot7\cdot13\cdot29$
1973 O'Nan-Sims	sporadic *O'Nan*	*O'N*	460,815,505,920 = $2^9\cdot3^4\cdot5\cdot7^3\cdot11\cdot19\cdot31$
1974 Fischer	sporadic *Monster*	M or F_1 (possible new simple group)	808,017,424,794,512,875,886, 459,904,961,710,757,005,754,368 000,000,000 = $2^{46}\cdot3^{20}\cdot5^9\cdot7^6\cdot11^2\cdot$ $13^3\cdot17\cdot19\cdot23\cdot29\cdot31\cdot41\cdot47\cdot59\cdot71$
1974 Fischer	sporadic *Baby Monster*	B or F_2 (possible new simple group)	4,154,781,481,226,426,191, 177,580,544,000,000 = $2^{41}\cdot3^{13}\cdot5^6\cdot7^2\cdot11\cdot13\cdot$ $17\cdot19\cdot23\cdot41\cdot47$
1974 Fischer-Smith-Thompson	sporadic *Fischer 3 or Thompson group*	F_3 or E	90,745,943,887,872,000 = $2^{15}\cdot3^{10}\cdot5^3\cdot7^2\cdot13\cdot19\cdot31$
1974 Fischer-Smith	sporadic *Fischer 5 or Harada group*	F_5 or F	273,030,912,000,000 = $2^{14}\cdot3^6\cdot5^6\cdot7\cdot11\cdot19$
1975 Janko	sporadic *Janko 4*	J_4 (possible new simple group)	86,775,571,046,077,562,880 = $2^{21}\cdot3^3\cdot5\cdot7\cdot11^3\cdot23\cdot29\cdot31\cdot37\cdot43$

This result has a number of important corollaries, the most important of which is a classification of all minimal simple groups. A few consequences of this corollary are mentioned in the next two sections. For his many profound contributions to simple group theory, Thompson was awarded the Fields Medal — the mathematical equivalent of the Nobel Prize — by the International Congress of Mathematicians in 1970.

15. The $p^aq^br^c$ problem

One of the corollaries of Thompson's result classifying the minimal simple groups seems to have put the solution of a very difficult, well-known problem within reach. This problem concerns the natural generalization of Burnside's p^aq^b theorem to three primes. Since there are eight known simple groups which have orders divisible by exactly three distinct primes the logical extension of Burnside's result would be the complete determination of all simple groups with orders of the form $p^aq^br^c$. The first steps in this direction were taken by Burnside, Frobenius and E. Maillet. For example, Burnside [**14**] showed that there is no simple group whose order has the form pq^br where $p < q < r$. Fifty years later Brauer and Hsio Tuan [**12**] used character theory to show that except for the groups $PSL(2,5)$ and $PSL(2,7)$, the restriction "$p < q < r$" was unnecessary. In 1962 and again in 1968 Brauer [**8**, **9**] returned to this problem and determined all simple groups whose orders have the form p^aq^br where $a = 1$ or 2 or the form 2^a3^b5.

In spite of the fact that Brauer and his predecessors had solved the "three-prime problem" in numerous special cases, the complete solution was far from sight until Thompson proved his result. Specifically, he proved that a simple group whose order is divisible by exactly three distinct primes must have one of $PSL(2,4)$, $PSL(2,7)$, $PSL(2,8)$, $PSL(2,17)$ or $PSL(3,3)$ as a section. From this it follows that such a group must have order of the form $2^a3^bp^c$ where p is 5, 7, 13 or 17. Then, since character theory is a natural tool for analyzing groups whose orders have a prime to the first power only, David Wales [**76**] used it in conjunction with the N-paper to determine all simple groups (8 of them) whose orders have the form 2^a3^bp. Finally, Kenneth Klinger and Geoffrey Mason are presently in the midst of showing (they hope) that there are no simple groups with orders of the form $2^a3^bp^c$ with $c > 1$. Of course, the completion of this work will finish the $p^aq^br^c$ problem.

16. The range problem to 1,000,000

At the suggestion of Brauer, Sister Michaels, in her 1963 Ph.D. dissertation [**62**], showed that there was no known simple group in the range 6232 to 20,000. Her work was superseded during the late 60's and early 70's when the range problem was taken up by M. Hall [**49**, **50**]. Using a wide assortment of methods from elementary to advanced as well as a computer he succeeded in eliminating all but 21 of the first 1,000,000 integers as possible orders for new simple groups. For the integers not eliminated by elementary considerations, the theory of modular characters and Thompson's result on minimal simple groups played an important role. The character theory yields integer equations which certain parameters of the group must satisfy and a computer was used to make the verifications. Every simple group must have a section which is a minimal simple group so the order of any simple group is divisible by the order of a minimal simple group.

Hall was able to show that from among Thompson's list of minimal simple groups only $PSL(2,5)$, $PSL(2,7)$, $PSL(2,8)$, $PSL(2,13)$, $PSL(2,17)$, $PSL(3,3)$, $PSL(2,23)$ and $PSL(2,27)$ could occur as a section of an unknown simple group of order less than 1,000,000. From a result of Gorenstein it then follows that such a group has order divisible by 840 or 2184. Eventually the list was pared down to 1146 integers which required individual consideration. The first paper [**49**] eliminates all but about 100 of these and the second paper [**50**] reduces the list to 21.

Then Paul Fong [**39**] classified all simple groups whose Sylow 2-subgroups have order at most 2^6 and this reduced Hall's list to 13 integers as possible orders for new simple groups of order less than 1,000,000. Finally, the range problem to 1,000,000 was recently finished when two students of Held,

Bert Beisiegel and Volker Stingl [3], eliminated the remaining integers on Hall's list by extending Fong's work as far as the case 2^{10}.

In conclusion we mention that even though the range problem is not the central one in simple group theory, this achievement is a dramatic illustration of how far the theory has progressed in the 84 years since Hölder determined the simple groups of order up to 200.

The author wishes to thank Professor Roger D. Coleman, Professor Warren J. Wong, and the editors for suggesting numerous changes in the manuscript. I am also grateful to Professor Robert L. Griess, Jr., for sending me information about the sporadic simple groups and some references.

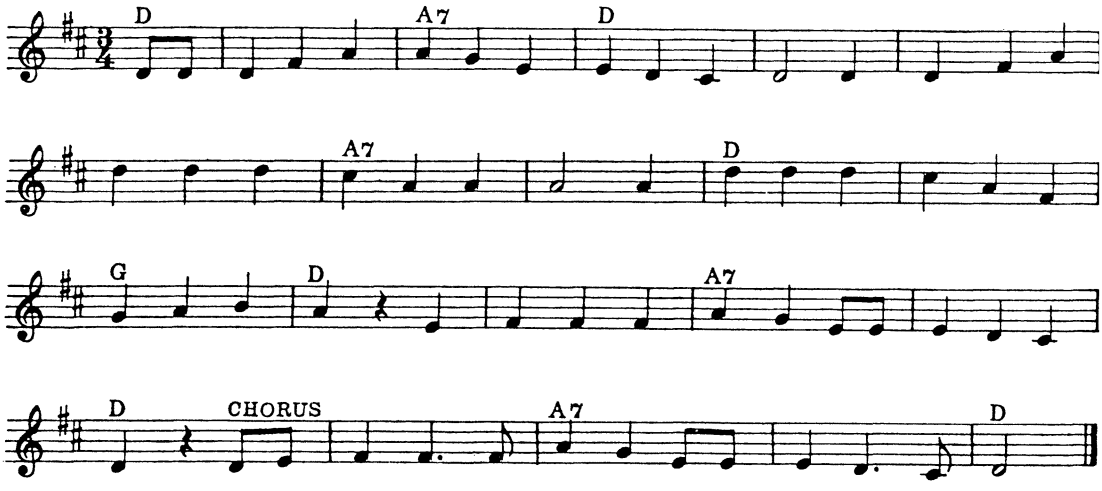
References

[1] E. Artin, The orders of the classical simple groups, Comm. Pure Appl. Math., 8 (1955) 455–472.
[2] ———, Geometric Algebra, Interscience, New York, 1957.
[3] B. Beisiegel and V. Stingl, The finite simple groups with Sylow 2-subgroups of order at most 2^{10}, to appear in Comm. Alg.
[4] E. T. Bell, Fifty years of algebra in America, 1888–1938, in Amer. Math. Soc. Semicentennial Publications, Vol. II, 1–34, New York, 1938.
[5] H. Bender, A group theoretic proof of Burnside's p^aq^b theorem, Math. Z., 126 (1972) 327–338.
[6] N. Biggs, Finite Groups of Automorphisms, Cambridge University Press, Cambridge, 1971.
[7] R. Brauer, On groups whose order contains a prime number to the first power I, Amer. J. Math., 64 (1942) 401–420.
[8] ———, On some conjectures concerning finite simple groups, in Studies in Mathematical Analysis and Related Topics, 56–61, Stanford U. Press, Palo Alto, 1962.
[9] ———, On simple groups of order $5 \cdot 3^a 2^b$, Bull. Amer. Math. Soc., 74 (1968) 900–903.
[10] R. Brauer and K. A. Fowler, Groups of even order, Ann. of Math., 62 (1955) 565–583.
[11] R. Brauer and W. F. Reynolds, On a problem of E. Artin, Ann. of Math., 68 (1958) 713–720.
[12] R. Brauer and H. Tuan, On simple groups of finite order, I, Bull. Amer. Math. Soc., 51 (1945) 756–766.
[13] W. Burnside, On a class of groups defined by congruences, Proc. London Math. Soc., 25 (1894) 113–139.
[14] ———, Notes on the theory of groups of finite order, Proc. London Math. Soc., 26 (1895) 191–214.
[15] ———, Notes on the theory of groups of finite order (continued), Proc. London Math. Soc., 26 (1895) 325–338.
[16] ———, On transitive groups of degree n and class $n-1$, Proc. London Math. Soc., 32 (1900) 240–246.
[17] ———, On some properties of groups of odd order, Proc. London Math. Soc., 33 (1901) 162–185.
[18] ———, On some properties of groups of odd order (second paper), Proc. London Math. Soc., 33 (1901) 257–268.
[19] ———, On groups of order p^aq^β, Proc. London Math. Soc., 2 (1904) 388–392.
[20] ———, Theory of Groups of Finite Order, 2nd ed., Dover, New York, 1955.
[21] R. Carter, Simple Groups of Lie Type, Wiley, London, 1972.
[22] C. Chevalley, Sur certains groupes simples, Tohoku Math. J., 7 (1955) 14–66.
[23] F. Cole, Simple groups from order 201 to order 500, Amer. J. Math., 14 (1892) 378–388.
[24] ———, Simple groups as far as order 660, Amer. J. Math., 15 (1893) 303–315.
[25] ———, List of the transitive substitution groups of ten and of eleven letters, Quart. J. Pure Appl. Math., 27 (1895) 39–50.
[26] ———, On simple groups of low order, Bull. Amer. Math. Soc., 30 (1924) 489–492.
[27] G. Cornell, N. Pelc, M. Wage, Simple groups of order less than 1000, J. of Undergraduate Mathematics, 5 (1973) 77–86.
[28] C. W. Curtis, The classical groups as a source of algebraic problems, Amer. Math. Monthly, 74 (1967) 80–91.
[29] L. E. Dickson, The analytic representation of substitutions on a power of a prime number of letters with a discussion of the linear group, Ann. of Math., 11 (1897) 161–183.
[30] ———, Proof of the non-isomorphism of the simple Abelian group on $2m$ indices and the orthogonal group on $2m+1$ indices for $m>2$, Quart. J. Pure Appl. Math., 32 (1900) 42–63.
[31] ———, Theory of linear groups in an arbitrary field, Trans. Amer. Math. Soc., 2 (1901) 363–394.
[32] ———, A new system of simple groups, Math. Ann., 60 (1905) 137–150.
[33] ———, Linear Groups with an Exposition of the Galois Field Theory, Dover, New York, 1958.
[34] J. Dieudonné, Sur les Groupes Classiques, Actualités Sci. Indust., No. 1040, Hermann, Paris, 1948.
[35] ———, On the automorphisms of the classical groups, Mem. Amer. Math. Soc., 2 (1951).
[36] W. Feit, The current situations in the theory of finite simple groups, Proc. Internat. Congr. Mathematicians (Nice, 1970), Gauthier-Villars, Paris, 1971.

[37] W. Feit, M. Hall, Jr., J. G. Thompson, Finite groups in which the centralizer of any non-identity element is nilpotent, Math. Z., 74 (1960) 1–17.
[38] W. Feit and J. G. Thompson, Solvability of groups of odd order, Pacific J. Math., 13 (1963) 775–1029.
[39] P. Fong, private communication.
[40] G. Frobenius, Über Gruppen des Grades p oder $p+1$, Berliner Sitzgsb., (1902) 351–369.
[41] E. Galois, Sur la théorie des nombres, J. Math. Pures Appl., 11 (1846) 398–407.
[42] G. Glauberman, Subgroups of finite groups, Bull. Amer. Math. Soc., 73 (1967) 1–12.
[43] D. Goldschmidt, A group theoretic proof of the p^aq^b theorem for odd primes, Math. Z., 113 (1970) 373–375.
[44] ———, Elements of order two in finite groups, Delta, 4 (1974) 45–58.
[45] D. Gorenstein, Some topics in the theory of finite groups, Rend. Mat. e Appl., (5) 23 (1964) 298–315.
[46] ———, Finite Groups, Harper and Row, New York, 1968.
[47] ———, Finite simple groups and their classification, Israel J. Math., 19 (1974) 5–66.
[48] D. Gorenstein and J. Walter, The characterization of finite groups with dihedral Sylow 2-subgroups, J. Algebra, 2 (1965) 85–151, 218–270, 354–393.
[49] M. Hall, A search for simple groups of order less than one million, Computational Problems in Abstract Algebra (John Leech, Ed.), 137–168, Pergamon Press, New York, 1969.
[50] ———, Simple groups of order less than one million, J. Algebra, 20 (1972) 98–102.
[51] T. Hawkins, The origins of the theory of group characters, Archive Hist. Exact Sci., 7 (1971) 142–170.
[52] ———, New light on Frobenius' creation of the theory of group characters, Archive Hist. Exact Sci., 12 (1974) 215–243.
[53] D. Held, A characterization of the alternating groups of degrees eight and nine, J. Algebra, 7 (1967) 218–237.
[54] ———, The simple group related to M_{24}, J. Algebra, 13 (1969), 253–296.
[55] I. N. Herstein, Topics in Algebra, 2nd ed., Xerox College, Lexington, Mass., 1975.
[56] O. Hölder, Die einfachen Gruppen im ersten und zweiten Hundert der Ordnungszahlen. Math. Ann., 40 (1892) 55–88.
[57] Z. Janko, A new finite simple group with Abelian Sylow 2-subgroups and its characterization, J. Algebra, 3 (1966) 147–186.
[58] C. Jordan, Traité des Substitutions, Gauthier-Villars, Paris, 1870.
[59] R. Lindberg and J. Robinson, A project in simple group theory, Senior paper, U. of Minn., Duluth, 1973.
[60] G. Ling and G. A. Miller, Proof that there is no simple group whose order lies between 1092 and 2001, Amer. J. Math., 22 (1900) 13–26.
[61] H. Matsuyama, Solvability of groups of order 2^ap^b, Osaka J. Math., 10 (1973) 375–378.
[62] E. Michaels, A study of simple groups of even order, Ph.D. dissertation, U. of Notre Dame, 1963.
[63] G. A. Miller, On the simple groups which can be represented as substitution groups that contain cyclical substitutions of a prime degree, Amer. Math. Monthly, 6 (1899) 102–103.
[64] ———, Sur plusieurs groupes simples, Bull. Soc. Math. France, 28 (1900) 266–267.
[65] ———, The simple group of order 2520, Bull. Amer. Math. Soc., 22 (1922) 98–102.
[66] E. H. Moore, A doubly-infinite system of simple groups, Bull. New York Math. Soc., 3 (1893) 73–78.
[67] M. B. Powell and G. Higman, editors, Finite Simple Groups, Academic Press, London, 1971.
[68] R. Ree, A family of simple groups associated with the simple Lie algebra of type (F_4), Amer. J. Math., 83 (1961) 401–420.
[69] ———, A family of simple groups associated with the simple Lie algebra of type (G_2), Amer. J. Math., 83 (1961) 432–462.
[70] I. M. Schottenfels, Two non-isomorphic simple groups of the same order 20,160, Ann. of Math., 2nd series, 1 (1900) 147–152.
[71] L. Siceloff, Simple groups from order 2001 to order 3640, Amer. J. Math., 34 (1912) 361–372.
[72] M. Suzuki, The nonexistence of a certain type of simple groups of odd order, Proc. Amer. Math. Soc., 8 (1957) 686–695.
[73] ———, A new type of simple groups of finite order, Proc. Nat. Acad. Sci. U.S.A., 46 (1960) 868–870.
[74] P. Telega, A computer project in simple group theory, Senior paper, U. of Minn., Duluth, 1975.
[75] J. G. Thompson, Nonsolvable finite groups all of whose local subgroups are solvable, Bull. Amer. Math. Soc., 74 (1968) 383–437; Pacific J. Math., 33 (1970) 451–537; Pacific J. Math., 39 (1971) 483–534; Pacific J. Math., 48 (1973) 511–592; Pacific J. Math., 50 (1974) 215–297; Pacific J. Math., 51 (1974) 573–630.
[76] D. Wales, Simple groups of order $13\cdot 3^a 2^b$, J. Algebra, 20 (1972) 124–143.
[77] J. H. Walter, The characterization of finite groups with abelian Sylow 2-subgroups, Ann. of Math., 89 (1969) 405–514.
[78] H. Wielandt, Finite Permutation Groups, Academic Press, New York, 1964.
[79] W. J. Wong, Recent work on finite simple groups, Math. Chronicle, 1 (1969) 5–12.

A Simple Song

For readers who have difficulty in keeping track of the history of simple group theory we offer the following summary in ballad form. This classic first appeared in print in 1973 in the *Amer. Math. Monthly* (p. 1028) where it is claimed to have been found scrawled on a table in Eckhart Library at the University of Chicago. The author, possibly for cause, is claimed to be unknown. The tune is that of "Sweet Betsy from Pike."



Oh, what are the orders of all simple groups?
I speak of the honest ones, not of the loops.
It seems that old Burnside their orders has guessed
Except for the cyclic ones, even the rest.

CHORUS: Finding all groups that are simple is no simple task.

Groups made up with permutes will produce some more:
For A_n is simple, if n exceeds 4.
Then, there was Sir Matthew who came into view
Exhibiting groups of an order quite new.

Still others have come on to study this thing.
Of Artin and Chevalley now we shall sing.
With matrices finite they made quite a list
The question is: Could there be others they've missed?

Suzuki and Ree then maintained it's the case
That these methods had not reached the end of the chase.
They wrote down some matrices, just four by four,
That made up a simple group. Why not make more?

And then came the opus of Thompson and Feit
Which shed on the problem remarkable light.
A group, when the order won't factor by two
Is cyclic or solvable. That's what is true.

Suzuki and Ree had caused eyebrows to raise,
But the theoreticians they just couldn't faze.
Their groups were not new: if you added a twist,
You could get them from old ones with a flick of the wrist.

Still, some hardy souls felt a thorn in their side.
For the five groups of Mathieu all reason defied;
Not A_n, not twisted, and not Chevalley,
They called them sporadic and filed them away.

Are Mathieu groups creatures of heaven or hell?
Zvonimir Janko determined to tell.
He found out that nobody wanted to know:
The masters had missed 1 7 5 5 6 0.

The floodgates were opened! New groups were the rage!
(And twelve or more sprouted, to greet the new age.)
By Janko and Conway and Fischer and Held
McLaughlin, Suzuki, and Higman, and Sims.

No doubt you noted the last lines don't rhyme.
Well, that is, quite simply, a sign of the time.
There's chaos, not order, among simple groups;
And maybe we'd better go back to the loops.

Counting by Correspondence

Insight into enumeration may follow from correspondences between sets to be counted and sets whose cardinalities are obvious.

ROMAE J. CORMIER
ROGER B. EGGLETON
Northern Illinois University

1. Introduction

How should mathematical objects be counted? Several quite distinct methods are popular — including summation of finite series, use of one-to-one correspondences, manipulation of generating functions, and application of the inclusion-exclusion principle, to name only the most common. Perhaps summation of finite series is the method used most widely, even though it tends to be a brute force approach relying on well-developed techniques of manipulation rather than on features which are inherent in the context of the particular problem. Its use reflects the viewpoint that the answer is of primary importance, while the way it is obtained is incidental: in other words, the end justifies the means. We are sure many besides ourselves have had the experience of solving a counting problem by summation of finite series and upon finding that the answer was relatively simple, have thought: "There *must* be a more natural way to derive this solution!" Our intention here is to show that the use of correspondences may provide more natural and more enlightening solutions to counting problems, by making use of inherent structural features.

As a paradigm, we take the problem of counting all nondecreasing sequences of k natural numbers, none of which exceeds n: that is, the problem of determining the cardinality of the set

$$S(n, k) = \{(a_1, a_2, \ldots, a_k) \in \mathbb{N}^k : 0 \leqq a_1 \leqq a_2 \leqq \cdots \leqq a_k \leqq n\},$$

where $\mathbb{N}$ is the set of natural numbers (including zero). The finite series method leads to

$$|S(n, k)| = \sum_{a_1=0}^{n} \sum_{a_2=a_1}^{n} \cdots \sum_{a_i=a_{i-1}}^{n} \cdots \sum_{a_k=a_{k-1}}^{n} 1,$$

which actually corresponds to counting by first imposing a particular ordering on $S(n, k)$, in this case the lexicographic ordering defined by

$$(a_1, a_2, \ldots, a_k) < (a_1', a_2', \ldots, a_k') \leftrightarrow \exists\, r\colon a_r < a_r' \;\&\; \forall\, i < r\colon a_i = a_i'.$$

The series first counts those sequences which differ only in the last term, then those which differ only in the last two terms, and so on. The summation is straightforward though not immediate. (One way the evaluation could be effected is to note that the series implies the identity $|S(n, k)| = \Sigma_{a=0}^{n} |S(n-a, k-1)| = \Sigma_{a=0}^{n} |S(a, k-1)|$. But this procedure is actually passing from summation of finite series to use of generating functions.)

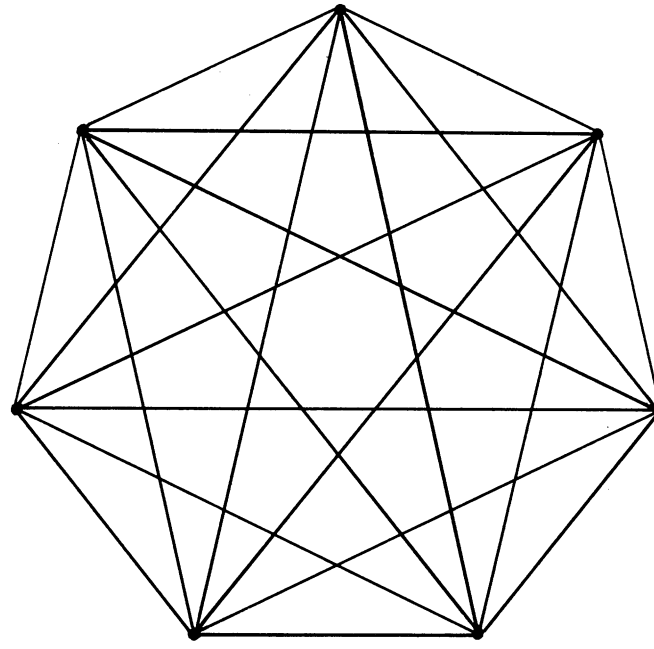
A Convex

Nondegenerate

Complete

Polygon

with seven sides contains a large number of triangles whose vertices are points of intersection of lines of the figure and whose sides are segments of these lines. In this figure there are 287 such triangles of which 35 have only original vertices, 140 have one crossing point in the interior as a vertex, 105 have two interior vertices, and only 7 are totally interior.



In contrast, the standard method of solving the problem is to note the one-to-one correspondence

$$(a_1, a_2, \ldots, a_i, \ldots, a_k) \leftrightarrow \{a_1, a_2 + 1, \ldots, a_i + i - 1, \ldots, a_k + k - 1\}$$

between the sequences in $S(n, k)$ and the k-subsets of $I(n + k) = \{i \in \mathbb{N}: i < n + k\}$, the initial segment of the natural numbers comprising all those smaller than $n + k$. The correspondence shows a way to interpret the nondecreasing condition on any sequence in $S(n, k)$ as ensuring that a related sequence is strictly increasing, so its terms are all different and it may be specified simply as a set, retrieval of the order being trivial. Let $I(n + k, k)$ denote the set of k-subsets of $I(n + k)$. Then we have the one-to-one correspondence $S(n, k) \leftrightarrow I(n + k, k)$, so immediately

$$|S(n, k)| = |I(n + k, k)| = \binom{n + k}{k}.$$

What determines whether a correspondence is good? The main requirement is that it should establish a relationship between the relatively unfamiliar set under study and a set which is more familiar and has a well-known structure. This is clearly the underlying reason for success in the above example. Another criterion for a good correspondence is that the structure of the given set should be made more transparent by reference to a familiar structure, any special peculiarities in the original structure being adequately reflected in the reference structure.

We shall now proceed to solve two triangle-counting problems by the correspondence method. (The corresponding finite series will not be presented.) As there will be need to use correspondences which are not necessarily one-to-one, we shall use the notation $A \overset{r}{\underset{s}{\leftrightarrow}} B$ to denote an r-to-s correspondence from the set A to the set B. If $r = 1$ or $s = 1$, we shall, for convenience, adopt the convention that such unitary values need not be explicitly specified.

2. Triangles in a complete polygon

A **complete n-gon** in the euclidean plane is a configuration which comprises n **vertices** (points), each pair joined by an **edge** (straight line segment). Let K_n be a complete n-gon which is **convex** (that is, the vertices determine a convex n-gon) and **nondegenerate** (that is, any point common to three

edges must be a vertex). Let $V(n)$ denote the set of vertices of K_n. A **triangle** in this configuration is any K_3 subconfiguration. Thus the vertices of such a triangle, which must not be collinear, are either vertices of K_n or crossing points of edges of K_n, and the edges of the triangle are contained in edges of K_n. Let $T(n)$ be the set of all triangles in K_n. Our goal is to determine $|T(n)|$.

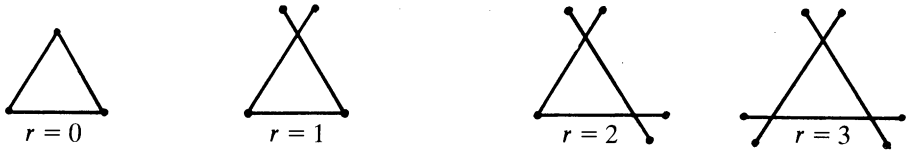
We begin by looking at the problem of counting the crossing points of K_n. This is one of Chrystal's combinatorial problems [1]. More recently it was posed as an elementary problem by Erdös [3], and two solutions were published. One, by Kaufman and Koch [7], uses correspondences; the other, by Rosenthal [9], uses summation of finite series. These two solutions provide a nice illustration of our theme. For completeness we shall present the correspondence argument.

Let $C(n)$ denote the set of crossing points in K_n. Each crossing point belongs to a unique pair of edges, which in turn determines a unique set of four vertices; conversely, each set of four vertices determines a K_4 subconfiguration, which contains just one crossing point. Thus $C(n) \leftrightarrow V(n, 4)$, where $V(n, k)$ denotes the set of k-subsets of $V(n)$. Hence

$$|C(n)| = |V(n, 4)| = \binom{n}{4}.$$

The success of this one-to-one correspondence suggests that subsets of $V(n)$ may be appropriate for counting $T(n)$. We shall say that a triangle of K_n is of **type** r (see FIGURE 1) if exactly r of its vertices are crossing points of K_n, and let $T(n, r)$ denote the set of type r triangles in K_n. Each triangle of type 0 obviously determines a unique set of three vertices of K_n, and conversely, so $T(n, 0) \leftrightarrow V(n, 3)$. Hence

$$|T(n, 0)| = |V(n, 3)| = \binom{n}{3}.$$



The four possible types of triangle in K_n

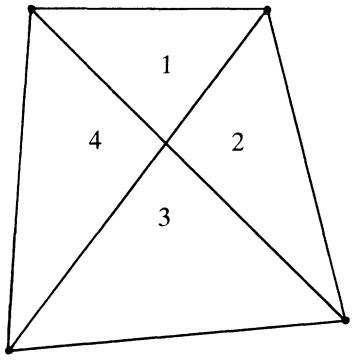
FIGURE 1.

Each triangle of type 1 determines a unique set of four vertices of K_n, those associated with the vertex of the triangle which is a crossing point of K_n. Each set of four vertices of K_n determines a K_4 subconfiguration, which contains four triangles of type 1. (See FIGURE 2.) Therefore there is a four-to-one correspondence $T(n, 1) \overset{4}{\leftrightarrow} V(n, 4)$, whence

$$|T(n, 1)| = 4|V(n, 4)| = 4\binom{n}{4}.$$

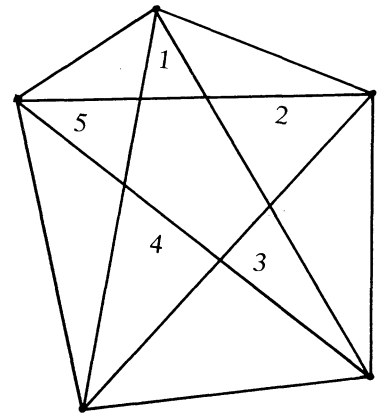
Each triangle of type 2 determines a unique set of five vertices of K_n, those associated with its two vertices which are crossing points of K_n (each has four associated vertices, and the two sets of four have three vertices in common). Each set of five vertices of K_n determines a K_5 subconfiguration, which contains five triangles of type 2. (See FIGURE 3.) Therefore we have a five-to-one correspondence $T(n, 2) \overset{5}{\leftrightarrow} V(n, 5)$, so

$$|T(n, 2)| = 5|V(n, 5)| = 5\binom{n}{5}.$$



K_4 contains 4 type 1 triangles

FIGURE 2.



K_5 contains 5 type 2 triangles

FIGURE 3.

Finally, each triangle of type 3 determines a unique set of six vertices of K_n, and each such set of six vertices determines a K_6 subconfiguration, which contains a unique triangle of type 3. Hence $T(n,3) \leftrightarrow V(n,6)$, and therefore

$$|T(n,3)| = |V(n,6)| = \binom{n}{6}.$$

Since $T(n) = \bigcup_{r=0}^{3} T(n,r)$ and the union is disjoint, we have

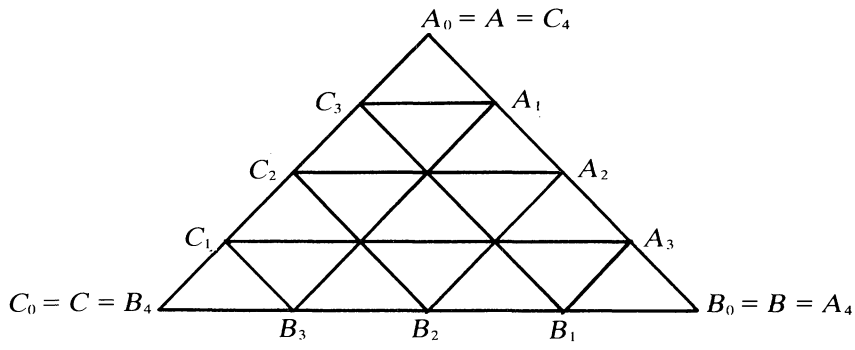
$$|T(n)| = \binom{n}{3} + 4\binom{n}{4} + 5\binom{n}{5} + \binom{n}{6}.$$

This clearly shows that type 3 triangles dominate when n is sufficiently large, and that $|T(n)| \sim (1/720)n^6$. In fact, type 0 dominates when $n = 3$, type 1 dominates when $5 \leqq n \leqq 7$, type 2 dominates when $9 \leqq n \leqq 34$, and type 3 dominates when $n \geqq 36$. (When $n = 4$, types 0 and 1 codominate; when $n = 8$, types 1 and 2 codominate; and when $n = 35$, types 2 and 3 codominate.) When $n \geqq 39$, more than 50% of the triangles are of type 3; when $n \geqq 279$, more than 90% are of type 3; and when $n \geqq 2979$ type 3 triangles account for more than 99% of the total.

Let us call a complete n-gon **regular** if its vertices determine a regular n-gon. We note that it was conjectured by Patten [**8**] and also by Steinhaus [**10**] that a regular complete n-gon is nondegenerate just if n is odd. This was proved by Heineken [**5**]. Therefore our results apply to regular complete n-gons with n odd. Independently Harborth [**4**] and Heineken [**6**] have shown that if n is even but not divisible by 3, a regular complete n-gon has no more than three edges passing through any crossing point other than the central point.

3. Triangles in a triangulated triangle

An nth-order **triangulated triangle** T_n (see FIGURE 4) may be constructed as follows. Let A, B, C be noncollinear points in the euclidean plane. Let A_i $(0 < i < n)$ be points which divide AB into n equal segments with $A_0 = A$, $A_n = B$. Similarly let B_i $(0 < i < n)$ divide BC into n equal segments with $B_0 = B$, $B_n = C$ and let C_i $(0 < i < n)$ divide CA into n equal segments with $C_0 = C$, $C_n = A$. The configuration T_n comprises the straight line segments A_iB_{n-i}, B_iC_{n-i} and C_iA_{n-i} where $i \in I(n)$, the initial segment of the natural numbers comprising all those smaller than n. A **triangle** in this



A fourth order triangulated triangle T_4

FIGURE 4.

configuration is any T_1 subconfiguration. Let $\Delta(n)$ be the set of all triangles in T_n. Edwards [2] has posed the problem of determining $|\Delta(n)|$. In keeping with the spirit of this paper, we shall now obtain $|\Delta(n)|$ via several correspondences.

Each triangle in $\Delta(n)$ has its sides parallel to those of the triangle ABC. If a particular triangle has its sides contained in the segments A_iB_{n-i}, B_jC_{n-j} and C_kA_{n-k}, we have a correspondence between that triangle and the ordered triple (i, j, k), which belongs to the cartesian product $I(n)^3$. However, not all ordered triples in $I(n)^3$ correspond to triangles in $\Delta(n)$. The condition that A_iB_{n-i} and B_jC_{n-j} intersect is equivalent to the condition that $i + j \leqq n$. Similarly we require $j + k \leqq n$ and $k + i \leqq n$. Also the condition that A_iB_{n-i}, B_jC_{n-j} and C_kA_{n-k} be concurrent is equivalent to the condition that $i + j + k = n$. Therefore

$$\Delta(n) \leftrightarrow \Sigma(n, 3) = \{(i, j, k) \in I(n)^3 : i + j \leqq n,\ j + k \leqq n,\ k + i \leqq n;\ i + j + k \neq n\}.$$

We partition $\Sigma(n, 3)$ into two classes, denoted by subscripts 0 and 1, according as $i + j + k$ is respectively less than or greater than n. The triples in $\Sigma_0(n, 3)$ all correspond to triangles in the same attitude as ABC, that is, with base nearest BC; the triples in $\Sigma_1(n, 3)$ all correspond to triangles in the inverted attitude to ABC, that is, with apex nearest BC. If $(i, j, k) \in \Sigma_0(n, 3)$, then $(i, i + j, i + j + k)$ is a nondecreasing sequence with no term exceeding $n - 1$, so in the notation of our paradigm, $\Sigma_0(n, 3) \leftrightarrow S(n - 1, 3)$. Since we know $S(n - 1, 3) \leftrightarrow I(n + 2, 3)$, we have

$$|\Sigma_0(n, 3)| = |I(n + 2, 3)| = \binom{n + 2}{3}.$$

Again, if $(i, j, k) \in \Sigma_1(n, 3)$ then $(i + j + k - n, i, i + j)$ is a nondecreasing sequence with no term exceeding n, and no term less than 1. By subtracting 1 from each term, we get a sequence in $S(n - 1, 3)$. However, not all sequences in $S(n - 1, 3)$ correspond in this way to members of $\Sigma_1(n, 3)$. The condition $k + i \leqq n$ must hold for (i, j, k), so if (a, b, c) is the corresponding sequence in $S(n - 1, 3)$ the appropriate condition is $a + b < c$. Thus

$$\Sigma_1(n, 3) \leftrightarrow \Lambda(n, 3) = \{(a, b, c) \in S(n - 1, 3) : a + b < c\}.$$

With any $(a, b, c) \in \Lambda(n, 3)$ we shall associate the ordered triples $(a, a + b + 1, c + 1)$ and $(b, a + b + 1, c + 1)$, each of which is a strictly increasing sequence of three natural numbers, none of which exceeds n. By suppressing order, each corresponds to a 3-set in $I(n + 1, 3)$. Conversely, any 3-set in $I(n + 1, 3)$ corresponds to just one member of $\Lambda(n, 3)$. Thus $(a, b, c) \in \Lambda(n, 3)$ corresponds to two members of $I(n + 1, 3)$ if $a < b$, and one member if $a = b$. (These conditions correspond to

whether or not the associated triangle has a vertex on AC.) Let $\Omega(n,2) = \{(2a, c): (a, a, c) \in \Lambda(n,3)\}$. Then for any $(a, b, c) \in \Lambda(n,3)$,

$$(a,b,c) \underset{2}{\leftrightarrow} \begin{cases} (a, a+b+1, c+1), (b, a+b+1, c+1) & \text{if} \quad a < b, \\ (a, 2a+1, c+1), (2a, c) & \text{if} \quad a = b, \end{cases}$$

defines a one-to-two correspondence $\Lambda(n,3) \underset{2}{\leftrightarrow} I(n+1,3) \cup \Omega(n,2)$, where the union is disjoint.

Similarly, for any $(2a, c) \in \Omega(n,2)$,

$$(2a,c) \underset{2}{\leftrightarrow} \begin{cases} (2a, c), (2a+1, c) & \text{if} \quad 2a+1 < c \\ (2a, 2a+1), a & \text{if} \quad 2a+1 = c \end{cases}$$

defines a one-to-two correspondence $\Omega(n,2) \underset{2}{\leftrightarrow} I(n,2) \cup I\left(\left[\frac{n}{2}\right]\right)$ where the union is disjoint. (The two cases in the correspondence depend on whether or not the associated triangle has a vertex on AB.)

We have now carried through the correspondences to sets which can be counted immediately, yielding

$$2|\Omega(n,2)| = |I(n,2)| + \left|I\left(\left[\frac{n}{2}\right]\right)\right| = \binom{n}{2} + \left[\frac{n}{2}\right],$$

$$2|\Lambda(n,3)| = |I(n+1,3)| + |\Omega(n,2)| = \binom{n+1}{3} + |\Omega(n,2)|,$$

$$|\Sigma_1(n,3)| = |\Lambda(n,3)|.$$

Because of the disjoint union $\Sigma_0(n,3) \cup \Sigma_1(n,3) = \Sigma(n,3)$, the original problem can now be resolved:

$$|\Delta(n)| = |\Sigma(n,3)| = |\Sigma_0(n,3)| + |\Sigma_1(n,3)|,$$

so

$$|\Delta(n)| = \binom{n+2}{3} + \frac{1}{2}\binom{n+1}{3} + \frac{1}{4}\binom{n}{2} + \frac{1}{4}\left[\frac{n}{2}\right].$$

From these results we also get two bonus facts, interesting in their own right: $\binom{n}{2}$ has the same parity as $\left[\frac{n}{2}\right]$, and $\binom{n+1}{3}$ has the same parity as $\frac{1}{2}\left\{\binom{n}{2} + \left[\frac{n}{2}\right]\right\}$.

The research of the second author was supported by the Foundation for Number Theory Computing.

References

[1] G. Chrystal, Problem 7, Exercise Set IV in Vol. 2 of Algebra: An Elementary Textbook, Black, Edinburgh, 1889.

[2] R. E. Edwards, Problem 889, this MAGAZINE, 47 (1974) 46–47.

[3] P. Erdös, Elementary Problem E750, Amer. Math. Monthly, 53 (1946) 591.

[4] Heiko Harborth, Diagonalen im regulären n-Eck, Elem. Math., 24 (1969) 104–109.

[5] Hermann Heineken, Regelmäßige Vielecke und ihre Diagonalen, Enseignement Math., Ser. II, 8 (1962) 275–278.

[6] ———, Regelmäßige Vielecke und ihre Diagonalen II, Rend. Sem. Mat. Univ. Padova, 41 (1968) 332–344.

[7] Norbert Kaufman and R. H. Koch, Solution to Elementary Problem E750, Amer. Math. Monthly, 54 (1947) 344.

[8] W. E. Patten, cited in editorial remarks on Elementary Problem E750, Amer. Math. Monthly, 54 (1947) 345.

[9] Arthur Rosenthal, Solution to Elementary Problem E750, Amer. Math. Monthly, 54 (1947) 344–345.

[10] H. Steinhaus, Problem 225, Colloq. Math., 5 (1958) 235.

The "Sales Tax" Theorem

SOLOMON W. GOLOMB
University of Southern California

Sometimes in mathematics there are several results which were obtained independently, and at first sight appear unrelated. Then someone finds a more general viewpoint that makes the earlier discoveries special cases of a single new theorem. In the process, it usually becomes clearer what was "really going on" in the original results.

One such example in the history of mathematics is the underlying similarity between the Chinese remainder theorem and the Lagrange interpolation formula. To state them similarly, the Chinese remainder theorem says: *given integers $m_1, m_2, \ldots, m_k$ which are "independent" in the sense that no two have a common prime factor, and given integer values $c_1, c_2, \ldots, c_k$, we can find an integer N which "takes the value c_i at m_i" for each $i = 1, 2, \ldots, k$, in the sense that $N \equiv c_i \pmod{m_i}$; and N is unique* modulo $M = m_1 m_2 \cdots m_k$. Correspondingly, the Lagrange interpolation theorem says: *given real numbers $x_1, x_2, \ldots, x_k$ which are "independent" (i.e., distinct points) on the real axis, and given real values $y_1, y_2, \ldots, y_k$, we can find a polynomial function $f(x)$ which "takes the value y_i at x_i" for each $i = 1, 2, \ldots, k$, in the sense that $f(x_i) = y_i$; and this $f(x)$ is unique for degree less than k.* The general result of which these are both special cases is called the "Theorem on Independence of Places" in algebraic geometry, and requires some specialized knowledge of "valuation theory" to state properly.

In this paper, we shall look at a similar situation, but in this case the general result will be easier to follow than the historical special cases which preceded it. One of these special cases is the following mysterious result in number theory, illustrated in TABLE 1:

THEOREM A. *Let p_n be the nth prime number, and let $\pi(n)$ be the number of primes not exceeding n. Then every positive integer occurs once and only once in either the sequence $\{n + \pi(n)\}$ or the sequence $\{p_n + n - 1\}$ and no integer occurs in both sequences.*

It turns out that this result has virtually nothing to do with the sequence of prime numbers! The general formualtion is:

THEOREM B. *Let $Q = \{q_n\}$ be any subsequence of the positive integers (thus $1 \leqq q_1 < q_2 < q_3 < \cdots$), and let $\tau(n)$ be the number of terms of the sequence Q which do not exceed n. (That is, if $q_k \leqq n$ while $q_{k+1} > n$, then $\tau(n) = k$.) Then the positive integers are exactly partitioned into the two non-overlapping sequences $\{n + \tau(n)\}$ and $\{q_n + n - 1\}$.*

Our main objective in this paper is to give an especially transparent proof of Theorem B, based on interpreting $\tau(n)$ as the "sales tax" to be paid on an item having a price of n. We then look at various applications of Theorem B, using the fact that q_n and $\tau(n)$ are "inverse functions", in view of the obvious identity $\tau(q_n) = n$.

To prove Theorem B, we regard the sequence $\{q_n\}$ as a "tax table", in the sense that the sales tax increases by one cent at every term of the sequence q_n (and at no other values). Thus the sales tax on the price q_k is exactly k. More generally, the sales tax on the price m, denoted $\tau(m)$, is the number of terms of $\{q_n\}$ not exceeding m.

n	p_n	$\pi(n)$	$n+\pi(n)$	p_n+n-1
1	2	0	1	2
2	3	1	3	4
3	5	2	5	7
4	7	2	6	10
5	11	3	8	15
6	13	3	9	18
7	17	4	11	23
8	19	4	12	26
9	23	4	13	31
10	29	4	14	38
11	31	5	16	41
12	37	5	17	48
13	41	6	19	53
14	43	6	20	56
15	47	6	21	61
16	53	6	22	68
17	59	7	24	75
18	61	7	25	78
19	67	8	27	85
20	71	8	28	90

The partition of the integers into $\{n+\pi(n)\}$ and $\{p_n+n-1\}$.

TABLE 1

n	$\sqrt{2}\,n$	$(2+\sqrt{2})n$	$\lfloor\sqrt{2}\,n\rfloor$	$\lfloor(2+\sqrt{2})n\rfloor$
1	1.4142	3.4142	1	3
2	2.8284	6.8284	2	6
3	4.2426	10.2426	4	10
4	5.6569	13.6569	5	13
5	7.0711	17.0711	7	17
6	8.4853	20.4853	8	20
7	9.8995	23.8995	9	23
8	11.3137	27.3137	11	27
9	12.7279	30.7279	12	30
10	14.1421	34.1421	14	34
11	15.5563	37.5563	15	37
12	16.9706	40.9706	16	40
13	18.3848	44.3848	18	44
14	19.7990	47.7990	19	47
15	21.2132	51.2132	21	51
16	22.6274	54.6274	22	54
17	24.0416	58.0416	24	58
18	25.4558	61.4558	25	61
19	26.8701	64.8701	26	64
20	28.2843	68.2843	28	68

The partition of the integers into $\{\lfloor\sqrt{2}n\rfloor\}$ and $\{\lfloor(2+\sqrt{2})n\rfloor\}$.

TABLE 2

From this point of view, the "total price" (including tax) on an item with a net price of n is $n+\tau(n)$. The sequence $\{n+\tau(n)\}$ thus consists of all numbers which can occur as "total prices". What numbers cannot occur as "total prices"? As the net price increases through one of the terms of $\{q_n\}$, say from q_n-1 to q_n, the total price increases from $(q_n-1)+(n-1)$ to q_n+n, thus skipping the value q_n+n-1. If m is not of the form q_n, then the total price goes from $(m-1)+\tau(m-1)$ to $m+\tau(m)$, increasing by only one cent, because in this case $\tau(m-1)=\tau(m)$. Thus the integers skipped in the sequence $\{n+\tau(n)\}$ are precisely the terms of the sequence $\{q_n+n-1\}$. This proves Theorem B.

We now turn to several applications of this "sales tax" theorem. Following Knuth [1], we use the notation $\lfloor x\rfloor$ to denote the largest integer $\leqq x$, and the notation $\lceil x\rceil$ to denote the smallest integer $\geqq x$. For real $\alpha>1$, the function $f(n)=\lfloor\alpha n\rfloor$ is one-to-one from the positive integers into the positive integers. The inverse function of $f(n)$ is given as follows: If $\lfloor\alpha n\rfloor=m$, then $n=\lceil m/\alpha\rceil$. (If $\lfloor\alpha n\rfloor=m$, then $\alpha n=m+\theta$ with $0\leqq\theta<1$. Hence $n=m/\alpha+\theta/\alpha=\lceil m/\alpha\rceil$, since $0<\theta/\alpha<1$ because $0\leqq\theta<1<\alpha$.) Similarly, we also have: If $\lceil\alpha n\rceil=m$, then $n=\lfloor m/\alpha\rfloor$.

We now apply Theorem B to this last result, letting $q_n=\lceil\alpha n\rceil$ for fixed $\alpha>1$, in which case $\tau(n)=\lfloor n/\alpha\rfloor$. Then every integer is either of the form $n+\lfloor n/\alpha\rfloor$ or of the form $\{\lceil\alpha n\rceil+n-1\}$. Clearly $n+\lfloor n/\alpha\rfloor=\lfloor(1+(1/\alpha))n\rfloor$, and $\lceil\alpha n\rceil+n-1=\lceil(\alpha+1)n\rceil-1=\lfloor(\alpha+1)n\rfloor$, where we introduce the symbol $\lfloor x\rfloor$ to denote the largest integer strictly less than x. (For example, $\lfloor 5\rfloor=4$.) By setting $u=1+1/\alpha$ and $v=u/(u-1)=\alpha+1$, we obtain:

THEOREM C. *Let u be a real number, $u>1$, and let $v=u/(u-1)$. Then every integer is either of the form $\lfloor un\rfloor$ or of the form $\lfloor vn\rfloor$, but not both.*

Note the symmetry between u and v: just as $v=u/(u-1)$, so too $u=v/(v-1)$. This relation can also be written as $(1/u)+(1/v)=1$. Note also that the restriction $u>1$ is essential, because if $0<u<1$ then $\lfloor un\rfloor$ will assume certain integer values more than once; while if $u=1$ then v is

undefined. When $1 < u < 2$, then $2 < v < \infty$, while when $u = 2$ then $v = 2$. In this last case $\{\lfloor un \rfloor\} = \{2n\}$ and $\{\lfloor vn \rfloor\} = \{2n - 1\}$, giving the ancient partition of the integers into even and odd. Here is another special case of Theorem C which appeared in [2].

THEOREM D. *Let $\lambda > 2$ be irrational, and let $\mu = \lambda/(\lambda - 1)$. Then every positive integer appears in one of the two non-overlapping sequences $\{\lfloor \lambda n \rfloor\}$ and $\{\lfloor \mu n \rfloor\}$.*

Our reasoning is that since λ is irrational, μ must also be irrational. Hence $\lfloor \mu n \rfloor$ is never an integer, and $\lfloor \mu n \rfloor = \lfloor \mu n \rfloor$ for all n. Thus Theorem D follows from Theorem C. For example, when $\lambda = \sqrt{2} = 1.41421\ldots$ we have $\mu = \sqrt{2}/(\sqrt{2} - 1) = 2 + \sqrt{2} = 3.41421\ldots$, and every positive integer occurs exactly once between the sequences $\{\lfloor \sqrt{2} n \rfloor\}$ and $\{\lfloor (2 + \sqrt{2}) n \rfloor\}$ (see TABLE 2).

Here are two more applications. Let $m > 1$ be a fixed positive integer, and let the sequence $\{q_n\}$ of Theorem B be given by $q_n = n^m$, $n = 1, 2, 3, \ldots$. Then every positive integer occurs exactly once between the two sequences $\{n^m + n - 1\}$ and $\{\lfloor n^{1/m} \rfloor + n\}$. (By Theorem B, it suffices to observe that $\tau(n)$, the number of perfect mth powers not exceeding n, is given by $\lfloor n^{1/m} \rfloor$.) Similarly, if $b > 1$ is a fixed positive integer, then every positive integer occurs exactly once between the two sequences $\{b^n + n - 1\}$ and $\{n + \lfloor \log_b n \rfloor\}$. These last two examples can be further generalized to non-integer values of m and b respectively.

At this point, readers are encouraged to find further examples of their own. Any subsequence $\{q_n\}$ of the positive integers may be used in Theorem B. For inspiration, the reader may consult Sloane's *Handbook of Integer Sequences* [3].

This research was supported in part by the U. S. Army Research Office under contract DA–ARO–D–31–124–73–G167.

References

[1] D. Knuth, Fundamental Algorithms, The Art of Computer Programming, vol. 1, Addison-Wesley, Reading, 1968, p. 609.
[2] I. M. Vinogradov, Elements of Number Theory, Dover, New York, 1954, problem 3, p. 29.
[3] N. J. A. Sloane, A Handbook of Integer Sequences, Academic Press, New York, 1973.

Symmetries of the Trihexaflexagon

MICHAEL GILPIN
Michigan Technological University

In this paper we compute the symmetry group of the trihexaflexagon, a figure formed by folding a strip of paper marked off into ten equilateral triangles. Flexagons, in general, are polygons made from folded strips of paper that can be "flexed" to change faces; the trihexaflexagon is the hexagonal shaped flexagon of three faces. Its symmetry group was probably known to the inventors of flexagons (Arthur Stone, Bryant Tuckerman, Richard Feynman and John Tukey) who according to Martin Gardner [2] worked out the complete theory of flexagons. It was probably also known to C. O. Oakley and R. J. Wisner who in [3] mention that one symmetry group associated with the trihexaflexagon is S_3.

The successive folds for making the trihexaflexagon are shown in FIGURE 1. When the position in FIGURE 1iii is reached, triangle C is placed in front of triangle B. Then triangle A is folded on top of triangle C and the two are pasted together. Two representations of the completely constructed trihexaflexagon are shown in FIGURE 2, where we have numbered the nine vertices of the

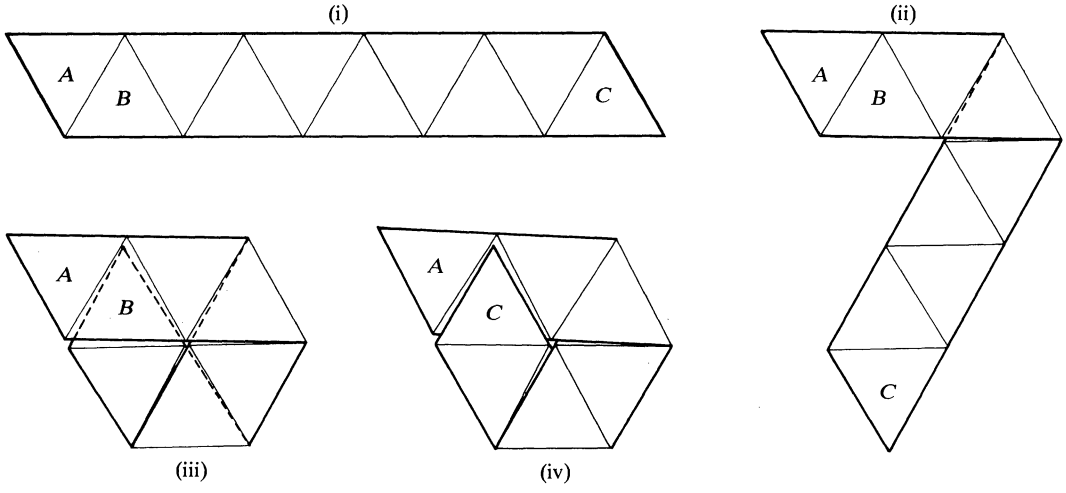


FIGURE 1

trihexaflexagon in succession along the edge of the Moebius band of which it is composed. Each vertex of the trihexaflexagon is determined by six vertices of equilateral triangles and each should therefore be labeled six times. FIGURE 2i is a slightly distorted view of the trihexaflexagon which clearly shows this numbering. FIGURE 2ii is a symbolic representation of the trihexaflexagon that we will use in subsequent illustrations.

To flex the trihexaflexagon, position it as shown in FIGURE 2. Then pinch it at the vertex labeled 1, thereby forcing together vertices 2 and 5. Next push vertex 8 to the meeting point of vertices 2 and 5. The trihexaflexagon can then be opened to a new face by pulling it apart at the point where vertices 3, 6 and 9 meet.

We begin our investigation of the symmetries of the trihexaflexagon by listing (and illustrating: see FIGURE 3) some of them:

I — the identity symmetry;
R — a rotation counterclockwise through 120°;
P — a pinch at the horizontal vertex followed by a clockwise rotation through 60°;
F — a half revolution about the horizontal axis followed by a clockwise rotation through 60°.

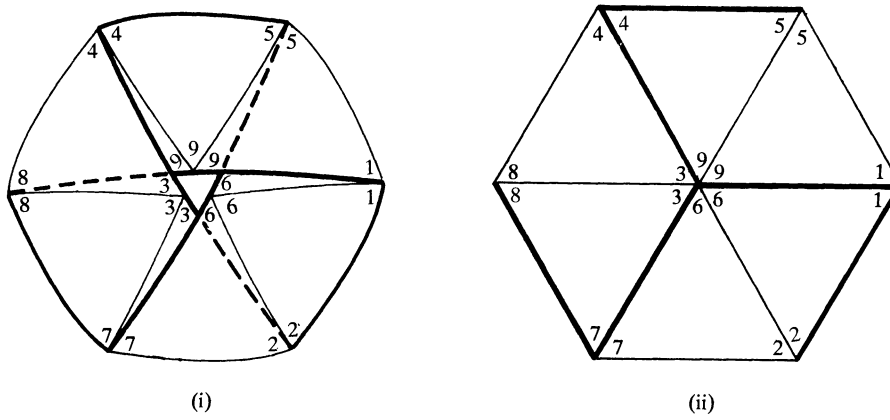
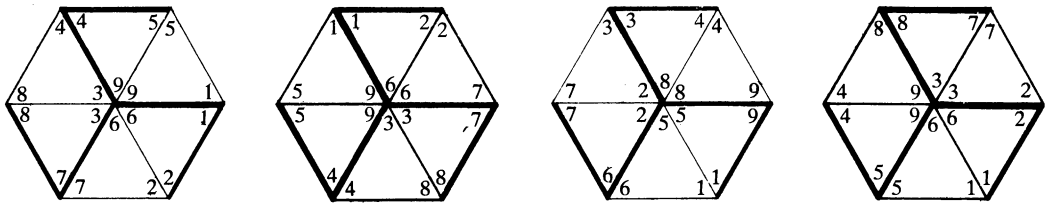


FIGURE 2



I *R* *P* *F*

FIGURE 3

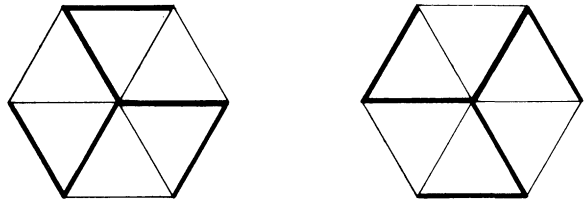
Finally if we take the length of the one edge of the Moebius band to be unity, then for each ε in [0, 1) a movement of the band uniformly along its path through the distance ε is a symmetry which we denote as U_ε. For concreteness let us agree that the direction of U_ε is governed by vertex 1 moving towards vertex 2 and so forth. Then it is easy to verify the relations

$$R^3 = F^2 = P^9 = I; \qquad (PF)^2 = (RF)^2 = I$$

$$U_{1/9} = P; \qquad U_{1/3} = R; \qquad (U_\varepsilon F)^2 = I \quad \text{for all} \quad \varepsilon \text{ in } [0, 1).$$

Now let G denote the full group of symmetries of the Moebius band of the trihexaflexagon. Then G is isomorphic to the group of symmetries of the unit circle—since uniform movements correspond to rotations and flip to flip. Its subgroup G_V consisting of all symmetries that send vertices into vertices on a Moebius band triangulated into nine equilaterial triangles is isomorphic to the symmetry group of a regular 9-sided polygon, namely, the dihedral group D_9. Since the symmetries P and F of the trihexaflexagon also generate D_9 (specifically, $P^9 = F^2 = (PF)^2 = I$), the group of symmetries of the trihexaflexagon must be D_9.

We now examine a more general class of movements of the trihexaflexagon. Consider the two standard positions indicated in FIGURE 4; let S denote the set of all possible positions represented by



(i) (ii)

FIGURE 4

FIGURE 4i, and let S' denote the set of all possible positions represented by FIGURE 4ii. We will call a movement of the trihexaflexagon from one standard position to another a pseudo-symmetry. Then, without going into detail, we can describe the system of pseudo-symmetries of the trihexaflexagon as follows: we have a collection of two objects $\{S, S'\}$ together with a set of pseudo-symmetries which map objects in this collection to other objects in the collection in such a way that

(1) Each object in this collection has associated to it an identity pseudo-symmetry having the usual properties of identities.

(2) Two pseudo-symmetries can be composed if and only if the second one starts where the first one stops.

(3) If f, g, and h are three pseudo-symmetries such that the composition "first f, then g, then h" makes sense, then both $(hg)f$ and $h(gf)$ make sense and $(hg)f = h(gf)$.

Systems that satisfy these conditions are called categories. Thus, the pseudo-symmetries of the trihexaflexagon form a category. In like fashion any hexaflexagon will have a set of standard positions and a collection of pseudo-symmetries which together form a category.

We can only marvel at this invention of Stone *et al.*, in that flexagons are of interest at many levels. The child wonders at an object that can change faces. A high school student can verify and discover the equations governing the symmetries of the trihexaflexagon. The modern algebra student has yet another interesting group to compute, and the graduate student has another example of categories.

The author wishes to thank the referee for several simplifications and Peter Ostlender for help with the drawings.

References

[1] G. Birkhoff and S. MacLane, A Survey of Modern Algebra, Macmillan, New York, 1953.
[2] Martin Gardner, The Scientific American Book of Mathematical Puzzles and Diversions, Simon and Schuster, New York, 1959.
[3] C. O. Oakley and R. J. Wisner, Flexagons, Amer. Math. Monthly, 64 (1957).

Calculating Commutators in Groups

EUGENE SPIEGEL
University of Connecticut

In every introduction to modern algebra, authors and teachers are certain to point out to students that a set which generates a group should not be confused with the group itself. In particular, the set H_G of commutators of a group G (elements of the form $xyx^{-1}y^{-1}$, $x, y \in G$) and the commutator subgroup G' generated by H_G needn't be the same. Carmichael [2], p. 39, gives an appealing illustration of this point with a group G of order 256 whose commutator subgroup G' has order 16, while the set of commutators of G has order 15.

When is it that $H_G = G'$? Of course if G is abelian, $G' = H_G = \{e\}$. In general H_G and G' possess two important properties in common; namely, both are invariant under conjugation and the taking of inverses. For $x \in G$, the set of conjugates of x is the set $C_x = \{axa^{-1} \mid a \in G\}$. Define B_x as the product of the sets C_x and $C_{x^{-1}}$, thus $B_x = \{st \mid s \in C_x, t \in C_{x^{-1}}\}$. The set B_x is a subset of H_G (this is proved in the lemma below) which is invariant in G under conjugation and the taking of inverses. Can $G' = B_x$ for some $x \in G$? This question is examined below for several specific groups which are frequently encountered in an elementary algebra course. It turns out that for all of these groups it is the case not only that $G' = B_x$ for some x, but also that in fact $G' = H_G$. This conclusion follows from a computational lemma and two theorems which we state below. The proofs of these results are given at the end, as our central purpose is to look at several common examples: the first five are applications of our theorems, while the last example shows that H_G may be equal to G' under even weaker conditions than we have stated in our theorems. Hence the example mentioned above (Carmichael [2]) is important: even though for all groups usually discussed in beginning courses in group theory $H_G = G'$, this "proof by example" is not valid.

For convenience let us summarize the notation we will need. G is always a group, H_G the set of its commutators, and G' the commutator subgroup of G. For any element x in G, C_x is the set of conjugates of x and B_x is the product $C_xC_{x^{-1}}$. Finally, let $Z(G)$ denote the center of G, that is $Z(G) = \{a \in G \mid xa = ax \text{ for all } x \in G\}$.

LEMMA. (a) $H_G = \bigcup_{x \in G} B_x$. (b) *If* $y \in B_x$, $C_y \subset B_x$. (c) *If* $y \in B_x$, $y^{-1} \in B_x$. (d) *If* $y = cx'$ *with* $c \in Z(G)$ *and* $x' \in C_x$, *then* $B_x = B_y$.

THEOREM 1. *If for some $x \in G$, B_x is a subgroup containing H_G, then $H_G = B_x = G'$.*

THEOREM 2. *Suppose G contains a normal abelian subgroup A with cyclic factor group G/A generated by xA, $x \in G$. Then $B_x = H_G = G'$.*

Example 1. Let $G = GL(2, \mathbb{C})$ be the group of 2×2 invertible complex matrices and $SL(2, \mathbb{C})$ the subgroup of G of matrices with determinant one. Since any commutator of G has determinant one, $G' \subset SL(2, \mathbb{C})$. Here we show that $G' = B_x = H_G$ where $x = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.

Let $a \in SL(2, \mathbb{C})$. Then a is similar to a matrix of the form $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} -1 & -1 \\ 0 & -1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{smallmatrix}\right)$ for $\alpha \neq 0$, $\alpha \in \mathbb{C}$ (Jordan canonical form). Set $x = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ so that $x = x^{-1}$. An element $g \in G$ is a conjugate of x if and only if x has trace zero and determinant one. (If g is not a scalar matrix, these numbers determine the eigenvalues of g, which in turn determine the Jordan form of g.)

Now $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{smallmatrix}\right) = rxr^{-1}sx^{-1}s^{-1}$ for some $r, s \in G$ if and only if $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{smallmatrix}\right) sxs^{-1} = rxr^{-1}$. But this means precisely that $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{smallmatrix}\right) sxs^{-1} \in C_x$, which is equivalent to the equality $\operatorname{tr}\left[\left(\begin{smallmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{smallmatrix}\right) sxs^{-1}\right] = 0$ for some $s \in G$. Select s such that $sxs^{-1} = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Then $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{smallmatrix}\right) \in B_x$. Similarly, we note that $\operatorname{tr}\left[\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\right] = \operatorname{tr}\left[\left(\begin{smallmatrix} -1 & -1 \\ 0 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\right] = 0$, so that $B_x = SL(2, \mathbb{C})$. By Theorem 1, $B_x = G'$.

Example 2. Here we generalize to fields F of characteristic not 2. As in the previous example, let $G = GL(2, F)$ be the group of 2×2 invertible F-matrices and $SL(2, F)$ the subgroup of G of matrices with determinant one. Once again we find an $x \in G$ such that $B_x = H_G = G' = SL(2, F)$.

If $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = rxr^{-1}sx^{-1}s^{-1}$ with $r, s \in G$, then $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) sxs^{-1} = rxr^{-1}$. But $-\operatorname{tr}(sxs^{-1}) = -\operatorname{tr}(x) = \operatorname{tr}(rxr^{-1}) = \operatorname{tr}(x)$, so $\operatorname{tr}(x) = 0$. Let the determinant of x be δ, with $\delta \neq 0$, $\delta \in F$. Also suppose $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) a = b$ with $a, b \in C_x$. Write $a = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & -\alpha \end{smallmatrix}\right)$ with $\alpha, \beta, \gamma \in F$ and $\det(a) = -\alpha^2 - \beta\gamma = \delta$. Then $\operatorname{tr}\left[\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & -\alpha \end{smallmatrix}\right)\right] = \alpha + \gamma - \alpha = \gamma = \operatorname{tr}(b) = 0$. This says that a is in upper triangular form and $\delta = -\alpha^2$. Then x is similar to $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & -\alpha \end{smallmatrix}\right)$. Let $x' = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. We can replace x by x', since $B_x = B_{x'}$ by part (d) of the Lemma.

Let $g \in SL(2, F)$. Then g is similar to a matrix of the form $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} -1 & -1 \\ 0 & -1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{smallmatrix}\right)\alpha \neq 0$, $\alpha \in F$, or $\left(\begin{smallmatrix} 0 & 1 \\ -1 & \beta \end{smallmatrix}\right)$ with $\beta \in F$ (rational canonical form). In the first three cases we proceed as in Example 1. In the last case, we observe that $\left(\begin{smallmatrix} 1 & -\beta \\ 0 & -1 \end{smallmatrix}\right) \in C_{x'}$ and $\left(\begin{smallmatrix} 0 & 1 \\ -1 & \beta \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & -\beta \\ 0 & -1 \end{smallmatrix}\right) \in C_{x'}$, since their traces are equal. Thus $B_{x'} = H_G = G' = SL(2, F)$. In fact, we have shown that if $x \in G$, $B_x = G'$, if and only if x is similar to a matrix of the form $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & -\lambda \end{smallmatrix}\right)$ for some non-zero λ in F.

Example 3. Let $D_n = \langle a, b \mid a^n = b^2 = 1, bab = a^{-1} \rangle$ be the dihedral group of order $2n$. Let A be the subgroup of D_n generated by a. Then A is a cyclic normal subgroup and D_n/A is of order 2. From Theorem 2 we can conclude $B_b = D_n'$.

Example 4. If p and q are primes, then a "Sylow subgroup type argument" shows that in fact any group of order p^3 or p^4 or pq or p^2q satisfies the hypothesis of Theorem 2. This accounts for all groups of order up to 30 except for groups of order 24.

Example 5. Let $G = S_n$ be the symmetric group on $\{1, 2, \ldots, n\}$, and A_n the group of even permutations in S_n. We first check that $G' = H_G = A_n$. From the definitions, $G' \subset A_n$. We must show $A_n \subset H_G$. Let $g \in A_n$. If we write g as the product of disjoint cycles, there are an even number of cycles each having an even number of letters. Because the conjugate of an element h, of S_n has a similar cycle decomposition as h, it is sufficient to check that $(1, 2, \ldots, 2k+1)$ and $(1, 2, \ldots, 2r)(2r+1, \ldots, 2r+2s)$ are each commutators in the letters involved. But

$$(1, 2, \ldots, 2k+1) = (1, 2, \ldots, k+1)(1, k+2, k+3, \ldots, 2k+1)$$

and

$$(1, 2, \ldots, 2r)(2r+1, \ldots, 2r+2s) = (1, 2, \ldots, r+1, 2r+1, 2r+2, \ldots, 2r+s)$$

$$\cdot (2r+1, r+2, r+3, \ldots, 2r, 1, 2r+s+1, 2r+s+2, \ldots, 2r+2s).$$

Thus $G' = H_G = A_n$.

There remains the question of whether there exists an element $x \in S_n$ such that $B_x = G' = A_n$. If $n = 3$, $S_3 \simeq D_3$ so the answer is positive. If $n = 4$, let $x = (1,2,3)(4)$. An easy calculation verifies that $B_x = A_4$. If $n \geqq 5$, the existence of such an x is essentially the question asked by Brenner [1]. Xu Cheng-Hao [3] showed that if $n \geqq 5$ and $m = [n/2]$, then the element $x = (1,2)(3,4,5,\ldots,2m)$ has the desired property.

Example 6. We show, finally, that it is possible to have a group G with the property that $H_G = G'$, but $B_x \neq H_G$ for any $x \in G$. For each positive integer i, associate a non-abelian group G_i. Let $G = \bigoplus_{i=1}^{\infty} G_i$. An element $x \in G$ can be thought of as a sequence $x = \{x_i\}$ with $x_i \in G_i$ and $x_i = e_i$ (e_i is the identity in G_i) if $i > N(x)$. ($N(x)$ is a positive integer depending upon x.) If $a \in B_x$ and $a = \{a_i\}$, $a_i \in G_i$, then $a_i = e_i$ if $i > N(x)$. Thus $B_x \neq H_G$. $G' = \bigoplus_{i=1}^{\infty} G_i'$ and by a slight abuse of notation $H_G = \bigoplus_{i=1}^{\infty} H_{G_i}$. This implies $G' = H_G$ if and only if $G_i' = H_{G_i}$ for $i = 1,2,\ldots$. To construct the required example we can just select $G_i = S_3$ for $i = 1,2,\ldots$.

Now we conclude with the proofs of the results stated earlier and used in the examples.

Proof of Lemma. (a) To show that $H_G = \bigcup_{x \in G} B_x$, let $x, y \in G$. Then $x \in C_x$, $yx^{-1}y^{-1} \in C_{x^{-1}}$ implies $xyx^{-1}y^{-1} \in B_x$, hence $H_G \subset \bigcup_{x \in G} B_x$. For the converse, it is sufficient to note the identity $(yxy^{-1})(zx^{-1}z^{-1}) = (yxy^{-1})(zy^{-1})(yxy^{-1})^{-1}(zy^{-1})^{-1}$.

(b) To show that $C_y \subset B_x$ whenever $y \in B_x$ we begin with the definition of B_x: $y = uv$ with $u \in C_x$, $v \in C_{x^{-1}}$. Let $w \in G$. Then $wyw^{-1} = (wuw^{-1})(wvw^{-1})$ which is contained in B_x.

(c) To show that $y^{-1} \in B_x$ whenever $y \in B_x$, let $y \in B_x$. Then $y = (rxr^{-1})(sx^{-1}s^{-1})$ with $r, s \in G$. Then $y^{-1} = (sxs^{-1})(rx^{-1}r^{-1})$ which belongs to B_x.

(d) To show that $B_x = B_y$ for $y = cx'$ where $c \in Z(G)$ and $x' \in C_x$, write $x' = txt^{-1}$ with $t \in G$, and let $(rxr^{-1})(sx^{-1}s^{-1}) \in B_x$. Then $(rxr^{-1})(sx^{-1}s^{-1}) = (rt^{-1}ytr^{-1})(st^{-1}y^{-1}ts^{-1}) \in B_y$.

Proof of Theorem 1. This follows almost immediately from the lemma part (a), since our hypotheses tell us that B_x is a subgroup containing all commutators, yet every element of B_x is a commutator.

Proof of Theorem 2. Since G/A is cyclic, any element $h \in G$ can be written in the form $h = ax^i$ for some $a \in A$, $i \in Z$. The proof of the theorem consists of the following sequence of observations.

(1) If $i \in Z$, $B_{x^i} = \{ax^ia^{-1}x^{-i} \mid a \in A\}$. Let $u \in B_{x^i}$, and write $u = rx^ir^{-1}sx^is^{-1}$ with $r = ax^j$, $s = bx^k$ and $a, b \in A$, $j, k \in Z$. Then $u = ax^jx^ix^{-j}a^{-1}bx^kx^{-i}x^{-k}b^{-1} = a(x^ia^{-1}bx^{-i})b^{-1} = ab^{-1}x^iba^{-1}x^{-i}$ since A is abelian and normal.

(2) If $i \in Z$, B_{x^i} is a normal subgroup of G. To prove this, let $ax^ia^{-1}x^{-i}$, $bx^ib^{-1}x^{-i} \in B_{x^i}$ with $a, b \in A$. Then

$$(ax^ia^{-1}x^{-i})(bx^ib^{-1}x^{-i}) = a(x^ia^{-1}x^{-i})b(x^ib^{-1}x^{-i}) = ab(x^ia^{-1}x^{-i})(x^ib^{-1}x^{-i}) = abx^ib^{-1}a^{-1}x^{-i}.$$

Since the right side of this equality belongs to B_{x^i}, so must the left side. By part (c) of the Lemma, B_{x^i} is a group, and from part (b) it is normal.

(3) If $i \in Z$, $B_{x^i} \subset B_x$. Let $ax^ia^{-1}x^{-i} \in B_{x^i}$ with $a \in A$. Since $axa^{-1}x^{-1} \in B_x$, $B_xaxa^{-1} = B_xx$. From (2), $B_xax^ia^{-1} = B_xx^i$, so that $B_xax^ia^{-1}x^{-i} = B_x$. Hence $B_{x^i} \subset B_x$.

(4) $H_G \subset B_x$. Let $c, d \in G$, $c = ax^i$, $d = bx^j$ with $a, b \in A$, $i, j \in Z$. Then

$$cdc^{-1}d^{-1} = ax^ibx^jx^{-i}a^{-1}x^{-j}b^{-1} = a(x^ibx^{-i})(x^ja^{-1}x^{-j})b^{-1} = (ax^ja^{-1}x^{-j})(b^{-1}x^ibx^{-i}) \in B_{x^j}B_{x^i}.$$

Because B_x is a subgroup of G containing B_{x^i} and B_{x^j}, $H_G \subset B_x$. Theorem 1 now implies Theorem 2.

References

[1] J. L. Brenner, Research problems I, group theory, Bull. Amer. Math. Soc., 66 (1960) 275.

[2] R. D. Carmichael, Introduction to the Theory of Groups of Finite Order, Dover, New York, 1956.

[3] Xu Cheng-Hao, The commutators of the alternating group, Sci. Sinica, 14 (1965) 339–342.

Making Change

ELWYN R. BERLEKAMP
University of California at Berkeley

The late John L. Kelly, Jr., of Bell Telephone Laboratories was a member of that set of mathematicians who could no more work without cigarettes and coffee than without paper and pencil. These essential supplies are commonly purchased from vending machines, those primitive computers which accept coins as inputs and dole out coins and essentials as output. Coffee is sold for ten cents per cup. The coffee machine accepts only nickels, dimes and quarters, and gives change only in nickels. If the machine has less than three nickels, then it is unable to make change for a quarter. The more enlightened machines turn on a light to notify the customer of this situation.

The light was on one day when Dr. Kelly and I approached. We both wanted coffee, but we could muster only two nickels and one quarter between us. "What odds will you wager," he asked, "that inserting these two nickels will turn the light out?" To that question we devote this paper.

We must first introduce some appropriate assumptions in order to make the problem more precise. Since the insertion of dimes does not affect the machine's nickel supply, we may consider customers who use dimes to be inconsequential. We assume that, if the light is out, each consequential customer will insert a quarter with probability q or two nickels with probability p. If the light is on, however, we assume that each potentially consequential customer will insert two nickels with probability p', or he will go away (or get a dime) with probability q'. Since $p' \geqq p$, we may set $p' = p + qb$; $q' = qa$. If we consider the unfortunate customer who finds the light on when he arrives with a quarter in hand, we may interpret a as the probability that he will go away (or use a dime) and $b = 1 - a$ as the probability that he will use two nickels. We further assume that different customers behave independently of each other, and that the machine is capable of holding an infinite number of nickels.

Finally, we assume that the probability distribution of nickels has reached a steady state. This assumption will be valid if the machine is allowed to operate for a long time without internal tampering with the number of nickels. If the vendor inserts or removes large numbers of nickels frequently, then we cannot expect this formulation of the problem to provide a reasonable answer to Dr. Kelly's question. However, the only vendor with whom I have discussed the situation stated that

Table of Notation

s_i	is the probability that the coffee machine has i nickels. Then $\Sigma s_{ki} = 1$.
P_L	is the probability that the light is on. Then $P_L = s_0 + s_1 + s_2$.
P_i	is the conditional probability of finding the machine with i nickels ($i = 0, 1$ or 2), given that the light is on. Then $P_i = s_i/P_L$.
p	is the probability that a consequential customer uses two nickels when the light is off.
q	is the probability that a consequential customer uses a quarter when the light is off. Then $q = 1 - p$, since customers who use dimes are inconsequential.
p'	is the probability that a potentially consequential customer uses two nickels when the light is off. Then $p' \geqq p$.
q'	is the probability that a potentially consequential customer does not put in two nickels when the light is on. Then $q' = 1 - p'$.
a	is the probability that a potential customer who arrives with quarter in hand and finds the light on goes away (or uses a dime). Then $a = q'/q$.
b	is the probability that a potential customer who arrives with quarter in hand and finds the light on uses two nickels. Then $b = 1 - a = (q - q')/q = (p' - p)/q$.
r	is the ratio of the probabilities p/q.

he usually confines his activities to refilling the coffee and removing the dimes and quarters, leaving the nickel supply unchanged. Under these circumstances our assumptions are not unreasonable. We will use them to determine the probability s_k that when the customer arrives, the coffee machine contains k nickels.

Let $S(z) = \Sigma s_k z^k$ be the generating function for the state probabilities. (We require, as usual, that $s_k \geqq 0$ and that $\Sigma s_k = 1$.) According to our assumptions, the generating function $S(z)$ must satisfy the equilibrium equation

$$S(z) = [S(z) - (s_0 + s_1 z + s_2 z^2)][pz^2 + qz^{-3}] + [s_0 + s_1 z + s_2 z^2][p'z^2 + q'].$$

In other words, if the light is off, two nickels will be added with probability p, while three will be subtracted with probability $q = 1 - p$, whereas if the light is on, two nickels will be added with probability p', while no change will occur with probability $q' = 1 - p'$. Solving for $S(z)$ gives

$$S(z) = \frac{(s_2 z^2 + s_1 z + s_0)(-qz^{-3} + (p' - p)z^2 + q')}{1 - pz^2 - qz^{-3}}.$$

To transform this expression into a more useful form, we will introduce some new variables. Let P_L be the probability that the light is on, and P_i $(i = 0, 1, 2)$ the conditional probabilities of finding the machine with i nickels if the light is on. Then $P_L = s_0 + s_1 + s_2$, and $P_i = s_i / P_L$. So we replace $p' - p$ by qb, s_i by $P_L P_i$, let $r = p/q$, and simplify, producing

$$S(z) = \frac{P_L[P_2 z^2 + P_1 z + P_0][b(z^4 + z^3) + (z^2 + z + 1)]}{[-r(z^4 + z^3) + (z^2 + z + 1)]}.$$

Since the probabilities must sum to one, we require $S(1) = 1 = P_L(2b + 3)/(3 - 2r)$. From this we deduce that the probability of the light being on must be

$$P_L = \frac{(3 - 2r)}{(2b + 3)} = \frac{(5q - 2)}{q(2b + 3)} = \frac{(3 - 5p)}{(1 - p)(2b + 3)}.$$

This expression is valid only when $r = p/q \leqq 3/2$, or, equivalently, when $p \leqq 3/5$. If $p \geqq 3/5$, then we expect the number of nickels in the machine to increase indefinitely, and the steady-state distribution for which we are searching will not exist. As p approaches 3/5, then P_L approaches zero, corresponding to the fact that if nickels are inserted almost as often as they are needed for change, then the machine should rarely run out. As p approaches zero, P_L approaches $1/(1 + 2b/3)$. Again, this agrees with our intuitive expectations. If b is 0, then nickels are never inserted and once the light goes on it will remain on forever. On the other hand, if $p = 0$, $b = 1$, then the machine states will follow the cycle: $0, 2, 4, 1, 3, 0, 2, 4, 1, 3, \ldots$. Since there are five states in this cycle, and three of them have the light on, P_L must be 3/5 in agreement with the formula. We also note that the conditional state probabilities when the light is on must be given by $P_0 = P_1 = P_2$ in this limiting situation.

For other values of p and b, however, our problem is far from solved. We have a formula for $S(z)$ in terms of the parameters P_L, P_0, P_1, P_2, r, and b, where r and b are presumed known and we have just derived an equation for P_L. But the conditional probabilities P_0, P_1, and P_2 remain unknown, except for the obvious restriction that they sum to one.

How can they be determined? Taking additional moments of the generating function (i.e., $S'(1)$ = the average number of nickels, etc.) will accomplish nothing since no information about the moments of the distribution is known. The original equilibrium equation already embodies all relations between the state probabilities. It appears that we have already used all the available information, and yet our answer still contains two undetermined parameters. Is the problem inadequately specified ? The reader is invited to stop and ponder our predicament before continuing.

As the experienced reader familiar with queueing theory will readily deduce, the key to the resolution of our difficulties lies hidden away in a condition that we have not yet exploited, namely, that $s_k \geqq 0$ for all k. Because of this requirement and the already-exploited condition that $\Sigma s_k = 1$, we

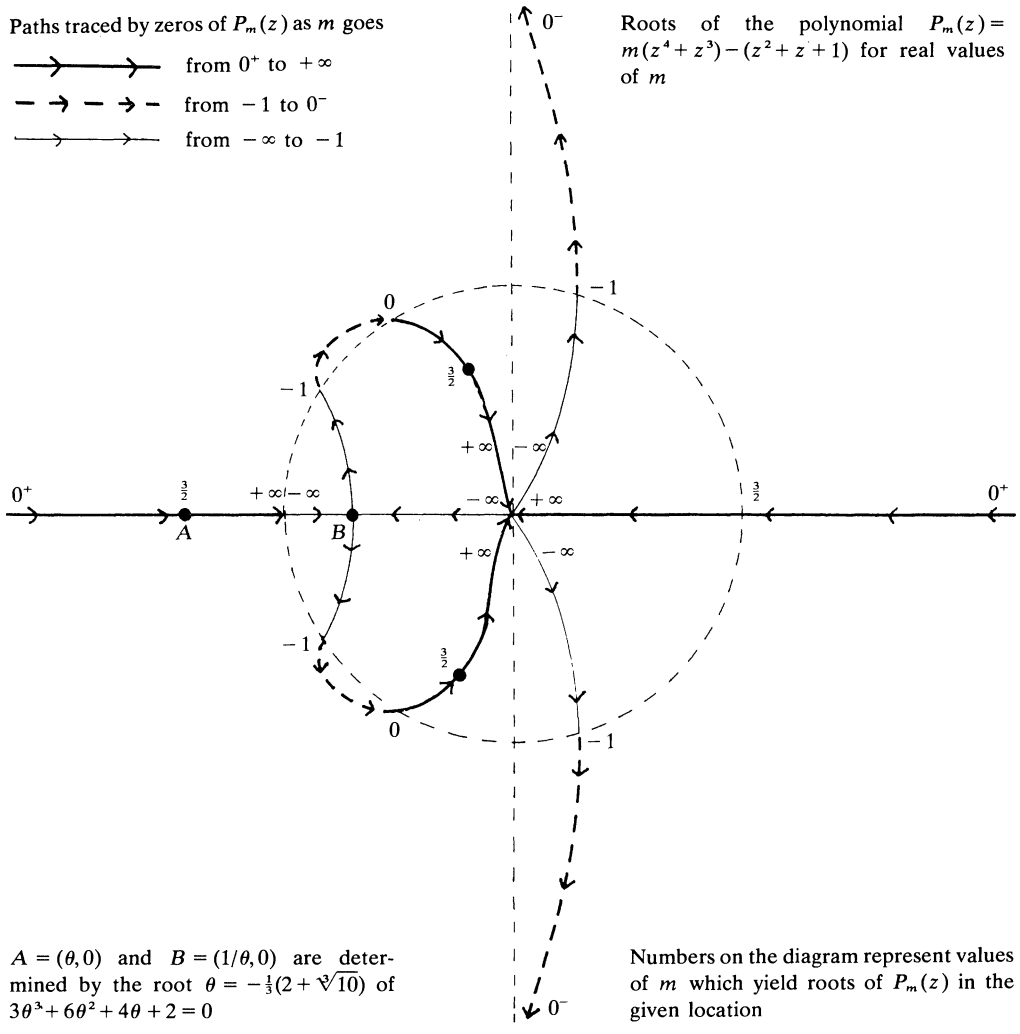


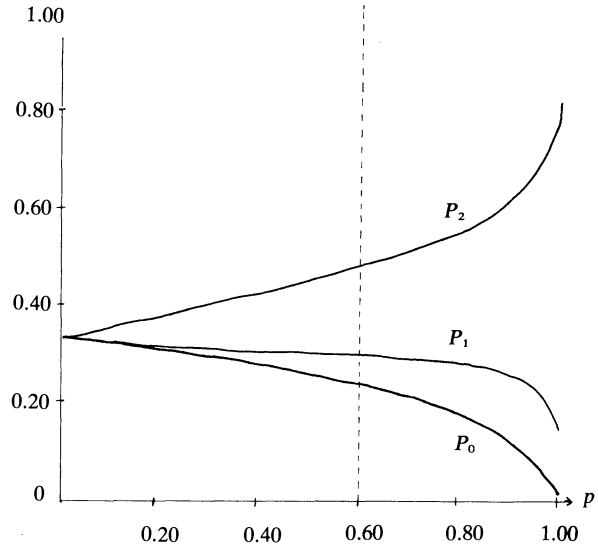
FIGURE 1

conclude that $S(z)$ must be analytic within the unit circle in the complex plane. For, if $|z| \leqq 1$, then $|S(z)| \leqq \Sigma |s_k| |z|^k \leqq \Sigma s_k = 1$. Thus, any poles of $S(z)$ must have magnitude greater than one.

To find these poles we have plotted in FIGURE 1 the location of possible roots of the expression $m(z^4+z^3)-(z^2+z+1)$ since it appears in both the numerator (where $-1 \leqq m \leqq 0$, since $m = -b$) and in the denominator (where $m \geqq 0$ since $m = r$) of $S(z)$. For positive m, the expression has one real positive root, one real negative root, and a pair of complex conjugate roots which have negative real part and lie within the unit circle. For m between -1 and 0, all four roots are complex and lie outside the unit circle. Hence the second numerator factor cannot have any roots in common with the denominator.

For positive r greater than 3/2, the real positive root of the denominator as well as the pair of complex roots lies within the unit circle. Since the first numerator factor $(P_2z^2+P_1z+P_0)$ is only a quadratic, it can cancel out at most two of these three roots. Thus, for $r \geqq 3/2$, the requirement that $S(z)$ be analytic within the unit circle cannot be satisfied. This verifies our earlier observation that if $r \geqq 3/2$ the nickel supply will increase indefinitely and no stationary state can exist.

p	P_0	P_1	P_2
ε	$\frac{1}{3}-\frac{\varepsilon}{9}$	$\frac{1}{3}-\frac{\varepsilon}{9}$	$\frac{1}{3}+\frac{2\varepsilon}{9}$
.119	0.32	0.32	0.36
.274	0.30	0.31	0.40
.388	0.28	0.30	0.42
.498	0.26	0.30	0.44
.600	0.23	0.29	0.48
.800	0.17	0.28	0.55
.954	0.08	0.23	0.69
.976	0.06	0.20	0.74
.9897	0.04	0.17	0.79
.9985	0.01	0.10	0.89
$1-\varepsilon$	$\varepsilon^{2/3}$	$\varepsilon^{1/3}$	$1-\varepsilon^{1/3}$



The conditional state probabilities when the light is on versus p, the consequential customer's probability of using two nickels.

FIGURE 2

On the other hand, if $r < 3/2$, then $S(z)$ will be analytic within the unit circle if and only if the coefficients P_0, P_1, and P_2 are chosen so that the two roots of the first numerator factor cancel out the pair of complex conjugate roots of the denominator. This requirement will give us two additional equations which will then enable us to determine P_0, P_1, and P_2. (Notice that P_0, P_1, and P_2 depend only on $r = p/q$, and not on a and b.)

Since the general quartic polynomial (and thus the denominator of $S(z)$) is factorable, it is possible to find these conditional probabilities explicitly. However, the solution turns out to be too cumbersome to be of much value. A more tractable approach is to solve the equations numerically, to obtain P_0, P_1, and P_2 as functions of p (see FIGURE 2). The graph has been continued for p between 3/5 and 1, even though the problem makes sense only if $0 < p < 3/5$. P_0, P_1 and P_2 all start from 1/3 when $p = 0$, in accord with our earlier observations. P_0 and P_1 both have an initial slope of $-1/9$; P_2 starts with an initial slope of 2/9 and remains very nearly linear for $0 < p < 3/5$.

The answer to Kelly's query is given by the graph of P_0, which represents the probability, depending on p, that his two nickels will not turn the light out.

Mersenne Primes and Group Theory

SHALOM FEIGELSTOCK
Bar-Ilan University, Israel

A well-known open question in number theory concerns the Mersenne primes, that is, those primes of the form $2^p - 1$ where p is itself prime: are there infinitely many Mersenne primes? We shall show that this question is, unexpectedly, related to a certain problem in group theory which is itself only a special case of a large class of similar problems.

Let π be any group property (e.g., commutativity, cyclicity, freeness, etc.). A group G will be said to be an hereditary π-group, if every subgroup of G is a π-group. Every group property gives rise to the problem of classifying the hereditary π-groups, a problem which has well-known solutions in a

number of special cases. For example the class of hereditary abelian (cyclic) groups is the class of all abelian (cyclic) groups, and the class of hereditary free groups is the class of all free groups. (This latter fact is sometimes known as the Neilsen-Schreier Theorem [**2**, Corollary 2.9].) However the class of hereditary finitely generated groups is not the class of all finitely generated groups (for a counterexample, see [**2**, p. 112, exercise 2]), even though all finitely generated abelian groups are hereditary finitely generated [**1**, Theorems 15.5 and 15.6].

If a group G is the multiplicative group of non-zero elements of a finite field we will call it a **field group.** We will show that, with a few exceptions, a group is a hereditary field group if and only if its order is a Mersenne prime. Thus there are infinitely many Mersenne primes if and only if there are infinitely many hereditary field groups.

The key to classifying the hereditary field groups is the following immediate consequence of a theorem of Skolem, [**1** vol. 11, Theorem 127.2, and p. 316 exercise 2]: *G is an hereditary field group if and only if G is cyclic and every divisor of $|G|$ (the order of G) is of the form $p^n - 1$, p a prime.* (Skolem's theorem is an important result in the unsolved problem of classifying the multiplicative groups of fields.)

THEOREM. *Let G be an hereditary field group. Then $|G|$ is even if and only if G is a cyclic group of order* 2, 4, 6, 8, 12, 16, 24, *or* 48; *and $|G|$ is odd if and only if $|G| = 1$ or $|G|$ is a Mersenne prime.*

Proof. From consideration of the field with 2^p elements it is clear that if $|G| = 1$, or if $|G|$ is a Mersenne prime then G is a (vacuously) hereditary field group. Conversely, suppose that $|G|$ is odd, and $|G| > 1$. Let $|G| = \prod_{i=1}^{r} p_i$, p_i a prime, $1 \leqq i \leqq r$. Skolem's theorem implies that $p_i = q_i^{n_i} - 1$, q_i a prime, $1 \leqq i \leqq r$. Since $|G|$ is odd, p_i must be odd. Hence $q_i = 2$ for $1 \leqq i \leqq r$. The reader may verify that $2^n - 1$ is prime if and only if n is a prime. Therefore p_i is a Mersenne prime for all i. To complete the proof it suffices to show that $r = 1$.

Suppose that $r > 1$. Then there exist Mersenne primes $2^p - 1$ and $2^q - 1$ such that $(2^p - 1)(2^q - 1)$ divides $|G|$. By Skolem's theorem $(2^p - 1)(2^q - 1) = r^n - 1$ for r a prime. But that says that $r^n - 1$ is odd hence $r = 2$. Clearly $n > 1$. Expanding this equation with $r = 2$ yields that $2^n = 2^{p+q} - 2^p - 2^q + 2$. Therefore $2^{n-1} = 2^{p+q-1} - 2^{p-1} - 2^{q-1} + 1$ so 2^{n-1} would be odd, a contradiction. Hence $r = 1$, and $|G|$ is a Mersenne prime.

In the case that $|G|$ is even, we may write $|G| = 2^k t$, $k > 0$, and t odd. There exists no prime p such that $p^n - 1 = 32$. Therefore, by Skolem's theorem, 2^5 does not divide $|G|$, and $k \leqq 4$. So if $t = 1$ then G is a cyclic group of order 2, 4, 8, or 16. If $t > 1$ then Skolem's theorem and the case that $|G|$ is odd imply that t is a Mersenne prime, $t = 2^p - 1$. Clearly $2t$ divides $|G|$ so that again by virtue of Skolem's theorem $2^{p+1} - 2 = q^m - 1$, or $2^{p+1} - 1 = q^m$, q a prime. So q^m must be odd, making q an odd prime. If $p = 2$, then $t = 3$, and $|G| = 12$, 24, or 48. If $p \neq 2$, then $2^{p+1} - 1 = (2^{(p+1)/2} - 1)(2^{(p+1)/2} + 1)$. The equality $2^{p+1} - 1 = q^m$ yields that $2^{(p+1)/2} - 1 = q^k$, $2^{(p+1)/2} + 1 = q^r$, and $k + r = m$. This implies that $q^r = q^k + 2$. Hence q^k divides 2. However q is an odd prime. Therefore $k = 0$, $q = 3$, and $r = 1$. Substituting in the equation $2^{p+1} - 1 = q^m$ yields $2^{p+1} = 4$, a contradiction.

To prove the converse statement it is necessary to prove that every divisor of the integers 2, 4, 6, 8, 12, 16, 24, or 48 is of the form $p^n - 1$. This simple exercise completes the proof.

We leave the reader with a suggestion for further study. Call a group a K-field group where K is a class of fields and G is the multiplicative group of a K-field. Our theorem, together with the results in [**1**, Section 127], should enable the reader to determine the hereditary K-field groups for classes K other than the class of finite fields.

References

[1] L. Fuchs, Infinite Abelian Groups, vol. 1 (1970), vol. 11 (1973), Academic Press, New York and London.

[2] W. Magnus, A. Karrass, D. Solitar, Combinatorial Group Theory, Interscience Publishers, Wiley, New York, London, and Sydney, 1966.

A Solvable Diophantine Equation

NORMAN WILDBERGER, student
University of Toronto

Despite the fact that the determination of whether there are integers x, y, z for which $x^n + y^n = z^n$, when n is a positive integer exceeding 2, is a notorious unsolved problem, there are other equations of a similar type for which an infinite family of solutions can be found by elementary means. In this paper we give a method for determining such solutions and the accompanying necessary conditions on the exponents. We will illustrate the procedure first by an example.

EXAMPLE 1. Find integer solutions for $x^4 + y^5 + w^7 - z^6 = 0$. Observing that 5 is relatively prime to 4, 7 and 6, choose integers a, b, c so that $d = c^6 - a^4 - b^7$ is non-zero. Then

$$a^4 + d + b^7 - c^6 = 0. \tag{1}$$

We choose a number divisible by 4, 7 and 6 and congruent to -1 modulo 5; for example 504. Multiply (1) by d^{504} to obtain

$$(ad^{126})^4 + (d^{101})^5 + (bd^{72})^7 - (cd^{84})^6 = 0.$$

Thus $(x, y, w, z) = (ad^{126}, d^{101}, bd^{72}, cd^{84})$ is a nontrivial solution of the given equation.

THEOREM. *Let $\{a_1, a_2, \ldots, a_n\}$ be any set of natural numbers of which at least one is relatively prime to the rest. Then any Diophantine equation of the form*

$$c_1 x_1^{a_1} + c_2 x_2^{a_2} + \cdots + c_n x_n^{a_n} = 0,$$

with integer coefficients $c_1, c_2, \ldots, c_n$ has infinitely many integer solutions $(x_1, x_2, \ldots, x_n)$.

Proof. Suppose, without loss of generality, that a_1 is relatively prime to each a_i $(i = 2, \ldots, n)$. Choose any integers $h_2, h_3, \ldots, h_n$ such that

$$u = -\sum_{i=2}^{n} c_i h_i^{a_i} \neq 0.$$

Then, for any positive integer k,

$$u^k = -\sum_{i=2}^{n} c_i h_i^{a_i} u^{k-1}. \tag{2}$$

Since a_1 is relatively prime to the other a_i, a general form of the Chinese Remainder Theorem [**1**, p. 47] allows us to choose k so that for some positive integers $f_1, f_2, \ldots, f_n$, $k = a_1 f_1$ and $k - 1 = a_2 f_2 = a_3 f_3 = \cdots = a_n f_n$. Thus

$$(u^{f_1})^{a_1} + \sum_{i=2}^{n} c_i (u^{f_i} h_i)^{a_i} = 0. \tag{3}$$

Again, by the general Chinese Remainder Theorem, p can be chosen so that for positive integers $g_1, g_2, \ldots, g_n$, $p - 1 = a_1 g_1$ and $p = a_2 g_2 = a_3 g_3 = \cdots = a_n g_n$. Multiplying (3) by c_1^p we obtain

$$c_1 (c_1^{g_1} u^{f_1})^{a_1} + \sum_{i=2}^{n} c_i (c_1^{g_i} u^{f_i} h_i)^{a_i} = 0.$$

Thus

$$(x_1, x_2, \ldots, x_n) = (c_1^{g_1} u^{f_1}, c_1^{g_2} u^{f_2} h_2, \ldots, c_1^{g_n} u^{f_n} h_n)$$

is a solution to the equation. The discretion in choosing u, f_i and g_i clearly allows for infinitely many solutions.

EXAMPLE 2. Find a solution to the equation $5x^6 + 2y^3 = 3z^5$. In the notation of the theorem, this is $3x_1^5 - 5x_2^6 - 2x_3^3 = 0$. If $h_2 = h_3 = 1$ we have $u = 7$. We may take $k = 25$ and $p = 6$ to obtain the relation

$$3(3\cdot 7^5)^5 - 5(3\cdot 7^4)^6 - 2(3^2\cdot 7^8)^3 = 0$$

so that a solution $(x, y, z) = (3\cdot 7^4, 3^2\cdot 7^8, 3\cdot 7^5)$ is determined.

In any solution found by the above method, $x_1, x_2, \ldots, x_n$ all have a common factor. Thus this method does not necessarily produce a complete set of solutions. Of course the method fails to provide solutions to the equation in Fermat's Theorem where all exponents are the same.

Reference

[1] J. E. Shockley, Introduction to Number Theory, Holt, Rinehart and Winston, New York, 1967.

Theodorus' Irrationality Proofs

ROBERT L. MCCABE
Southeastern Massachusetts University

The Pythagorean proof of the irrationality of $\sqrt{2}$ is well known. If we assume that $\sqrt{2} = a/b$ and that a/b is in lowest terms, then $2b^2 = a^2$ implies that a^2 and hence that a is even, that is, $a = 2k$. Substituting $2k$ for a yields $2b^2 = 4k^2$, or $b^2 = 2k^2$. Therefore b^2 and hence b is even. But a and b cannot both be even if a/b is in lowest terms.

This proof, using only the concepts of even and odd, generalizes to almost all other square roots thereby shedding light on an ancient problem concerning incommensurables, or irrational numbers.

Plato's *Theaetetus* contains a famous passage on irrational numbers. Theaetetus says [**1**, p. 25],

> Theodorus here was drawing some figures for us in illustration of roots, showing that squares containing three square feet and five square feet are not commensurable in length with the unit of the foot, and so, selecting each one in its turn *up to* the square containing seventeen square feet; and at that he stopped. [Italics added]

The question that has puzzled historians ever since is this: What manner of proof did Theodorus use? It is certain that Theodorus knew the proof of the irrationality of $\sqrt{2}$, but did not have any general method of proof available, else why stop at 17?

Van der Waerden [**6**, pp. 142–145], following Zeuthen, suggests a method of proof involving ratios, which after a few transformations begin to cycle themselves endlessly thus leading to proofs by contradiction. The method works well *through* 17 (18 is trivial since $\sqrt{18} = 3\sqrt{2}$), and 19 is "quite complicated" because it requires eight ratios before the endless cycle begins. Van der Waerden may be right, but his suggestion is not convincing.

Hardy and Wright [**2**, pp. 41–43] suggest the familiar method involving remainders. For example, if $\sqrt{5} = a/b$ with a/b in lowest terms, then $5b^2 = a^2$ implies that 5 divides a^2. Does 5 then divide a? There are five possibilities: a is of the form $5n$, $5n + 1$, $5n + 2$, $5n + 3$, or $5n + 4$. Squaring each shows that only $(5n)^2$ turns out to be divisible by 5. Thus $a = 5n$. Substituting, we easily find that 5 also

divides b so that a/b was not in lowest terms. Unfortunately this proof works for all squares and Hardy and Wright can give no explanation for Theodorus stopping at 17 except that "he may well have been quite tired."

Heath [**3**, p. 133] suggests that Theodorus "may have adapted the Pythagorean proof in the case of $\sqrt{2}$, by substituting $3, 5, \ldots$ for 2." This suggestion is quite perceptive, for the adaptation of the famous even-odd Pythagorean proof fails precisely at 17 — a remarkable fact that has apparently gone unnoticed. (Prof. Dirk Struik has called my attention to an unpublished Ph.D. disssertation [**4**] of W. R. Knorr in which a similar conclusion is reached. Knorr suggests a rather long detour through Pythagorean triples, but in essence argues for an even-odd mode of proof.)

The following theorem, also proved by Knorr, sheds light on the matter and may well supply the answer to Theodorus' method.

THEOREM. *If p is a positive integer which can be written in any one of the following forms, $4n+2$, $4n+3$, or $8n+5$, for $n = 0, 1, 2, \ldots$, then $\sqrt{p}$ can be proved irrational using only even-odd techniques.*

The proof is easy. Assuming, for example, that $\sqrt{8n+5} = a/b$ in lowest terms, we have $(8n+5)b^2 = a^2$, so that both b and a must be odd. Letting $b = 2j+1$ and $a = 2k+1$, substituting and simplifying, we get the equation

$$8nj^2 + 8nj + 2n + 5(j^2 + j) + 1 = k^2 + k. \tag{1}$$

Since the combinations $j^2 + j$ and $k^2 + k$ are always even, for all j and k, we have a contradiction—for now the left side of (1) is odd and the right side even. The proofs for $4n+2$ and $4n+3$ are similar.

Thus only numbers of the form $4n+1$ are not covered by the theorem (those of the form $4n$ are easily reduced to one of the forms since $\sqrt{4n} = 2\sqrt{n}$). But, since $8n+5 = 4(2n+1)+1$, it is only numbers of the form $4(2n)+1 = 8n+1$ whose square roots cannot easily be handled by this method. These numbers are

$$1, 9, 17, 25, 33, 41, 49, 56, \ldots .$$

The list contains all the odd perfect squares because they are all of the form

$$8\left(\frac{x(x+1)}{2}\right) + 1, \quad \text{where} \quad x = 0, 1, 2, \ldots . \tag{2}$$

Therefore the theorem excludes precisely those positive integers *not* of the form (2).

Thus a direct application of the Pythagorean even-odd technique fails precisely at 17. A straightforward attack on 17, using only even and odd techniques, proves fruitless, or at least too tedious to warrant continuing. The first step is to assume $\sqrt{17} = a/b$ with a/b in lowest terms. Letting $b = 2j+1$ and $a = 2k+1$ leads to the equation

$$17(j^2 + j) + 4 = k^2 + k. \tag{3}$$

There is no even-odd contradiction here, hence there are four possibilities: j and k are both even, j is even and k is odd, etc. Trying each in turn leads again to still further possibilities. The reader is encouraged to try it.

The coincidence here seems too strong to resist. Knowing the Greeks' powerful belief in the theory of the even and the odd, it seems reasonable that Theodorus may well have tried using this theory on the general problem of irrationals.

The translation of Plato is critical. Referring to the quotation from the *Theaetetus*, van der Waerden [**6**, p. 142] translates the latter part of

> ... καὶ οὕτω κατα μίαν ἑκάστην προαιρούμενος μέχρὶ τῆς ἑπτακαιδεκάποδος — ἐν δὲ ταύτη πως ἐνέσχετο.

as

... up to the one of 17 feet; here something stopped him (or: here he stopped).

Heath [3, p. 132] writes

... up to seventeen square feet, 'at which point for some reason he stopped'.

The Greek word μεχρι meaning "up to" or "until" has more the sense of "just short of" [5, p. 1123] and suggests what I believe to be true: Theodorus generalized the Pythagorean even-odd concept and, as shown by the theorem, got stuck at 17. He did not have the algebra necessary to prove the theorem in its full generality and would have handled each integer separately. It is worth quoting van der Waerden again [6, p. 109]:

> For the Pythagoreans, even and odd are not only the fundamental concepts of arithmetic, but indeed the basic principles of all nature.

And,

> Plato always defines Arithmetica as 'the theory of the even and the odd'.

References

[1] H. N. Fowler, Plato: Theaetetus — Sophist, Cambridge, Mass., 1921.
[2] G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, 4th ed., London, 1960.
[3] T. L. Heath, A Manual of Greek Mathematics, London, 1931.
[4] W. R. Knorr, The Pre-Euclidean Theory of Incommensurable Magnitudes, unpublished doctoral dissertation, Dept. of the History of Science, Harvard University, 1972.
[5] H. G. Liddell and R. Scott, A Greek-English Lexicon, revised and augmented by H. S. Jones, Oxford, 1968.
[6] B. L. van der Waerden, Science Awakening, Groningen, Holland (English translation by Arnold Dresden), 1954.

A Multiplicative Metric

Doris J. Schattschneider
Moravian College

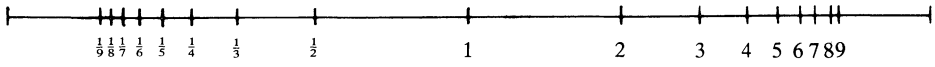
Collectors of metrics are a strange breed known to inhabit topological caves and caverns of analysis, ready to pounce on unwary examples. This specimen was hatched late at night while searching for metrics to spring on an elementary topology class.

For $x, y \in R$, define

$$d(x,y) = \begin{cases} \dfrac{|x-y|}{|x|+|y|}, & \text{if } x, y \text{ are not both } 0, \\ 0, & \text{if } x = y = 0. \end{cases}$$

Clearly $d(x,y) \geqq 0$ and $d(x,y) = d(y,x)$ for all $x, y \in R$. The triangle inequality $d(x,y) \leqq d(x,z) + d(z,y)$ also holds, but its proof can be an algebraic nightmare. We give a version below which reduces calculations somewhat; perhaps a reader will be able to supply a short, neat proof.

First, note that the well-known inequality $|x-y| \leqq |x| + |y|$ implies that $d(x,y) \leqq 1$ for all x, y in R. It is easy to verify the following:



THE "MULTIPLICATIVE METRIC" VIEWED THROUGH EUCLIDEAN EYES: The point 1 serves as the center of the positive reals, and as $r \to \infty$, the distances $d(1/r, 1)$ and $d(1, r)$ remain equal to each other, increasing to 1 in value. Distance along the negative part of the real axis behaves in a similar fashion, but no distance is "added" if we cross the origin: -1, for instance, is no further away from $+1$ than is 0 since $d(-1, +1) = d(0, 1) = 1$.

(i) $d(0, x) = 1$ for all $x \neq 0$.
(ii) $d(x, y) = 1$ if x and y have opposite signs.
(iii) $0 < d(x, y) < 1$ if and only if $x \neq y$ and both are positive or both negative.
(iv) $d(x, y) = d(-x, -y)$.

To prove the triangle inequality, we consider separately three cases. If any two of the three real numbers x, y, z are equal, the inequality clearly holds. If any one of the three numbers x, y, z equals zero or differs in sign from the other two, then properties (i)–(iii) imply that $d(x, y) \leqq 1$ and $d(x, z) + d(z, y) > 1$. Finally, if the inequality were proved for x, y, z all positive, then property (iv) would imply that it also holds for x, y, z all negative, and we would be done.

So let x, y, z be all positive. The triangle inequality then holds if and only if the sum

$$\frac{|x - z|}{x + z} + \frac{|z - y|}{z + y} - \frac{|x - y|}{x + y}$$

is non-negative. Simplifying, this is equivalent to the following sum being non-negative: $y^2|x - z| + x^2|z - y| - z^2|x - y| + cd$, where $c = xz + yz + xy$ and $d = |x - z| + |z - y| - |x - y|$. Since interchanging x and y does not change the sum to be considered, computations to prove the sum is non-negative are necessary only for the three cases when $0 < x < y$. Verification of these three cases is left to the reader.

The metric d possesses both predictable and unexpected geometric properties. Since it is a bounded metric, it behaves predictably like a "perspective" metric. If we consider points equally spaced at intervals of length c along the real line R in the usual metric, these points seem to converge at $\pm\infty$ in the d-metric. In fact, for any $c \in R$, $\lim_{x \to +\infty} d(x, x + c) = \lim_{x \to -\infty} d(x, x + c) = 0$. It is also easy to verify the expected property that for any $c \in R$, $\lim_{x \to +\infty} d(x, c) = \lim_{x \to -\infty} d(x, c) = 1$. But we have already noted in (i), (ii) above that the d-distance from $x \neq 0$ on R to zero or to any other point having opposite sign is 1. Thus, for example, if c is a positive number, it is as "far" to zero as it is to the "ends" of the line R; in fact, it is no "farther" d-distance to any negative number than it is to zero.

Property (iv) states that the d-metric is invariant under the mapping which sends real numbers to their additive inverses; this property is shared by the usual metric on R. However, the d-metric is also invariant under the mapping which sends real numbers to their multiplicative inverses! It is easy to verify that $d(x, y) = d(1/x, 1/y)$ for all $x, y \neq 0$. An interesting consequence of this property is that for any $c \neq 0$, $d(c, 1) = d(1/c, 1)$ and $d(c, -1) = d(1/c, -1)$.

The metric d has yet another invariance property. The multiplicative group $G = R - \{0\}$ can be considered as a group of automorphisms of R, where each $a \in G$ is identified with the mapping $x \to ax$, $x \in R$. For each $a \in G$, and $x, y \in R$, it is easily shown that $d(ax, ay) = d(x, y)$; thus the metric d is invariant under the group G. An interesting comparison should be noted here. It is well known that the isometries of R with the usual metric are the translations $x \to x + a$, the reflection $x \to -x$, and the glide-reflections $x \to -x + a$. The multiplicative counterparts of these mappings, the

dilations $x \to ax$, $a \neq 0$ the inversion $x \to 1/x$ and their composites $x \to a/x$, $a \neq 0$ are isometries of R with the d metric. (These last mappings are extended to all of R by sending $0 \to 0$.) It is an interesting exercise to verify that these are the only isometries of R with the d-metric. For this reason, we might label d a "multiplicative metric" for R.

The topology on R induced by the d-metric also contains some surprises, and reflects the invariance properties of d. For any point $x \in R$ we denote by $B(x; r)$ the open ball of radius r about x, namely the set $\{y \in R \mid d(x, y) < r\}$. If $r \geqq 1$, the ball $B(x; r)$ is all of R. For $0 < r < 1$, the ball about 0 is just the one-point set $\{0\}$, and for all $x \neq 0$, the sets $B(x; r)$ are open intervals in the usual topology on R. The precise description of these balls is

$$B(x; r) = \left\{ y \,\middle|\, x\frac{(1-r)}{(1+r)} < y < x\frac{(1+r)}{(1-r)} \right\}$$

if $x > 0$, and $B(x; r) = \{-y \mid y \in B(-x; r)\}$ if $x < 0$. Since $(1-r)/(1+r) = d(1, r)$, the ball $B(1; r)$ is just the set $\{y \mid d(1, r) < y < 1/d(1, r)\}$, and in general, if $x \neq 0$, $B(x; r) = \{xy \mid y \in B(1; r)\}$. Thus every open ball of radius $r < 1$ is a "multiplicative translate" of the open ball $B(1; r)$. (Compare this with the usual topology on R in which every open ball is an "additive translate" of an open ball about 0.) For example, $B(1; 1/2) = \{y \mid 1/3 < y < 3\}$; therefore, $B(1/3; 1/2) = (1/3)B(1; 1/2) = \{y \mid 1/9 < y < 1\}$, and $B(3/2; 1/2) = (3/2)B(1; 1/2) = \{y \mid 1/2 < y < 9/2\}$.

Since for $x \neq 0$, an open ball of radius r about x is an open interval in the usual topology on R, and the length of this interval approaches 0 as r approaches 0, the topology induced by d on $R - \{0\}$ is just the usual topology. The metric d disconnects the whole space R; in fact, R is the disjoint union of the three connected components R_+, $\{0\}$, R_-. No sequence can converge to 0 except the constant sequence $\{0\}$. However, a sequence converges to a limit $c \neq 0$ in the d-topology if and only if it converges to c in the usual topology on R.

The reader is invited to continue this investigation of the topological properties of the d-metric, and to attempt to extend it in a natural way to higher dimensions. Since the given proof of the triangle inequality depends on the ordering of R, it is not immediately apparent whether an analogous "multiplicative" metric can be defined on R^n.

Groups of Singular Matrices

COLONEL JOHNSON, JR.
Southern University

Groups of matrices under the operation of multiplication provide good examples in even the most elementary course of linear or abstract algebra. Since it is invariably assumed that the "identity" of such a group is the identity matrix (i.e., the diagonal matrix with all diagonal entries equal to 1), the only such groups studied are then necessarily composed of non-singular matrices.

What happens when we throw away this prejudicial assumption? Are there non-trivial groups of matrices containing singular matrices? If so, what do they look like? There are, as we show below, many examples of such groups. We will characterize all such groups, and discover (not surprisingly) that such groups are isomorphic to groups of non-singular matrices.

We begin our study by establishing notation and listing some elementary examples. All matrices considered in this note are square, of order n. The identity of a group of matrices under multiplication will be denoted as E, and the inverse of a matrix A in such a group will be denoted as A'; thus $AA' = A'A = E$. Only when A is non-singular will the usual notation A^{-1} for the inverse of A be

used. I_k will denote the diagonal $k \times k$ matrix with all diagonal entries equal to 1. The letters G and H will be reserved for groups of matrices.

EXAMPLE 1. Let $G = \{A\}$, where the matrix A satisfies $A^2 = A$. Then G is a group with one element where $A = A' = E$. Three special cases are: (i) $A = 0$, the zero matrix;

$$\text{(ii)} \quad A = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}, \; k \leqq n; \quad \text{and} \quad \text{(iii)} \quad A = \begin{pmatrix} I_k & M \\ 0 & 0 \end{pmatrix}, \; k \leqq n;$$

where M is any $k \times n - k$ matrix.

EXAMPLE 2. Let H be a group of non-singular matrices of order $k < n$. Then the set of matrices of the form $\left(\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $A \in H$, is a group G of singular matrices. Here

$$E = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}' = \begin{pmatrix} A^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

EXAMPLE 3. Let H be a group of non-singular matrices of order $k \leqq n$. Then the set of matrices of the form $\left(\begin{smallmatrix} A & AM \\ 0 & 0 \end{smallmatrix}\right)$, $A \in H$, where M is a fixed $n \times n - k$ matrix is a group which we shall denote by $G(H; M)$. (Note that if $n = k$, then $\left(\begin{smallmatrix} A & AM \\ 0 & 0 \end{smallmatrix}\right) = A$.) Here

$$E = \begin{pmatrix} I_k & M \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} A & AM \\ 0 & 0 \end{pmatrix}' = \begin{pmatrix} A^{-1} & A^{-1}M \\ 0 & 0 \end{pmatrix}.$$

All groups in Examples 1 and 2 are of this type for a suitable choice of the group H and the matrix M.

EXAMPLE 4. Two groups G_1 and G_2 of matrices of order n are **conjugate** if there exists a non-singular matrix P of order n such that $G_2 = \{P^{-1}AP \mid A \in G_1\}$. Given a group $G(H; M)$ of singular matrices, and a non-singular matrix P of order n, the conjugate group $G = \{P^{-1}AP \mid A \in G(H; M)\}$ is also a group of singular matrices. The group G differs from $G(H; M)$ only by a change of basis. In fact, as we will show, $G(H, M)$ differs from $G(H, 0)$ only by a change of basis. We will use changes of bases to show in our main theorem that all groups of singular matrices are of the form of Example 2.

Known properties of groups of non-singular matrices suggest some questions to be asked about general matrix groups. Must all elements of a matrix group have the same rank? Does every singular matrix belong to a matrix group? The answer to the first question is quite simple: all elements of a matrix group must have the same rank since if A and B are matrices of order n, then rank B = rank $A'AB \leqq$ rank A. The answer to the second question, however, is more complex: *A matrix B is an element of a group of matrices if and only if rank B^2 = rank B.*

One part of this characterization (the "only if" part) follows from the answer to our first question. So we only need to prove that if rank B^2 = rank B, then B belongs to a group of matrices. View B as a linear transformation on a vector space V. From the assumption, $\dim(\ker B^2) = \dim(\ker B)$. Since $\ker B \subseteq \ker B^2$, it follows that $\ker B^2 = \ker B$, Image $B \cap \ker B = \{0\}$, and $V = \text{Image } B \oplus \ker B$. Let $\alpha_1, \ldots, \alpha_m$ be a basis for Image B and let $\beta_1, \beta_2, \ldots, \beta_m$ be a basis for $\ker B$. The set $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m$ is a basis for V and the matrix of the linear transformation B relative to this basis is of the form $\left(\begin{smallmatrix} S & 0 \\ 0 & 0 \end{smallmatrix}\right)$ where S is a non-singular matrix of order k. From Example 2, $\left(\begin{smallmatrix} S & 0 \\ 0 & 0 \end{smallmatrix}\right)$ belongs to a group G of matrices. If P is the matrix of change of basis, then, from Example 4, the matrix B belongs to the group $G_1 = \{PAP^{-1} \mid A \in G\}$.

We note that if B is a matrix of order n, then rank B^m is a monotonically decreasing (integer-valued) function of m. Thus there exists a positive integer m (which can be shown to be less than or equal to n) such that $(B^m)^2 = B^m$. This means that B^m belongs to a group of matrices. This group is trivial only in the case that B is nilpotent, i.e., $B^k = 0$ for some $k \leqq n$. For a matrix E to be the identity of a group G says that $E^2 = E$, that is, E is a projection. This is a stronger hypothesis than the remark that rank E^2 = rank E and forms the basis for our main result.

THEOREM. *If G is a matrix group where each element is of rank k, then G is isomorphic to a group $G(H;0)$ for some group H of non-singular matrices of order k.*

Proof. Since $E^2 = E$ it follows that E, viewed as a linear transformation, is the identity when restricted to its image. This together with the above basis argument implies that in a properly chosen basis, E is of the form $\left(\begin{smallmatrix} I_k & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Let P be the change of basis matrix and $G_1 = \{P^{-1}NP, N \in G\}$. Thus $P^{-1}EP = \left(\begin{smallmatrix} I_k & 0 \\ 0 & 0 \end{smallmatrix}\right)$ is the identity for G_1. Since $\left(\begin{smallmatrix} I_k & 0 \\ 0 & 0 \end{smallmatrix}\right)S = S$ for any $S \in G_1$ it follows that $S = \left(\begin{smallmatrix} A & B \\ 0 & 0 \end{smallmatrix}\right)$ where A is of order k and B is a $k \times (n-k)$ matrix. Since $S\left(\begin{smallmatrix} I_k & 0 \\ 0 & 0 \end{smallmatrix}\right) = S$ we may conclude that S must be of the form $\left(\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Finally if $S' = \left(\begin{smallmatrix} A_1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ we see that $AA_1 = I_k$ so A is non-singular of order k.

The obvious fact that $G(H,0)$ is isomorphic to the group H justifies our claim that every group of matrices is isomorphic to a group of non-singular matrices. Furthermore a less judicious choice of basis in the above proof (extend the basis $\alpha_1, \ldots, \alpha_m$ of the image of E to a basis of V without consideration of the kernel of E) would have resulted in our finding that

$$E = \begin{pmatrix} I_k & M \\ 0 & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} A & AM \\ 0 & 0 \end{pmatrix}, \qquad S' = \begin{pmatrix} A^{-1} & A^{-1}M \\ 0 & 0 \end{pmatrix}.$$

So in fact, the pattern $G(H,M)$ of Example 3 is just that of $G(H,0)$ up to a choice of basis even though computations of conjugations $P^{-1}SP$ will not readily yield this information.

The author wishes to acknowledge helpful suggestions of Herman Harris, John Dyer, and the editors.

The Conway Stones: What the Original Hebrew May Have Been

DANIEL M. BERRY
University of California at Los Angeles

MOSHE YAVNE
Jet Propulsion Laboratory, Pasadena

In a recent book, *Surreal Numbers* [3], Donald E. Knuth demonstrates a new kind of mathematics textbook called a mathematical novelette and uses this to teach John Horton Conway's axioms for the arithmetic of integer, rational, real, surreal, and transfinite numbers. The main idea behind the new style text is that it teaches in an entertaining way the methodology of mathematical creativity and discovery. This particular novelette deals with two bored ex-students, Bill and Alice, who are hiding from civilization in caves off the Indian Ocean and stumble on an old Hebrew stone tablet. With the remnants of his knowledge of Hebrew, Bill sets down to produce a translation.

> Okay, Alice, I've got it. There are a few doubtful places, a couple signs I don't recognize; you know, maybe some obsolete word forms. Overall I think I know what it says, though I don't know what it means. Here's a fairly literal translation.

Bill's translation is reproduced in FIGURES 1 and 2.

The novelette narrates Bill's and Alice's brainstorming as they struggle to figure out what the stone tablets mean. They try, scrap, and re-try many potential theorems. Some of these theorems are relatively easy to discover and to prove, while others are much more difficult to discover or to prove, taking several failures before they hit on the right idea. As they slowly get the hang of working with the axiomatized system, the "two ex-students turn on to pure mathematics and find total happiness."

When we read this, we began to wonder what the original Hebrew wording of the tablets might have been. We figured that with the second author's native command of Hebrew and imperfect knowledge of English, with the first author's native command of English and more imperfect knowledge of Hebrew, and with some previous cooperation on a literary venture, we had the resources to do the job. So, armed with Hertz's Pentateuch [2], Alcalay's English-Hebrew and Hebrew-English Dictionaries [1], Knuth's book, paper, pencils, and erasers, we sat down and argued our way to the "original" Hebrew version of the Conway Stones.

We adopted the following basic constraints for the translation: The narrative wording had to be as biblical as possible. Any time an obviously biblical idiom or phrase was used, we would try to use the original Hebrew for that idiom or phrase. (This necessitated finding the phrase in the English half of the Pentateuch and then using the corresponding Hebrew.) However, the mathematical portions of the text were to be translated into semi-modern Hebrew as precise as the English. Ambiguities (or puns) in the English wording were to show up as similar ambiguities (or puns) in the Hebrew wording.

1	In the beginning, everything was void, and J.H.W.H. Conway began to create numbers.	1	בראשית היה תהו ובהו, כנוי יהוה החל לברא את
2	Conway said, "Let there be two rules which	2	המספרים. ויאמר כנוי "יהיו שני חוקים אשר יצרו את
3	bring forth all numbers large and small. This	3	כל המספרים הגדולים כקטנים. ויהי זה החוק הראשון:
4	shall be the first rule: Every number corresponds to two sets of previously created numbers, such that no member of the left set is greater than or equal to any member of the right set.	4	כל מספר מקביל לשתי קבוצות של מספרים אשר נבראו בעבר, ואשר מקימים את התכונה ששום גורם בקבוצה השמאלית גדול או שוה לגורם מהקבוצה
5	And the second rule shall be this: One number	5	הימנית. ויהי החוק השני כדלקמן: אחד המספרים הנו
6	is less than or equal to another number if and only if no member of the first number's left set is greater than or equal to the second number, and no member of the second number's right set	6	קטן או שוה למספר אחר אם ורק אם אף אחד מהגורמים של הקבוצה השמאלית של המספר הראשון הינו גדול או שוה למספר השני, ואף אחד מהגורמים בקבוצה הימנית של המספר השני הינו קטן או שוה
7	is less than or equal to the first number." And Conway examined these two rules he had made,	7	למספר הראשון". ויבחן כנוי את שני החוקים שיצר,
8	and behold! they were very good.	8	והנה? ראה כי טוב.
9	And the first number was created from the void	9	ונברא המספר הראשון מהקבוצה הריקה השמאלית
10	left set and the void right set. Conway called this number "zero," and said that it shall be a sign to separate positive numbers from negative	10	ומהקבוצה הריקה הימנית. ויקרא כנוי את המספר הזה "אפס", והיה לאות המבדיל בין המספרים החיוביים
11	numbers. Conway proved that zero was less than or equal to zero, and he saw that it was good.	11	והשליליים. כנוי הוכיח שאפס הינו קטן או שוה לעצמו,
12	And the evening and the morning were the day	12	וראה כי טוב. ויהי ערב ויהי בוקר יום אפס. ביום הבא
13	of zero. On the next day, two more numbers were created, one with zero as its left set and	13	שני מספרים נוספים נבראו, אחד עם אפס בקבוצה
14	one with zero as its right set. And Conway called the former number "one," and the latter	14	השמאלית ואחד עם אפס בקבוצה הימנית. ויקרא כנוי למספר הראשון "אחד" ולשני קרא "מינוס אחד".
15	he called "minus one." And he proved that minus one is less than but not equal to zero and	15	והוכיח כי מינוס אחד הינו קטן ושונה מאפס ואפס הינו
16	zero is less than but not equal to one. And the evening...	16	קטן ושונה מאחד. ויהי ערב...

FIGURE 1

The name of the creator J.H.W.H. Conway posed a problem that Alice noticed:

> But 'Conway' ... that's not a Hebrew name. You've got to be kidding.

Actually Knuth is kidding. The name is clearly intended to suggest the name of the mathematician John Horton Conway of Cambridge University. Bill offers some help, replying to Alice's remark above:

> No, honest. Of course the old Hebrew writing doesn't show any vowels, so the real name might be Keenawu or something; maybe related to the Khans? I guess not. Since I'm translating into English, I just used an English name. Look, here are the places where it shows up on the stone. The J.H.W.H. might also stand for "Jehowah."

J.H.W.H. is obviously intended to be translated as יהוה (letters: YHWH). But there is no Hebrew word or name for Conway. However, what Bill says gives us a few clues. By removing the vowels

1	...day. And Conway said, "Let the numbers be	1	...יום. ויאמר כנוי "נוסיף את המספרים זה לזה באופן
2	added to each other in this wise: The left set of the sum of two numbers shall be the sums of all left parts of each number with the other; and in like manner the right set shall be from the right	2	הבא: הקבוצה השמאלית של סכום שני המספרים תהה הסכומים של כל החלקים השמאליים של כל מספר עם משנהו; ובצורה דומה הקבוצה הימנית תהה מהחלקים
3	parts, each according to his kind." Conway proved that every number plus zero is unchanged, and he saw that addition was good.	3	הימניים, כל אחד בהתאם למינהו". כנוי הוכיח שכל מספר שנוסף לו אפס אינו משתנה, וירא כי טובה
4	And the evening and the morning were the third day.	4	התוספת. ויהי ערב ויהי בקר יום שלישי.
5	And Conway said, "Let the negative of a number have as its sets the negatives of the number's opposite sets; and let subtraction be	5	ויאמר כנוי, "תהינה הקבוצות של מספר שלילי, השליליים של הקבוצות המוחלפות של המספר;
6	addition of the negative." And it was so. Conway	6	והפחתה תהה ההוספה של השלילי". וכן היה. ויוכיח
7	proved that subtraction was the inverse of addi-	7	כנוי שההפחתה הינה הפעולה ההפכית של ההוספה,
8	tion, and this was very good. And the evening and the morning were the fourth day.	8	וירא כי טוב. ויהי ערב ויהי בקר יום רביעי.
9	And Conway said to the numbers, "Be fruitful	9	ויאמר כנוי למספרים "כפלו ורבו. יהה חלקו של מספר
10	and multiply. Let part of one number be multiplied by another and added to the product of the first number by part of the other, and let	10	אחד מוכפל במשנהו ומוסף למכפלת המספר הראשון
11	the product of the parts be subtracted. This shall be done in all possible ways, yielding a number in the left set of the product when the parts are of the same kind, but in the right set when they	11	בחלקו של משנהו, ותהה מכפלת החלקים מופחתת. וזה הדבר יעשה בכל הדרכים האפשריות, ויניב מספר מהקבוצה השמאלית של המכפלה כאשר החלקים הינם מאותו הסוג, אבל מהקבוצה הימנית כאשר הינם מהסוג
12	are of opposite kinds." Conway proved that	12	ההפוך". כנוי הוכיח שכל מספר פעם אחד אינו משתנה.
13	every number times one is unchanged. And the evening and the morning were the fifth day.	13	ויהי ערב ויהי בקר יום חמישי.
14	And behold! When the numbers had been	14	כפתור ופרח! כאשר המספרים נבראו לאורך מספר
15	created for infinitely many days, the universe	15	ימים אינסופי היקום עצמו הופיע. ויהי ערב ויהי בקר
16	itself appeared. And the evening and the morning were א day.	16	יום א.
17	And Conway looked over all the rules he had made for numbers, and saw that they were very,	17	ויבחן כנוי את כל החוקים שיצר למספרים, וירא כי
18	very good. And he commanded them to be for signs, and series, and quotients, and roots.	18	טובים מאד, מאד. וצוה אותם לסימנים, טורים, ומנות ושורשים.
19	Then there sprang up an infinite number less	19	מזה נובע מספר אינסופי שהינו קטן מהאינסוף.
20	than infinity. And infinities of days brought forth multiple orders of infinities.	20	ואינסוף הימים יצרו רמות רבות של האינסוף.

FIGURE 2

from Keenawu, we get the consonants KNW which is either קנו (letters: QNW) or כנו (letters: Kh NW) in Hebrew; the difference is in which of the two letters, Quf or Khaph (both sounding like K), is used as the first letter. The possibility that the name KNW might be related to the name Khan (Cohen, spelled כהן) suggests that the first letter should be Khaph, and thus KNW should be כנו. By vocalizing כנו as כְּנוּי (pronounced: Kinu-i) we call to mind the biblical phrase כנוי יהוה (pronounced: Kinu-i YHWH) meaning "named YHWH." However, by vocalizing כנו as כָּנְוֵי (pronounced: Kan-vay) we get a modern Hebrew spelling for "Conway." (Also, as Julian Bigelow of the Institute for Advanced Studies has observed, another slight change to כנו gives כְּנוּת, a possible spelling of "Knuth.")

There were a number of phrases which were clearly intended to be biblical idioms. These include

In the beginning...	בראשית
Everything was void	היתה תהו ובהו
Let there be...	יהי...
And behold!	והנה or כפתור ופרח
and he saw that it was good (and variants)	וירא כי טוב
...a sign to separate...	אות המבדיל בין
and he said...	ויאמר
each according to its kind	כל אחד בהתאם למינהו
and the evening and the morning were the X day	ויהי ערב ויהי בוקר יום X
Be fruitful and multiply	פרו ורבו

There were a few places, however, where the direct biblical idiom did not fit grammatically. For example, in one place where it was written, "and he saw that it was good," וירא כי טוב, the "it" was feminine. In this case the gender of טוב was changed to get וירא כי טובה.

In a few places, the biblical idiom was involved in an English *double-entendre,* e.g., the English word had another, mathematical, meaning. The well-known phrase "be fruitful and multiply" has as its original form פרו ורבו (pronounced PRU UR'VU). Here "multiply" is taken in a sexual sense; however, a mathematical sense is implied by its appearance in the tablets. We translated the phrase as כפלו ורבו (pronounced KIFLU UR'VU) which literally means "multiply and multiply" (mathematically and sexually) but which alliterates with the original.

Another phrase, "null and void", תהו ובהו, is used to describe the state of the universe before creation. Later "void" is for the empty set. To call the set תהו ובהו would be sheer nonsense. Recognizing that the mathematics had to be as precise as in English, we simply called the empty set the "empty set" הקבוצה הריקה.

The "original" Hebrew of the Conway Stones and its English translation according to Bill are given in FIGURES 1 and 2 in the style of the Hertz Pentateuch, in two columns side-by-side, with each sentence numbered consecutively.

The authors wish to thank Professors Tamar Alexander, Arnold Band, and Herbert Davidson of the Near Eastern Languages Department at UCLA for refereeing an argument between the two authors. The authors, of course, take full responsibility for any decisions taken. The authors also wish to thank Nitza Abramovitz, Jean Campbell, Brenda Ramsey and Judy Estrin for typing portions of this paper.

References

[1] Reuben Alcalay, The Complete Hebrew-English Dictionary, The Complete English-Hebrew Dictionary, Massadah Publishing Co., Tel Aviv, 1965.

[2] Dr. H. H. Hertz, (Ed.), The Pentateuch and Haftorahs (Second Ed.), Soncino Press, London, 1964.

[3] D. E., Knuth, Surreal Numbers, Addison-Wesley, Reading, 1974.

PROBLEMS

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

Proposals

To be considered for publication, solutions should be mailed before April 1, 1977.

973. Correction to proposal appearing in March, 1976: The last hypothesis should read $N^2 \nmid b^p - 1$.

985. Correction to proposal appearing in May, 1976: The third term of the series should be $3/(k+4)!$.

988. A given equilateral triangle ABC is projected orthogonally from a given plane P to another plane P'. Show that the sum of the squares of the sides of triangle $A'B'C'$ is independent of the orientation of the triangle ABC in plane P. [*Murray S. Klamkin, University of Waterloo.*]

989. Let $r \geq 0$, $s \geq 0$, and $r + s \leq n$. Find the number of sequences of positive integers $(a_1, a_2, \ldots, a_n)$ such that for $1 \leq k \leq n$, $a_k \leq k$ where $a_k = 1$ for r values of k, and $a_k = k$ for s values of k. [*L. Carlitz and Richard Scoville, Duke University.*]

990. Prove that the identity $f(x+1)/g(x+1) - f(x)/g(x) = h(1/x)$ is not satisfied by any non-constant polynomials f, g, and h. [*Harry W. Hickey, Arlington, Virginia.*]

991. Let a and b be elements of a finite ring such that $ab^2 = b$. Prove that $bab = b$. [*F. S. Carter, Portland State University.*]

992. Call a vertex of a convex hexagon *ordinary* if it is the intersection of at least three diagonals or sides of different lengths. Otherwise, let the vertex be called *exceptional.*

(a) Prove that at least one vertex of a convex hexagon is ordinary.

(b)* What is the maximum number of exceptional vertices that a convex hexagon can have?

[*Kenneth Fogarty, Erwin Just, and Norman Schaumberger, Bronx Community College.*]

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

993. Let g be a continuous function from $[0, 1]$ to $[0, 1]$ with $g(0) = 0$. If for each x in $[0, 1]$ there is a positive integer $n(x)$ such that $g^{n(x)}(x) = x$ (the $n(x)$th iterate of g), then show that $g(x) = x$ for all x in $[0, 1]$. [*F. David Hammer, University of Illinois at Chicago Circle.*]

994. For n and m positive integers, evaluate $\int_0^1 (-1)^{[nt]}(-1)^{[mt]}dt$, where [ ] denotes the greatest integer function. [*Peter Ørno, The Ohio State University.*]

995. Call an $n \times n$ matrix ($n \geqq 2$) *R-symmetric* if the interchange of any two distinct rows yields a symmetric matrix. Find a characterization of all R-symmetric matrices. [*Edward T. H. Wang, Wilfred Laurier University, Canada.*]

Quickies

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

Q638. Let a, b, and c denote the sides of an arbitrary triangle with respective medians m_a, m_b, and m_c. Determine all integral p and q so that

$$\left(\frac{\sqrt{3}}{2}\right)^p (a^p m_a^q + b^p m_b^q + c^p m_c^q) \geqq \left(\frac{\sqrt{3}}{2}\right)^q (a^q m_a^p + b^q m_b^p + c^q m_c^p).$$

[*Murray S. Klamkin, University of Waterloo.*]

Q639. Let $k_1, k_2, \ldots, k_n$ be any set of n integers and let $m_1, m_2, \ldots, m_n$ be any permutation of this set. Then $|k_1 - m_1| + |k_2 - m_2| + \cdots + |k_n - m_n|$ is even. [*Frank Gillespie, Southwest Missouri State University.*]

Solutions

Charlemagne's Magic Squares **May 1975**

943. Early in his reign as Emperor of the West, Charlemagne ordered a pentagonal fort to be built at a strategic point of his domain. As good luck charms, he had a third order magic square with all prime elements engraved on each wall. The five magic squares were different from each other, but they had the same magic constant — the year in which the fort was completed. The fort proved its ability to resist attack midway through his reign.

On this evidence, reconstruct the magic squares. [*Charles W. Trigg, San Diego, California.*]

Solution: We begin by recalling two facts about 3×3 magic squares: (1) If S is the magic constant and A is the entry in the middle cell of the magic square, then $A = S/3$, and (2) a third order magic square is always composed of three sets each containing three numbers where the difference between

the numbers of each trio is the same (in what follows we shall represent this difference by d) and the difference between the last term of the nth trio and the first term of the $(n+1)$st trio ($n = 1, 2$) is the same (we shall represent this difference by v). [Thus d is a positive integer and v is an integer.]

Since Charlemagne reigned as Emperor of the West from 800 to 814 and the fort proved its ability to resist attack midway through his reign, $800 \leqq S \leqq 807$ so $267 \leqq S/3 \leqq 269$. But $A = S/3$ is a prime number. Thus $A = 269$ and $S = 807$. There are thirteen triples of prime numbers with 269 as one element and having a sum of 807. They are: 257, 269, 281; 227, 269, 311; 191, 269, 347; 179, 269, 359; 149, 269, 389; 137, 269, 401; 107, 269, 431; 89, 269, 449; 71, 269, 467; 59, 269, 479; 47, 269, 491; 29, 269, 509; and 17, 269, 521. Inspection of a listing of the twenty-seven distinct numbers which appears immediately above, ordered from smallest to largest, and their differences, yields the following five sets of trios of numbers:

(i) 47, 59, 71; 257, 269, 281; 467, 479, 491 ($d = 12$, $v = 186$)
(ii) 107, 149, 191; 227, 269, 311; 347, 389, 431 ($d = 42$, $v = 36$)
(iii) 71, 149, 227; 191, 269, 347; 311, 389, 467 ($d = 78$, $v = 36$)
(iv) 47, 137, 227; 179, 269, 359; 311, 401, 491 ($d = 90$, $v = -48$)
(v) 17, 137, 257; 149, 269, 389; 281, 401, 521 ($d = 120$, $v = -108$).

From these triples the following third order magic squares are obtained:

(i)
479	71	257
47	269	491
281	467	59

(ii)
389	191	227
107	269	431
311	347	149

(iii)
389	227	191
71	269	467
347	311	149

(iv)
401	227	179
47	269	491
359	311	137

(v)
401	257	149
17	269	521
389	281	137

Hence we have reconstructed Charlemagne's five magic squares. To obtain the third square, for example, this procedure was employed: (a) Enter the three given trios of numbers as rows 1, 2, and 3. (b) Interchange the entries of each pair of symmetrically located elements which is not on a diagonal. (c) Rotate each of the entries on the outer perimeter of the square one cell counterclockwise.

(a)
71	149	227
191	269	347
311	389	467

(b)
71	389	227
347	269	191
311	149	467

(c)
389	227	191
71	269	467
347	311	149

Bob Prielipp
University of Wisconsin–Oshkosh

Comment on 943. In Charlemagne's day, negative numbers were either unknown or immoral. Therefore the solutions would have to be as above. However, things have changed and the following solution might well have been placed on the pentagon walls by a recent administration:

347	−109	569
491	269	47
−31	647	191

311	−103	599
557	269	71
−61	641	227

359	−109	557
467	269	71
−19	647	179

389	−61	479
359	269	179
59	599	149

359	−31	479
389	269	149
59	569	179

Scott Smith
Bellingham, Washington

Also solved by Walter Bluger (Canada), Stephen C. Currier, Clayton W. Dodge, Timothy P. Farrell & John M. Samoylo, Ken Jackman, University of Santa Clara Problem Solving Group, Kenneth M. Wilke, and the proposer.

An Old Bridge Problem **May 1975**

944. Compute the total number of distinct auctions in contract bridge. [*Richard Johnsonbaugh, Chicago State University and R. Rangarajan, Tata Institute of Fundamental Research, independently.*]

Solution: One possible auction is 4 consecutive passes. Each other auction follows the pattern $n_1P + n_2(B + S)$, where $mX + nY$ signifies m repetitions of act X followed by n repetitions of act Y and where

B signifies a bid (e.g., 1 club, 5 no-trump),

P signifies "pass",

S signifies a legitimate combination of intervening passes, doubles (D), and redoubles (R), e.g., $2P + D$, P, $-$, $D + R + 2P$.

Depending on the number of times the bid is doubled, S may assume one of three forms:

(1) a_1P (bid not doubled),

(2) $a_2P + D + a_1P$ (bid doubled, not redoubled),

(3) $a_2P + D + a_3P + R + a_1P$ (bid doubled and redoubled).

In all of the above, a_1 can be 0, 1, or 2, whereas a_2 and a_3 are restricted to 0 and 2. (One may not double one's partner!) Thus, form 1 has 3 possibilities, form 2 has 6 ($= 2 \cdot 3$), and form 3 has 12 ($= 2 \cdot 2 \cdot 3$), for a total of 21 possibilities for S.

Since three consecutive passes are permitted at the beginning of the auction without the hand being passed out, n_1 assumes one of the four values 0, 1, 2, or 3. The total number of bids, n_2, must be between 1 and 35, inclusive, 35 being the largest number of different possible bids in a single auction.

There are $\binom{35}{j}$ combinations of j B's. Since any j bids may be ordered only one way in a single auction, and we have seen that there are 21 possibilities for each of the j S's (except the final S which has only seven, as the auction must end in three passes), it follows that the total number of auctions is

$$4\sum_{j=1}^{35}\binom{35}{j}21^{j-1}7 = \frac{4}{3}\sum_{j=1}^{35}\binom{35}{j}21^{j} = \frac{4}{3}(22^{35}-1).$$

Counting the auction with 4 consecutive passes, we get

$$\frac{4}{3}(22^{35}-1)+1$$

or

128, 745, 650, 347, 030, 683, 120, 231, 926, 111, 609, 370, 563, 122, 697, 557

auctions in all.

ALAN FRANK
Oberlin College

Also solved by J. L. Caldwell & R. L. Raymond, Steven R. Conrad, Milton Eisner, Carl P. McCarty, T. E. Moore, C. C. Oursler, Temple University Problem Solving Group, B. L. Schwartz, and the proposers.

Conrad and Moore found several references for this problem in Amer. Math. Monthly: E 801 (1948, 578), E 1409 (1960, 925), and W. E. Langlois, *The Number of Possible Auctions at Bridge*, (1962, 634).

Square in a Triangle **September 1975**

945. 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. [*Alan Wayne, Pasco-Hernando Community College, Florida.*]

Solution: Let the sides of the desired triangle be of lengths ka, kb, and kc where (a, b, c) is a primitive Pythagorean triple. Let the side of the desired square be x. We wish to minimize the area of the triangle which is $\frac{1}{2}k^2ab$. The similar triangles formed by removing the square yield the proportion $(ka - x)/x = x/(kb - x)$. Solving for x we have $x = kab/(a + b)$. Since $a + b$ is relatively prime to both a and b and x is to be an integer, $a + b$ divides k. To minimize $\frac{1}{2}k^2ab$ we take $k = a + b$. Then the area is equal to $\frac{1}{2}ab\,(a + b)^2$ which is a minimum for a and b equal to 3 and 4. Thus the sides of the desired Pythagorean triangle are of lengths 21, 28, and 35.

GLADWIN BARTEL
Otero Junior College,
LaJunta, Colorado

Also solved by Mangho Ahuja, Leon Bankoff, George Berzsenyi, Walter Bluger (Canada), Milo Bryn & Ken Yocom, Eliot William Collins, Romae Cormier, Joseph D'Mello & V. Srinivas (India), Thomas Elsner, Irwin Feinstein, Robert Fisk, Donald C. Fuller, Jack Garfunkel, Richard Gibbs, Michael Goldberg, M. G. Greening (Australia), G. A. Heuer, Lew Kowarski, Sidney Kravitz, Henry S. Lieberman, Janice A. McGoldrick, Marilyn Sieben McIntosh, Paul Moulton, R. B. Nelson & R. W. Owens, C. C. Oursler, Leonard L. Palmer, D. E. Penney, Aron Pinker, Ellis J. Rich, Harry D. Ruderman, L. Van Hamme (Belgium), Edward T. H. Wang (Canada), Martin C. Weiss, Kenneth M. Wilke, G. W. Williams, Jr., Gregory Wulczyn, Gene Zirkel, Aleksandras Zujus, and the proposer. There was one unsigned solution.

Expected Length **September 1975**

946. 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? [*M. H. Hoehn, Santa Rosa, California.*]

Solution: If we condition on the three mutually exclusive events:

$\mathcal{S}$ — both points on the same side of the square;

$\mathcal{A}$ — points on adjacent sides of the square;

$\mathcal{O}$ — points on opposite sides of the square;

we find that the expected value of the length L of the line segment joining the points is such that

$$E[L] = \frac{1}{4}E[L \mid \mathcal{S}] + \frac{1}{2}E[L \mid \mathcal{A}] + \frac{1}{4}E[L \mid \mathcal{O}].$$

Since the two points have independent uniform distribution,

$$E[L \mid \mathcal{S}] = \int_0^1 \int_0^1 |y - x|\,dx\,dy = \int_0^1 y^2\,dy = \frac{1}{3},$$

$$E[L \mid \mathcal{A}] = \int_0^1 \int_0^1 \sqrt{x^2 + y^2}\,dx\,dy = \frac{2}{3}\int_0^{\pi/4} \sec^3\theta\,d\theta$$

$$= \frac{1}{3}(\sqrt{2} + \ln(\sqrt{2} + 1)), \quad \text{and}$$

$$E[L \mid \mathcal{O}] = \int_0^1 \int_0^1 \sqrt{1 + (y - x)^2}\,dx\,dy = 2\int_0^1 \int_0^y \sqrt{1 + x^2}\,dx\,dy$$

$$= 2\int_0^1 (1 - x)(1 + x^2)^{1/2}\,dx = \sqrt{2} + \ln(\sqrt{2} + 1)$$

$$- \frac{2}{3}(2\sqrt{2} - 1).$$

Combining these conditional expectations, we find that

$$E[L] = \frac{1}{12}(3 + \sqrt{2} + 5\ln(\sqrt{2}+1)) \doteq 0.7351.$$

P. J. PEDLER
Mount Lawley College
Australia

Also solved by D. Bell, Joseph D'Mello & V. Srinivas (India), Leon Gerber, Michael Goldberg, G. A. Heuer, Paul Rogers, George Schillinger, and the proposer. One person, whose signature we could not make out, submitted a correct solution along with the computer printout of a computer program using randomly generated test pairs.

Minimum Perimeter **September 1975**

947. 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. [*Steve Moore and Mike Chamberlain, University of Santa Clara.*]

Solution: Let the x intercept of a line through (a, b) be $a + x$ and the y intercept be $b + y$. The perimeter of the right triangle is then

$$P = a + x + b + y + \sqrt{a^2 + y^2} + \sqrt{b^2 + x^2}.$$

Similar triangles imply $xy = ab$. Thus

$$P = a + b + x + \frac{ab}{x} + \left(1 + \frac{a}{x}\right)\sqrt{b^2 + x^2},$$

$$\frac{dP}{dx} = \frac{1}{x^2\sqrt{b^2 + x^2}}[x^3 - ab^2 + (x^2 - ab)\sqrt{b^2 + x^2}],$$

$$\frac{d^2P}{dx^2} = bx^{-3}[b(b^2 + x^2)^{-3/2}(x^3 + 2ab^2 + 3ax^2) + 2a].$$

For $x > 0$,

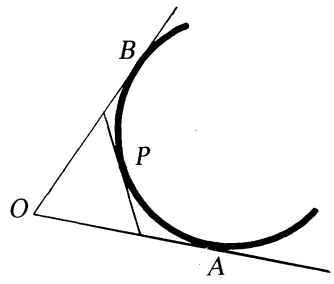
$$\frac{d^2P}{dx^2} > 0 \quad \text{and} \quad \frac{dP}{dx} = 0 \quad \text{iff} \quad x = \frac{b(a + \sqrt{2ab})}{b + \sqrt{2ab}}.$$

Thus the minimum perimeter occurs for this x and the equation of the line is $(b + \sqrt{2ab})x + (a + \sqrt{2ab})y = \sqrt{2ab}(a + b + \sqrt{2ab})$.

JOHN OMAN
Oshkosh, Wisconsin

Editor's comment. The solution above is typical of all the correct solutions using calculus. M. S. Klamkin sent us a comment along with a reference (Problem E1008, Amer. Math. Monthly, 1952, page 639) to the allied problem of determining a triangle of given perimeter for a general angle. We expand Klamkin's comment as follows:

Let the given point be P and construct the larger circle which contains P and is tangent to the coordinate axes at A and B as in the figure. We see that the perimeter of the triangle formed by the tangent to the circle at P is equal to $\overline{OA} + \overline{OB}$. This triangle has the minimal perimeter since any other line through P will have a corresponding tangent circle associated with it for which the sum $\overline{OA'} + \overline{OB'}$ is larger.



Also solved by Mangho Ahuja, Leon Bankoff, Gerald Bergum & Ken Yocom, Bern Problem Solving Group (Switzerland), Walter Bluger (Canada), Robert X. Brennan, Eliot W. Collins, Romae J. Cormier, Ragnar Dybvik (Norway), Thomas E. Elsner, Irwin K. Feinstein, Donald C. Fuller, Richard A. Gibbs, Michael Goldberg, M. S. Klamkin (Canada), Lew Kowarski, Henry S. Lieberman, Leonard L. Palmer, D. E. Penney, Gary D. Peterson, Aron Pinker, Benjamin L. Schwartz, J. M. Stark, John A. Tierney, Kenneth M. Wilke, Gregory Wulczyn, Atila Yanik, Aleksandras Zujus and the proposers.

Complete Residue Systems **September 1975**

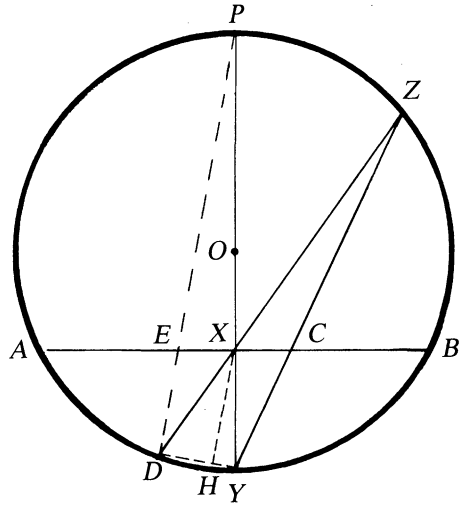
948. Let Z_n be the ring of integers modulo n. For what values of n different from 2 do there exist permutations f and g on Z_n such that the pointwise product fg is also a permutation on Z_n? [*Bob Prielipp and N. J. Kuenzi, Oshkosh, Wisconsin.*]

Editor's comment. No such permutations exist for $n > 2$. This was proven by S. Chowla and T. Vijayaraghavan, On complete residue systems, Quart. J. Math., vol. 19, pp. 193–194. For a shorter proof see W. J. Coles and F. R. Olson, *A note on complete residue systems*, Amer. Math. Monthly, 1954, page 622.

Also solved by Bern Problem Solving Group (Switzerland), R. J. Cormier & R. B. Eggleton, Joseph D'Mello & V. Srinivas (India), Temple University Problem Solving Group, and Edward T. Wong.

Another Butterfly Problem **September 1975**

949. In a circle with center O, OXY is perpendicular to chord AB (as shown). Prove $DX \leqq CY$. [*P. Erdös, Hungarian Academy of Science, and M. S. Klamkin, University of Waterloo.*]



Solution I: Draw DY and choose H on DY so that XH is perpendicular to DY. We have that $\angle XDH = \angle ZDY = \pi/2 - \angle XYZ$ since the intercepted arcs form a semi-circle. Thus, right triangle XDH is similar to right triangle XCY and so $XH : DX = XY : CY$. Since $XH \leqq XY$, $DX \leqq CY$.

MARK KLEIMAN, Student
Stuyvesant High School
New York, N.Y.

Solution II: Let the end of the diameter be P and let PD intersect AB at E. Then, $EX = XC$ by the "butterfly problem". (See, for example, Steven R. Conrad, *Another simple solution to the butterfly problem*, this MAGAZINE, 46 (1973) 278–280.) Applying the law of sines to triangles DEX and YXC to obtain

$$\frac{DX}{\sin DEX} = \frac{EX}{\sin D} = \frac{XC}{\sin Y} = \frac{CY}{\sin YXC},$$

we find $DX = CY \sin DEX \leqq CY$.

DONALD BATMAN
Socorro, New Mexico

Also solved by Mango Ahuja, Leon Bankoff, Bern Problem Solving Group (Switzerland), Walter Bluger (Canada), Charles Chouteau, Romae J. Cormier, Joseph D'Mello & V. Srinivas (India), Hüseyin Demir (Turkey), Mark A. Flood, Donald C. Fuller, Leon Gerber, Michael Goldberg, Leonard D. Goldstone, M. G. Greening (Australia), N. G. Gunderson, Carolyn Hazel, Václav Konečný (Czechoslovakia), Lew Kowarski, Graham Lord (Canada), John Oman, Adam Riese, Harry D. Ruderman, Léo Sauvé (Canada), J. M. Stark, A. W. Walker (Canada), Kenneth M. Wilke, Ken Yocom, and the proposers. M. S. Klamkin showed that CY/DX is an increasing function of XC.

This problem was also proposed in the May 1975 issue of the Ontario Secondary School Mathematics Bulletin and the October 1975 issue of Eureka, a publication of the Carleton-Ottawa Mathematics Association. The progenitor of the problem in all cases was Paul Erdös.

Answers

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

Q638. It is known that the medians m_a, m_b, m_c form a triangle with respective medians $3a/4$, $3b/4$, $3c/4$. Consequently for any side-median inequality in the terms a, b, c, m_a, m_b, m_c we have a dual median-side inequality in the terms $m_a, m_b, m_c, 3a/4, 3b/4, 3c/4$. Dualizing the inequality of the problem merely reverses the inequality sign, producing equality. Only $p = q$ yields this identity. (Note that p and q need not be integers.)

Q639. If a is any integer, $(-1)^{|a|} = (-1)^a$. Hence,

$$(-1)^{\Sigma|k_i - m_i|} = \Pi(-1)^{|k_i - m_i|} = \Pi(-1)^{(k_i - m_i)} = (-1)^{\Sigma(k_i - m_i)} = (-1)^0 = 1.$$

Thus, $\Sigma|k_i - m_i|$ must be even.

NEWS & LETTERS

SOLUTION OF THE FOUR COLOR PROBLEM

The four color conjecture, one of the most popular and appealing unsolved problems of mathematics, was verified this summer by an intricate computer-based analysis carried out by Kenneth Appel and Wolfgang Haken of the University of Illinois. While it may take a year or longer for others to verify every detail of their work--the proof contains several hundred pages of what even the authors term "ridiculous detail" and subsumes over 1000 hours of computer calculation--the general outline of their method and the initial confirmation of their major calculations is accepted by most graph theorists as complete and correct.

The conjecture was first posed in 1853 by Francis Guthrie, a mathematics student at University College, London: can every map in the plane be colored with four colors so that adjacent regions receive different colors? The first "proof" was published in 1879 by A.B. Kempe, but it proved to be incorrect. Appel's and Haken's new proof is simply a very elaborate correction of Kempe's oversight.

Kempe began by showing, correctly, that it suffices to verify the conjecture for "normal" maps in which precisely three regional boundaries meet at each vertex: vertices where more than three boundaries meet can be separated into several trivalent vertices, and the resulting map will be more difficult to color than was the original because more regions are adjacent. He then used Euler's formula $V - E + F = 2$ relating the vertices, boundaries (or edges) and regions (or faces) of a map to show that any normal map must contain regions with fewer than six neighbors.

This isn't too hard. Let e_i denote the number of edges of region i. Then the total number E of edges is precisely $\frac{1}{2}\Sigma e_i$. Moreover, in a normal map, $2E = 3V$, where V denotes the total number of vertices. Hence

$$\Sigma(6-e_i) = 6F - 2E = 6F - 6E + 4E$$
$$= 6F - 6E + 6V = 12.$$

If each e_i were six or greater, the left side would be zero or less. Hence some region must have fewer than six neighbors.

Armed with this information concerning "unavoidable" regions, Kempe tried to show that whenever a region with fewer than six sides appeared in a normal map, the map could be reduced to a smaller one whose coloring was no easier if the reduced map could be four-colored then so could the original map. His proof was simple and correct for regions of 2, 3 or 4 sides, but for pentagonal regions his proof was incomplete: in 1890 P.J. Heawood gave an example of a normal map with 25 regions that contained a pentagon that could not be reduced according to the methods used in Kempe's proof. Heawood's map could, of course, be colored, but not by Kempe's method.

Heawood's analysis of the Kempe proof demonstrated that the problem was far more subtle than had at first been believed. The problem attracted the attention of amateur and professional mathematicians throughout the world. But not even the best minds of the twentieth century could solve it. Laborious modifications of Kempe's reducibility arguments revealed only that maps of size no larger than 40 or so must be four-colorable. These efforts climaxed, in a way, with Martin Gardner's publication of a hoax counter-example in the April 1, 1975 issue of *Scientific American*.

Coloring problems are at present customarily translated into graphs by duality: each region is replaced by a

point in it (the capital of the country, to speak geographically) and each boundary by a line joining two points. Then a normal map becomes a triangulated graph (one whose faces are all triangles), and the Euler formula for graphs leads to various Kempe-type reduction arguments.

One particularly convenient way to work with the Euler formula on triangulated graphs is to assign a "charge" to each vertex v of $6 - n(v)$, where $n(v)$ is the number of edges that meet at vertex v. Then the Euler formula implies, as above, that the sum of these charges for any triangulated graph is 12. Since positive charge occurs only on vertices of degree less than 6 (i.e., in regions with fewer than six neighbors), we can be sure--from the positive total--that such vertices are unavoidable in triangulated graphs. Kempe tried unsuccessfully to show that this set of unavoidable configurations was also reducible; where Kempe failed, Haken and Appel succeeded. But to do so they had to replace his one flawed case (pentagonal regions) with 1,936 complex configurations.

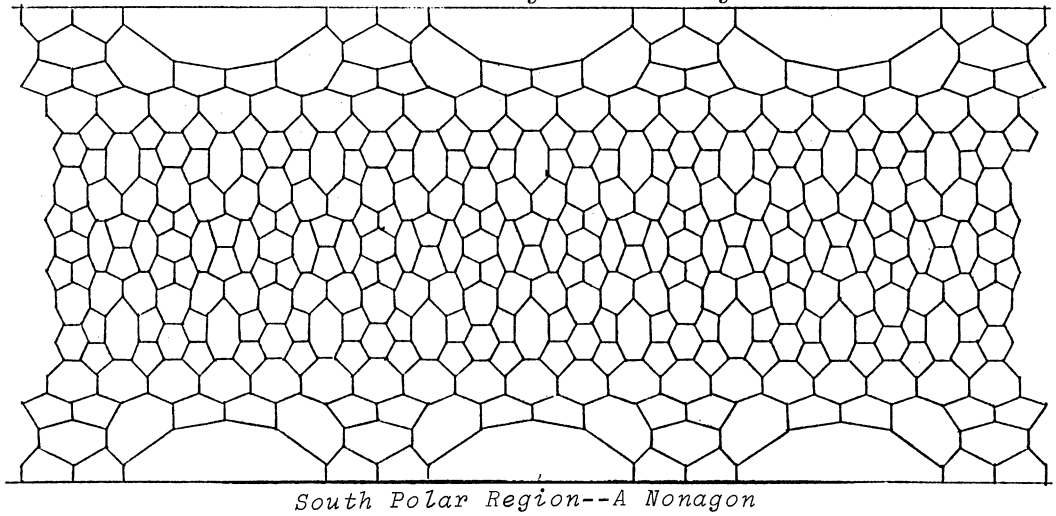
The strategy behind the Appel-Haken proof is to redistribute the Eulerian charge on the graph using a specially designed "discharging algorithm" in such a way as to minimize the number of positive vertices--always subject to the constraint that the total charge must remain unchanged. Then each of the configurations identified by the remaining positive charges must be reduced, just as Kempe had reduced the cases of 2, 3 and 4 sided regions.

The first step in this program--the design of an appropriate discharging algorithm--took about three and one-half years of counterpoint between man and machine. For each vertex with positive charge produced by the draft algorithm, Haken and Appel tried to find a reducible configuration around it (using specially designed computer algorithms for assistance). If no reducible configuration was discovered in reasonable time--perhaps half an hour of computer search--they assumed that none existed and went back to modify the discharging algorithm to avoid such situations.

When they were finally convinced that they had an algorithm that located only reducible configurations, Appel and Haken began the systematic verification of reducibility for all cases of positive charge produced by the algorithm. The resulting catalogue contained 1,936 reducible configurations, each requiring a search of up to 500,000 options to verify reducibility. This last phase of the work took six months, and was completed in June, 1976.

Final checking--part of which was carried out by the researchers' teen-age children--took the entire month of July,

North Polar Region--A Nonagon



South Polar Region--A Nonagon

and the results were communicated to the *Bulletin of the American Mathematical Society* on July 26, 1976: "Every Planar Map Is Four Colorable." This official announcement is scheduled to appear in the September 1976 issue of the *Bulletin*.

The Haken-Appel proof has sent several different shock waves through the mathematical world. The verification of a century-old conjecture that had baffled the twentieth century's best mathematicians is an astounding accomplishment. But a solution based on computerized case analyses involving nearly 2000 cases and 10 billion logical options is the complete antithesis of the idealized "elegant" mathematical proof. (Perhaps this is why Haken's personal presentation of the result to an audience of several hundred mathematicians at the University of Toronto in August was greeted with no more than mildly polite applause.)

The Haken-Appel proof is the first example of a major mathematical problem solved by an essential symbiosis of theoretician and computer. Many mathematicians feel that this result is only the prelude to better, shorter, more conceptual proofs. "We aren't going to go through eternity," vowed one mathematician, "saying 'And the computer said...'."

In defense of their formidable method, Haken and Appel observe that theirs is close to an optimal proof within the Kempe tradition of seeking unavoidable sets of reducible configurations. Every configuration that must be reduced is surrounded by a ring of neighbors that determine its reducibility. The size of this ring has great bearing on the difficulty of establishing reducibility.

Several years ago Edward F. Moore of the University of Wisconsin developed a strategy for disproving the conjecture (if indeed it were false) by creating maps that exclude all known reducible configurations almost as fast as such configurations were discovered. The map on p. 220 (flattened out from the surface of a sphere: each polar region bounded by the top and bottom lines is a nonagon) is one Moore created in 1963 that contains no reducible regions whose ring size is smaller than 12. The Haken-Appel proof requires configurations of ring size no larger than 14.

The Moore graph, therefore, shows that no proof based on an unavoidable set of reducible configurations can be even moderately short, since it must deal with ring sizes at least as large as 12; the Haken-Appel proof, while somewhat longer than necessary, deals with configurations only two sizes larger.

The fact that the Haken-Appel proof appeared between ring sizes 12 and 14 confirms some general probability estimates concerning the likely occurrence of reducible configurations in randomly drawn plane maps: it is quite likely that maps do exist that contain no reducible configurations of ring size smaller than 13, but it is also very likely that every map contains a reducible configuration of ring size not exceeding 14. These estimates establish upper and lower bounds on the length of any Kempe-like proof of the four color problem. Haken and Appel's proof fits right in between these theoretical limits.

Speculators on the market of mathematical problems might be inclined now to support computer attacks on all famous unsolved problems. But the crucial first step in any computer attack is a difficult theoretical maneuver --the reduction from an infinite to a finite number of cases. This is possible in the four color problem because of the intricate geometry of maps: the behavior of graphs of size n very strongly influences the behavior of graphs of size $n + 1$. With luck and insight, it is possible to develop a finite number of cases that cover all infinitely many possible maps.

Such is not likely to be the case with problems in number theory such as Fermat's conjecture, for the behavior of prime numbers appears to be much more loosely knit than is the geometry of graphs. So the finitization of problems in number theory will either be very difficult or perhaps impossible. And no computer assault can work until the finitization theory is complete.

But even if a problem is finite, it may be impractical to implement on even the fastest computers. Had the Haken-Appel attack turned out to require configurations of ring size 15, the time required for computer search of reducibility would have made the present proof totally impossible.

The Haken-Appel result points, therefore, to the existence of a new class of mathematical theorems that are true, but for which no simple proof exists. Exploration of this realm by mathematician-computer teamwork is, according to Haken and Appel, a major challenge for mathematicians in the final quarter of this century. "This work has changed my view of what mathematics is," said Haken. "I hope it will do the same for others."

--Lynn Arthur Steen

LESTER R. FORD AWARDS

Authors of six expository papers appearing in the 1975 issue of the *American Mathematical Monthly* and *Mathematics Magazine* received Lester R. Ford Awards at the August meeting of the Mathematical Association of America at the University of Toronto. Each award is in the amount of $100. The award winning papers are:

M.L. Balinski and H.P. Young, The Quota Method of Apportionment, *Amer. Math. Monthly*, 82 (1975) 701-730.

E.A. Bender and J.R. Goldman, On the Applications of Mobius Inversion in Combinatorial Analysis, *Amer. Math. Monthly*, 82 (1975) 789-803.

Branko Grunbaum, Venn Diagrams and Independent Families of Sets, *Mathematics Magazine*, 48 (1975) 12-23.

J.E. Humphreys, Representations of LS(2,p), *Amer. Math. Monthly*, 82 (1975) 21-39.

J.B. Keller and D.W. McLaughlin, The Feynman Integral, *Amer. Math. Monthly*, 82 (1975) 451-465.

J.J. Price, Topics in Orthogonal Functions, *Amer. Math. Monthly*, 82 (1975) 594-609.

STATISTICS LECTURERS

The Visiting Lecturer Program in Statistics, now its fourteenth year, attempts to provide information to students and college faculty about the nature and scope of modern statistics, and to provide advice about careers, graduate study, and college curricula in statistics. Leading teachers and research workers in statistics--from universities, industry and government--have agreed to participate as lecturers. Lecture topics include subjects in experimental and theoretical statistics, as well as in such related areas as probability theory, information theory and stochastic models in the physical, biological and social sciences.

The Visiting Lecturer Program is sponsored by the American Statistical Association, the Biometric Society and the Institute of Mathematical Statistics. Partial support is also provided by International Business Machines Corporation. The program is available to colleges and universities in the United States and Canada. Inquiries should be addressed to H.T. David, Visiting Lecturer Program in Statistics, Department of Statistics, Iowa State University, Ames, Iowa 50011.

SCHOOL MATHEMATICS COMPETITIONS

The New York City Interscholastic Mathematics League currently holds five contests each semester. Although official entrance into the NYCIML is limited to secondary schools in New York City, the league welcomes unofficial entry by schools from outside New York City. The dues are $15 per team per semester.

Any school interested in joining the league on an unofficial basis, or in using its problems as the basis for their own minileague, may do so by sending a check for one full year's dues. Those who wish to receive one copy of each of the contests on a regular basis can do so for a $5 league fee, payable in advance.

Further details can be obtained from Steven R. Conrad, President, NYCIML, 39 Arrow Street, Selden, New York 11784.

PERMUTATION NUMBERS

The question discussed by Jeffrey Jaffe in "Permutation Numbers" (this *Magazine*, March 1976, pp. 80-84) may be traced back to the year 1894 through an article by Victor Thebault (Sur un nouveau theorem d'arithmetique, *C. R. Acad. Sci.*, Paris, Ser. A-B 213 (1941) 967-970; MR 5: 141) although previous discussions appear to neglect the question of the necessity of the condition that the sum of the digits of n be less than b-1.

On the other hand, Jaffe's proof of this fact is far from convincing. To simply say "Similarly ..." after discussing an easy special case is not very satisfying. In fact the complete story on the failure of An to be a permutation number can be told based on Jaffe's method.

I will survey the conclusions when $(m,b-1) = 1$, and $g > m$ using the notation of "Permutation Numbers": (i) if $g = 2m$, then 0 is omitted and $(b-1)$ occurs twice. (ii) otherwise $b-m-1$ will be omitted and $b-m$ repeated and (a) if $q > 2m$, $q-2m$ is omitted and $q-2m-1$ repeated or (b) if $q < 2m$, $q-2m+b$ is omitted and $q-2m+b-1$ repeated. Thus at most 2 digits are omitted, and no digit occurs more than 3 times. In fact, if $q = m+1$, then $b-m$ will occur 3 times; otherwise, no digit occurs more than twice.

Richard T. Bumby
Rutgers University
New Brunswick
New Jersey 08903

1976 INTERNATIONAL MATHEMATICAL OLYMPIAD

The 18th International Mathematical Olympiad, which took place in Leinz, Austria, on July 12 and 13, 1976, consisted of the following six problems. The United States team, coached by Samuel Greitzer and Murray Klamkin, finished third behind Hungary and England.

1. In a plane convex quadrilateral of area 32 cm^2 the sum of the lengths of two opposite sides and one diagonal is equal to 16 cm. Determine all possible lengths of the other diagonal.

2. Let $P_1(x) = x^2 - 2$ and $P_j(x) = P_1(P_{j-1}(x))$ for $j = 2, 3, \ldots$. Show that for any positive integer n, the roots of the equation $P_n(x) = x$ are all real and distinct.

3. A rectangular box can be filled completely with unit cubes. If one places as many cubes as possible, each with volume 2, in the box so that their edges are parallel to the edges of the box, one can fill exactly 40% of the box. Determine the dimensions of all such boxes.

4. Determine, with proof, the largest number which is the product of positive integers whose sum is 1976.

5. Consider the system of p equations in q unknowns, where $q = 2p$,

$$a_{11}x_1 + \ldots + a_{1q}x_q = 0$$
$$a_{21}x_1 + \ldots + a_{2q}x_q = 0$$
$$\ldots$$
$$a_{p1}x_1 + \ldots + a_{pq}x_q = 0$$

with every coefficient a_{ij} a member of the set $\{-1,0,+1\}$. Prove that there exists a solution $(x_1,\ldots,x_q)$ of the system such that

(a) all $x_j (j = 1,\ldots,q)$ are integers;

(b) there is at least one value of j for which $x_j \neq 0$;

(c) $|x_j| \leq q (j = 1,\ldots,q)$.

6. A sequence $\{u_n\}$ is defined by

$$u_0 = 2,\ u_1 = 5/2,$$
$$u_{n+1} = u_n(u_{n-1}^2 - 2) - u_1$$

for $n = 1,2,\ldots$. Prove that for positive integral n

$$[u_n] = 2^{(2^n-(-1)^n)/3}$$

where $[x]$ denotes the greatest intege less than or equal to x.

THE CARUS MATHEMATICAL MONOGRAPHS

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

These numbers are currently available:

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

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

Orders should be sent to:

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

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