MATHEMATICS MAGAZINE

ŒUVRES

MATHÉMATIQUES

D'ÉVARISTE GALOIS,

PUBLIÉES

SOUS LES AUSPICES DE LA SOCIÉTÉ MATHÉMATIQUE DE FRANCE,

AVEC UNE INTRODUCTION

PAR

M. ÉMILE PICARD,

MEMBRE DE L'INSTITUT.

PARIS,

GAUTHIER-VILLARS ET FILS, IMPRIMEURS-LIBRAIRES

DU BUREAU DES LONGITUDES, DE L'ÉCOLE POLYTECHNIQUE,

Quai des Grands-Augustins, 55.

1897

(Tous droits réservés.)

Vol. 59 No. 4

October, 1986

EVOLUTION OF GROUP THEORY • THE LAPLACIAN

RECIPROCITY LAWS • GENERALIZED INVERSES

MATHEMATICS MAGAZINE

Volume 59, Number 4, October 1986
(ISSN—0025-570X)

COVER: Title page of Galois' *Oeuvres Mathématiques*, a collection that runs to a total of 61 pages!

ARTICLE

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

EDITORIAL POLICY

The aim of *Mathematics Magazine* is to provide lively and appealing mathematical exposition. This is not a research journal and, in general, the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for an article for the *Magazine*. Articles should include examples, applications, historical background, and illustrations, where appropriate. They should be attractive and accessible to undergraduates and would, ideally, be helpful in supplementing undergraduate courses or in stimulating student investigations. Articles on pedagogy alone, unaccompanied by interesting mathematics, are not suitable. Neither are articles consisting mainly of computer programs unless these are essential to the presentation of some good mathematics. Manuscripts on history are especially welcome, as are those showing relationships between various branches of mathematics and between mathematics and other disciplines.

In addition to articles and notes the *Magazine* solicits proofs without words, mathematical verse, anecdotes, cartoons, and other such material consistent with the level and aims described above. Letters and comments are also welcome.

The full statement of editorial policy appears in this *Magazine*, Vol. 54, pp. 44-45, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, nor published by another journal or publisher.

Send new manuscripts to: G. L. Alexanderson, Editor, Mathematics Magazine, Santa Clara University, Santa Clara, CA 95053. Manuscripts should be typewritten and double spaced and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit the original and one copy and keep one copy. Illustrations should be carefully prepared on separate sheets in black ink, the original without lettering and two copies with lettering added.

AUTHOR

Israel Kleiner ("The Evolution of Group Theory: A Brief Survey") received his Ph.D. in ring theory under J. Lambek at McGill University, and has been at York University since 1965. He has been involved in preservice and in-service teacher education at the undergraduate and graduate levels, among other things. One of his major interests during the past ten years has been in the history of mathematics and its uses in the teaching of mathematics.

PHOTO CREDITS

p. 197, G. L. Alexanderson; p. 199, Polya Collection; p. 200, Springer-Verlag, New York; p. 203, Konrad Jacobs, Erlangen; p. 206, Springer-Verlag, New York; p. 208, G. L. Alexanderson; p. 209, Polya Collection; p. 211, Israel Kleiner; p. 233, Mrs. W. L. Hart.

The *Mathematics Magazine* (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D. C. 20036 and Montpelier, VT, bimonthly except July/August.

The annual subscription price for the MATHEMATICS MAGAZINE to an individual member of the Association is $11 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $22. Student, unemployed and emeritus members receive a 50% discount; new members receive a 30% dues discount for the first two years of membership.) The non-member/library subscription price is $28 per year. Bulk subscriptions (5 or more copies) are available to colleges and universities for classroom distribution to undergraduate students at a 42% discount ($6.50 per copy—minimum order $32.50).

Subscription correspondence and notice of change of address should be sent to the Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D. C. 20036. Back issues may be purchased, when in print, from P. and H. Bliss Company, Middletown, CT 06457. Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

Advertising correspondence should be addressed to Ms. Elaine Pedreira, Advertising Manager, The Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D. C. 20036.

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

Postmaster: Send address changes to Mathematics Magazine, Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D. C. 20077-9564.

PRINTED IN THE UNITED STATES OF AMERICA

The Evolution of Group Theory: A Brief Survey

ISRAEL KLEINER
York University
North York, Ontario, Canada M3J 1P3

This article gives a brief sketch of the evolution of group theory. It derives from a firm conviction that the history of mathematics can be a useful and important integrating component in the teaching of mathematics. This is not the place to elaborate on the role of history in teaching, other than perhaps to give one relevant quotation:

> Although the study of the history of mathematics has an intrinsic appeal of its own, its chief raison d'être is surely the illumination of mathematics itself. For example the gradual unfolding of the integral concept from the volume computations of Archimedes to the intuitive integrals of Newton and Leibniz and finally to the definitions of Cauchy, Riemann and Lebesgue—cannot fail to promote a more mature appreciation of modern theories of integration.
>
> —C. H. Edwards [**11**]

The presentation in one article of the evolution of so vast a subject as group theory necessitated severe selectivity and brevity. It also required omission of the broader contexts in which group theory evolved, such as wider currents in abstract algebra, and in mathematics as a whole. (We will note *some* of these interconnections shortly.) We trust that enough of the essence and main lines of development in the evolution of group theory have been retained to provide a useful beginning from which the reader can branch out in various directions. For this the list of references will prove useful.

The reader will find in this article an outline of the origins of the main concepts, results, and theories discussed in a beginning course on group theory. These include, for example, the concepts of (abstract) group, normal subgroup, quotient group, simple group, free group, isomorphism, homomorphism, automorphism, composition series, direct product; the theorems of J. L. Lagrange, A.-L. Cauchy, A. Cayley, C. Jordan-O. Hölder; the theories of permutation groups and of abelian groups. At the same time we have tried to balance the technical aspects with background information and interpretation.

Our survey of the evolution of group theory will be given in several stages, as follows:

1. Sources of group theory.
2. Development of "specialized" theories of groups.
3. Emergence of abstraction in group theory.
4. Consolidation of the abstract group *concept*; dawn of abstract group *theory*.
5. Divergence of developments in group theory.

Before dealing with each stage in turn, we wish to mention the context within mathematics as a whole, and within algebra in particular, in which group theory developed. Although our "story" concerning the evolution of group theory begins in 1770 and extends to the 20th century, the major developments occurred in the 19th century. Some of the general mathematical features of that century which had a bearing on the evolution of group theory are: (a) an increased concern for rigor; (b) the emergence of abstraction; (c) the rebirth of the axiomatic method; (d) the view of mathematics as a human activity, possible without reference to, or motivation from, physical situations. Each of these items deserves extensive elaboration, but this would go beyond the objectives (and size) of this paper.

Up to about the end of the 18th century, algebra consisted (in large part) of the study of solutions of polynomial equations. In the 20th century, algebra became a study of abstract, axiomatic systems. The transition from the so-called classical algebra of polynomial equations to the so-called modern algebra of axiomatic systems occurred in the 19th century. In addition to group theory, there emerged the structures of commutative rings, fields, noncommutative rings, and vector spaces. These developed alongside, and sometimes in conjunction with, group theory. Thus Galois theory involved both groups and fields; algebraic number theory contained elements of group theory in addition to commutative ring theory and field theory; group representation theory was a mix of group theory, noncommutative algebra, and linear algebra.

1. Sources of group theory

There are four major sources in the evolution of group theory. They are (with the names of the originators and dates of origin):

(a) Classical algebra (J. L. Lagrange, 1770)
(b) Number theory (C. F. Gauss, 1801)
(c) Geometry (F. Klein, 1874)
(d) Analysis (S. Lie, 1874; H. Poincaré and F. Klein, 1876)

We deal with each in turn.

(a) *Classical Algebra* (J. L. Lagrange, 1770)

The major problems in algebra at the time (1770) that Lagrange wrote his fundamental memoir "Réflexions sur la résolution algébrique des équations" concerned polynomial equations. There were "theoretical" questions dealing with the existence and nature of the roots (e.g., Does every equation have a root? How many roots are there? Are they real, complex, positive, negative?), and "practical" questions dealing with methods for finding the roots. In the latter instance there were exact methods and approximate methods. In what follows we mention exact methods.

The Babylonians knew how to solve quadratic equations (essentially by the method of completing the square) around 1600 B.C. Algebraic methods for solving the cubic and the quartic were given around 1540. One of the major problems for the next two centuries was the algebraic solution of the quintic. This is the task Lagrange set for himself in his paper of 1770.

In his paper Lagrange first analyzes the various known methods (devised by F. Viète, R. Descartes, L. Euler, and E. Bézout) for solving cubic and quartic equations. He shows that the common feature of these methods is the reduction of such equations to auxiliary equations—the so-called resolvent equations. The latter are one degree lower than the original equations. Next Lagrange attempts a similar analysis of polynomial equations of arbitrary degree n. With each such equation he associates a "resolvent equation" as follows: let $f(x)$ be the original equation, with roots $x_1, x_2, \ldots, x_n$. Pick a rational function $R(x_1, x_2, \ldots, x_n)$ of the roots and coefficients of $f(x)$. (Lagrange describes methods for doing so.) Consider the different values which $R(x_1, x_2, \ldots, x_n)$ assumes under all the $n!$ permutations of the roots $x_1, x_2, \ldots, x_n$ of $f(x)$. If these are denoted by $y_1, y_2, \ldots, y_k$, then the resolvent equation is given by $g(x) = (x - y_1) \cdot (x - y_2) \cdots (x - y_k)$. (Lagrange shows that k divides $n!$—the source of what we call Lagrange's theorem in group theory.) For example, if $f(x)$ is a quartic with roots x_1, x_2, x_3, x_4, then $R(x_1, x_2, x_3, x_4)$ may be taken to be $x_1 x_2 + x_3 x_4$, and this function assumes three distinct values under the 24 permutations of x_1, x_2, x_3, x_4. Thus the resolvent equation of a quartic is a cubic. However, in carrying over this analysis to the quintic, he finds that the resolvent equation is of degree six!

Although Lagrange did not succeed in resolving the problem of the algebraic solvability of the quintic, his work was a milestone. It was the first time that an association was made between the

Evariste Galois

solutions of a polynomial equation and the permutations of its roots. In fact, the study of the permutations of the roots of an equation was a cornerstone of Lagrange's general theory of algebraic equations. This, he speculated, formed "the true principles for the solution of equations." (He was, of course, vindicated in this by E. Galois.) Although Lagrange speaks of permutations without considering a "calculus" of permutations (e.g., there is no consideration of their composition or closure), it can be said that the germ of the group concept (as a group of permutations) is present in his work. For details see **[12]**, **[16]**, **[19]**, **[25]**, **[33]**.

(b) *Number Theory* (C. F. Gauss, 1801)

In the *Disquisitiones Arithmeticae* of 1801 Gauss summarized and unified much of the number theory that preceded him. The work also suggested new directions which kept mathematicians occupied for the entire century. As for its impact on group theory, the *Disquisitiones* may be said to have initiated the theory of finite abelian groups. In fact, Gauss established many of the significant properties of these groups without using any of the terminology of group theory. The groups appear in four different guises: the additive group of integers modulo m, the multiplicative group of integers relatively prime to m, modulo m, the group of equivalence classes of binary quadratic forms, and the group of nth roots of unity. And though these examples appear in number-theoretic contexts, it is as abelian groups that Gauss treats them, using what are clear prototypes of modern algebraic proofs.

For example, considering the nonzero integers modulo p(p a prime), Gauss shows that they are all powers of a single element; i.e., that the group Z_p^* of such integers is cyclic. Moreover, he determines the number of generators of this group (he shows that it is equal to $\phi(p-1)$, where ϕ is Euler's ϕ-function). Given any element of Z_p^* he defines the order of the element (without using the terminology) and shows that the order of an element is a divisor of $p-1$. He then uses this result to prove *P*. Fermat's "little theorem," namely, that $a^{p-1} \equiv 1 \bmod p$ if p does not divide a, thus employing group-theoretic ideas to prove number-theoretic results. Next he shows that if t is a positive integer which divides $p-1$, then there exists an element in Z_p^* whose order is t—essentially the converse of Lagrange's theorem for cyclic groups.

Concerning the nth roots of 1 (which he considers in connection with the cyclotomic equation), he shows that they too form a cyclic group. In connection with this group he raises and answers many of the same questions he raised and answered in the case of Z_p^*.

The problem of representing integers by binary quadratic forms goes back to Fermat in the early 17th century. (Recall his theorem that every prime of the form $4n+1$ can be represented as a sum of two squares x^2+y^2.) Gauss devotes a large part of the *Disquisitiones* to an exhaustive study of binary quadratic forms and the representation of integers by such forms. (A **binary quadratic form** is an expression of the form $ax^2+bxy+cy^2$, with a, b, c integers.) He defines a composition on such forms, and remarks that if K and K^1 are two such forms one may denote their composition by $K+K^1$. He then shows that this composition is associative and commutative, that there exists an identity, and that each form has an inverse, thus verifying all the properties of an abelian group.

Despite these remarkable insights one should not infer that Gauss had the concept of an abstract group, or even of a finite abelian group. Although the arguments in the *Disquisitiones* are quite general, each of the various types of "groups" he considers is dealt with separately—there is no unifying group-theoretic method which he applies to all cases. For details see [**5**], [**9**], [**25**], [**30**], [**33**].

(c) *Geometry* (F. Klein, 1872)

We are referring here to Klein's famous and influential (but see [**18**]) lecture entitled "A Comparative Review of Recent Researches in Geometry," which he delivered in 1872 on the occasion of his admission to the faculty of the University of Erlangen. The aim of this so-called Erlangen Program was the classification of geometry as the study of invariants under various groups of transformations. Here there appear groups such as the projective group, the group of rigid motions, the group of similarities, the hyperbolic group, the elliptic groups, as well as the geometries associated with them. (The affine group was not mentioned by Klein.) Now for some background leading to Klein's Erlangen Program.

The 19th century witnessed an explosive growth in geometry, both in scope and in depth. New geometries emerged: projective geometry, noneuclidean geometries, differential geometry, algebraic geometry, n-dimensional geometry, and Grassmann's geometry of extension. Various geometric methods competed for supremacy: the synthetic versus the analytic, the metric versus the projective. At mid-century, a major problem had arisen, namely, the classification of the relations and inner connections among the different geometries and geometric methods. This gave rise to the study of "geometric relations," focusing on the study of properties of figures invariant under transformations. Soon the focus shifted to a study of the transformations themselves. Thus the study of the geometric relations of figures became the study of the associated transformations. Various types of transformations (e.g., collineations, circular transformations, inversive transformations, affinities) became the objects of specialized studies. Subsequently, the logical connections among transformations were investigated, and this led to the problem of classifying transformations and eventually to Klein's group-theoretic synthesis of geometry.

Klein's use of groups in geometry was the final stage in bringing order to geometry. An intermediate stage was the founding of the first major theory of classification in geometry, beginning in the 1850's, the Cayley-Sylvester Invariant Theory. Here the objective was to study invariants of "forms" under transformations of their variables. This theory of classification, the

Felix Klein

precursor of Klein's Erlangen Program, can be said to be *implicitly* group-theoretic. Klein's use of groups in geometry was, of course, explicit. (For a thorough analysis of implicit group-theoretic thinking in geometry leading to Klein's Erlangen Program, see [33].) In the next section (2-(c)) we will note the significance of Klein's Erlangen Program (and his other works) for the evolution of group theory. Since the Program originated a hundred years after Lagrange's work and eighty years after Gauss' work, its importance for group theory can best be appreciated *after* a discussion of the evolution of group theory beginning with the works of Lagrange and Gauss and ending with the period around 1870.

(d) *Analysis* (S. Lie, 1874; H. Poincaré and F. Klein, 1876)

In 1874 Lie introduced his general theory of (continuous) transformation groups—essentially what we call Lie groups today. Such a group is represented by the transformations

$$x_i' = f_i(x_1, x_2, \dots, x_n, a_1, a_2, \dots, a_n), \qquad i = 1, 2, \dots, n,$$

where the f_i are analytic functions in the x_i and a_i (the a_i are parameters, with both x_i and a_i real or complex). For example, the transformations given by

$$x' = \frac{ax+b}{cx+d}, \quad \text{where } a, b, c, d, \text{ are real numbers and } ad - bc \neq 0,$$

define a continuous transformation group.

Lie thought of himself as the successor of N. H. Abel and Galois, doing for differential equations what they had done for algebraic equations. His work was inspired by the observation

Sophus Lie

that almost all the differential equations which had been integrated by the older methods remain invariant under continuous groups that can be easily constructed. He was then led to consider, in general, differential equations that remain invariant under a given continuous group and to investigate the possible simplifications in these equations which result from the known properties of the given group (cf. Galois theory). Although Lie did not succeed in the actual formulation of a "Galois theory of differential equations," his work was fundamental in the subsequent formulation of such a theory by E. Picard (1883/1887) and E. Vessiot (1892).

Poincaré and Klein began their work on "automorphic functions" and the groups associated with them around 1876. Automorphic functions (which are generalizations of the circular, hyperbolic, elliptic, and other functions of elementary analysis) are functions of a complex variable z, analytic in some domain D, which are invariant under the group of transformations

$$z' = \frac{az+b}{cz+d}, \quad (a, b, c, d \text{ real or complex and } ad - bc \neq 0)$$

or under some subgroup of this group. Moreover, the group in question must be "discontinuous" (i.e., any compact domain contains only finitely many transforms of any point). Examples of such groups are the modular group (in which a, b, c, d are integers and $ad - bc = 1$), which is associated with the elliptic modular functions, and Fuchsian groups (in which a, b, c, d are real and $ad - bc = 1$) associated with the Fuchsian automorphic functions. As in the case of Klein's Erlangen Program, we will explore the consequences of these works for group theory in section 2-(c).

2. Development of "specialized" theories of groups

In §1 we outlined four major sources in the evolution of group theory. The first source—classical algebra—led to the theory of permutation groups; the second source—number theory—led to the theory of abelian groups; the third and fourth sources—geometry and analysis—led to the theory of transformation groups. We will now outline some developments within these specialized theories.

(a) *Permutation Groups*

As noted earlier, Lagrange's work of 1770 initiated the study of permutations in connection with the study of the solution of equations. It was probably the first clear instance of implicit group-theoretic thinking in mathematics. It led directly to the works of P. Ruffini, Abel, and Galois during the first third of the 19th century, and to the concept of a permutation group.

Ruffini and Abel proved the unsolvability of the quintic by building upon the ideas of Lagrange concerning resolvents. Lagrange showed that a necessary condition for the solvability of the general polynomial equation of degree n is the existence of a resolvent of degree less than n. Ruffini and Abel showed that such resolvents do not exist for $n > 4$. In the process they developed a considerable amount of permutation theory. (See [**1**], [**9**], [**19**], [**23**], [**24**], [**25**], [**30**], [**33**] for details.) It was Galois, however, who made the fundamental conceptual advances, and who is considered by many as the founder of (permutation) group theory.

Galois' aim went well beyond finding a method for solvability of equations. He was concerned with gaining insight into general principles, dissatisfied as he was with the methods of his predecessors: "From the beginning of this century," he wrote, "computational procedures have become so complicated that any progress by those means has become impossible" [**19**, p. 92].

Galois recognized the separation between "Galois theory" (i.e., the correspondence between fields and groups) and its application to the solution of equations, for he wrote that he was presenting "the general principles and just one application" of the theory [**19**, p. 42]. "Many of the early commentators on Galois theory failed to recognize this distinction, and this led to an emphasis on applications at the expense of the theory" (Kiernan, [**19**]).

Galois was the first to use the term "group" in a technical sense—to him it signified a collection of permutations closed under multiplication: "if one has in the same group the substitutions S and T one is certain to have the substitution ST" [**33**, p. 111]. He recognized that the most important properties of an algebraic equation were reflected in certain properties of a group uniquely associated with the equation— "the group of the equation." To describe these properties he invented the fundamental notion of normal subgroup and used it to great effect. While the issue of resolvent equations preoccupied Lagrange, Ruffini, and Abel, Galois' basic idea was to bypass them, for the construction of a resolvent required great skill and was not based on a clear methodology. Galois noted instead that the existence of a resolvent was equivalent to the existence of a normal subgroup of prime index in the group of the equation. This insight shifted consideration from the resolvent equation to the group of the equation and its subgroups.

Galois defines the group of an equation as follows [**19**, p. 80]:

> Let an equation be given, whose m roots are $a, b, c, \ldots$. There will always be a group of permutations of the letters $a, b, c, \ldots$ which has the following property: 1) that every function of the roots, invariant under the substitutions of that group, is rationally known [i.e., is a rational function of the coefficients and any adjoined quantities]. 2) conversely, that every function of the roots, which can be expressed rationally, is invariant under these substitutions.

The definition says essentially that the group of the equation consists of those permutations of the roots of the equation which leave invariant all relations among the roots over the field of coefficients of the equation—basically the definition we would give today. Of course the definition does not guarantee the existence of such a group, and so Galois proceeds to demonstrate it. Galois next investigates how the group changes when new elements are adjoined to the "ground field" F. His treatment is amazingly close to the standard treatment of this matter in a modern algebra text.

Galois' work was slow in being understood and assimilated. In fact, while it was done around 1830, it was published posthumously in 1846, by J. Liouville. Beyond his technical accomplishments, Galois "challenged the development of mathematics in two ways. He discovered, but left unproved, theorems which called for proofs based on new, sophisticated concepts and calculations. Also, the task of filling the gaps in his work necessitated a fundamental clarification of his methods and their group theoretical essence" (Wussing, [**33**]). For details see [**12**], [**19**], [**23**], [**25**], [**29**], [**31**], [**33**].

The other major contributor to permutation theory in the first half of the 19th century was Cauchy. In several major papers in 1815 and 1844 Cauchy inaugurated the theory of permutation groups as an autonomous subject. (Before Cauchy, permutations were not an object of independent study but rather a useful device for the investigation of solutions of polynomial equations.) Although Cauchy was well aware of the work of Lagrange and Ruffini (Galois' work was not yet published at the time), Wussing suggests that Cauchy "was definitely not inspired directly by the contemporary group-theoretic formulation of the solution of algebraic equations" [**33**].

In these works Cauchy gives the first systematic development of the subject of permutation groups. In the 1815 papers Cauchy uses no special name for sets of permutations closed under multiplication. However, he recognizes their importance and gives a name to the number of elements in such a closed set, calling it "diviseur indicatif." In the 1844 paper he defines the concept of a group of permutations generated by certain elements [**22**, p. 65].

> Given one or more substitutions involving some or all of the elements $x, y, z, \ldots$ I call the products of these substitutions, by themselves or by any other, in any order, *derived* substitutions. The given substitutions, together with the derived ones, form what I call a *system of conjugate substitutions.*

In these works, which were very influential, Cauchy makes several lasting additions to the terminology, notation, and results of permutation theory. For example, he introduces the permutation notation $\begin{pmatrix} x & y & z \\ x & z & y \end{pmatrix}$ in use today, as well as the cyclic notation for permutations; defines the product of permutations, the degree of a permutation, cyclic permutation, transposition; recognizes the identity permutation as a permutation; discusses what we would call today the direct product of two groups; and deals with the alternating groups extensively. Here is a sample of some of the results he proves.

(i) Every even permutation is a product of 3-cycles.

(ii) If p (prime) is a divisor of the order of a group, then there exists a subgroup of order p. (This is known today as "Cauchy's theorem", though it was stated without proof by Galois.)

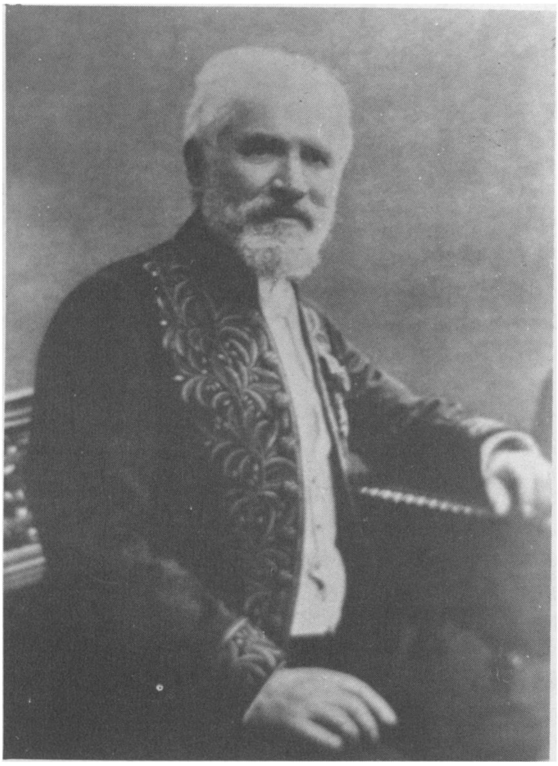
(iii) Determined all subgroups of S_3, S_4, S_5, S_6 (making an error in S_6.)

(iv) All permutations which commute with a given one form a group (the centralizer of an element).

It should be noted that all these results were given and proved in the context of permutation groups. For details see [**6**], [**8**], [**23**], [**24**], [**25**], [**33**].

The crowning achievement of these two lines of development—a symphony on the grand themes of Galois and Cauchy—was Jordan's important and influential *Traité des substitutions et des équations algébriques* of 1870. Although the author states in the preface that "the aim of the work is to develop Galois' method and to make it a proper field of study, by showing with what facility it can solve all principal problems of the theory of equations," it is in fact group theory per se—not as an offshoot of the theory of solvability of equations—which forms the central object of study.

The striving for a mathematical synthesis based on key ideas is a striking characteristic of Jordan's work as well as that of a number of other mathematicians of the period (e.g., F. Klein). The concept of a (permutation) group seemed to Jordan to provide such a key idea. His approach



Camille Jordan

enabled him to give a unified presentation of results due to Galois, Cauchy, and others. His application of the group concept to the theory of equations, algebraic geometry, transcendental functions, and theoretical mechanics was also part of the unifying and synthesizing theme. "In his book Jordan wandered through all of algebraic geometry, number theory, and function theory in search of interesting permutation groups" (Klein, [20]). In fact, the aim was a survey of all of mathematics by areas in which the theory of permutation groups had been applied or seemed likely to be applicable. "The work represents... a review of the whole of contemporary mathematics from the standpoint of the occurrence of group-theoretic thinking in permutation-theoretic form" (Wussing, [33]).

The *Traité* embodied the substance of most of Jordan's publications on groups up to that time (he wrote over 30 articles on groups during the period 1860–1880) and directed attention to a large number of difficult problems, introducing many fundamental concepts. For example, Jordan makes explicit the notions of isomorphism and homomorphism for (substitution) groups, introduces the term "solvable group" for the first time in a technical sense, introduces the concept of a composition series, and proves part of the Jordan-Hölder theorem, namely, that the indices in two composition series are the same (the concept of a quotient group was not explicitly recognized at this time); and he undertakes a very thorough study of transitivity and primitivity for permutation groups, obtaining results most of which have not since been superseded. Jordan also gives a proof that A_n is simple for $n > 4$.

An important part of the treatise is devoted to a study of the "linear group" and some of its subgroups. In modern terms these constitute the so-called classical groups, namely, the general linear group, the unimodular group, the orthogonal group, and the symplectic group. Jordan considers these groups only over finite fields, and proves their simplicity in certain cases. It should be noted, however, that he considers these groups as permutation groups rather than groups of matrices or linear transformations (see [**29**], [**33**]).

Jordan's *Traité* is a landmark in the evolution of group theory. His permutation-theoretic point of view, however, was soon to be overtaken by the conception of a group as a group of transformations (see (c) below). "The *Traité* marks a pause in the evolution and application of the permutation-theoretic group concept. It was an expression of Jordan's deep desire to effect a conceptual synthesis of the mathematics of his time. That he tried to achieve such a synthesis by relying on the concept of a permutation group, which the very next phase of mathematical development would show to have been unduly restricted, makes for both the glory and the limitations of the *Traité*..." (Wussing, [**33**]). For details see [**9**], [**13**], [**19**], [**20**], [**22**], [**24**], [**29**], [**33**].

(b) *Abelian Groups*

As noted earlier, the main source for abelian group theory was number theory, beginning with Gauss' *Disquisitiones Arithmeticae*. In contrast to permutation theory, group-theoretic modes of thought in number theory remained implicit until about the last third of the 19th century. Until that time no explicit use of the term "group" was made, and there was no link to the contemporary, flourishing theory of permutation groups. We now give a sample of some implicit group-theoretic work in number theory, especially in algebraic number theory.

Algebraic number theory arose in connection with Fermat's conjecture concerning the equation $x^n + y^n = z^n$, Gauss' theory of binary quadratic forms, and higher reciprocity laws. Algebraic number fields and their arithmetical properties were the main objects of study. In 1846 G. L. Dirichlet studied the units in an algebraic number field and established that (*in our terminology*) the group of these units is a direct product of a finite cyclic group and a free abelian group of finite rank. At about the same time E. Kummer introduced his "ideal numbers," defined an equivalence relation on them, and derived, for cyclotomic fields, certain special properties of the number of equivalence classes (the so-called class number of a cyclotomic field; in our terminology, the order of the ideal class group of the cyclotomic field). Dirichlet had earlier made similar studies of *quadratic* fields.

In 1869 E. Schering, a former student of Gauss, investigated the structure of Gauss' (group of) equivalence classes of binary quadratic forms. He found certain fundamental classes from which all classes of forms could be obtained by composition. In group-theoretic terms, Schering found a basis for the abelian group of equivalence classes of binary quadratic forms.

L. Kronecker generalized Kummer's work on cyclotomic fields to arbitrary algebraic number fields. In a paper in 1870 on algebraic number theory, entitled "Auseinandersetzung einiger Eigenschaften der Klassenzahl idealer complexer Zahlen," he began by taking a very abstract point of view: he considered a finite set of arbitrary "elements," and defined an abstract operation on them which satisfied certain laws—laws which we may take nowadays as axioms for a finite abelian group:

> Let $\theta', \theta'', \theta''', \ldots$ be finitely many elements such that with any two of them we can associate a third by means of a definite procedure. Thus, if f denotes the procedure and θ', θ'' are two (possibly equal) elements, then there exists a θ''' equal to $f(\theta', \theta'')$. Furthermore, $f(\theta', \theta'') = f(\theta'', \theta')$, $f(\theta', f(\theta'', \theta''')) = f(f(\theta', \theta''), \theta''')$ and if θ'' is different from θ''' then $f(\theta', \theta'')$ is different from $f(\theta', \theta''')$. Once this is assumed we can replace the operation $f(\theta', \theta'')$ by multiplication $\theta' \cdot \theta''$ provided that instead of equality we employ equivalence. Thus using the usual equivalence symbol " $\sim$ " we define the equivalence $\theta' \cdot \theta'' \sim \theta'''$ by means of the equation $f(\theta', \theta'') = \theta'''$.

Kronecker aimed at working out the laws of combination of "magnitudes," in the process giving an implicit definition of a finite abelian group. From the above abstract considerations

Kronecker deduces the following consequences:

(i) If θ is any "element" of the set under discussion, then $\theta^k = 1$ for some positive integer k. If k is the smallest such then θ is said to "belong to k". If θ belongs to k and $\theta^m = 1$ then k divides m.

(ii) If an element θ belongs to k, then every divisor of k has an element belonging to it.

(iii) If θ and θ' belong to k and k' respectively, and k and k' are relatively prime, then $\theta\theta'$ belongs to kk'.

(iv) There exists a "fundamental system" of elements $\theta_1, \theta_2, \theta_3, \ldots$ such that the expression $\theta_1^{h_1}\theta_2^{h_2}\theta_3^{h_3} \cdots (h_i = 1, 2, 3, \ldots, n_i)$ represents each element of the given set of elements just once. The numbers $n_1, n_2, n_3, \ldots$ to which, respectively, $\theta_1, \theta_2, \theta_3, \ldots$ belong, are such that each is divisible by its successor; the product $n_1 n_2 n_3 \cdots$ is equal to the totality of elements of the set.

The above can, of course, be interpreted as well known results on finite abelian groups; in particular (iv) can be taken as the basis theorem for such groups. Once Kronecker establishes this general framework, he applies it to the special cases of equivalence classes of binary quadratic forms and to ideal classes. He notes that when applying (iv) to the former one obtains Schering's result.

Although Kronecker did not relate his implicit definition of a finite abelian group to the (by that time) well established concept of a permutation group, of which he was well aware, he clearly recognized the advantages of the abstract point of view which he adopted:

> The very simple principles... are applied not only in the context indicated but also frequently, elsewhere—even in the elementary parts of number theory. This shows, and it is otherwise easy to see, that these principles belong to a more general and more abstract realm of ideas. It is therefore proper to free their development from all inessential restrictions, thus making it unnecessary to repeat the same argument when applying it in different cases. ... Also, when stated with all admissible generality, the presentation gains in simplicity and, since only the truly essential features are thrown into relief, in transparency.

The above lines of development were capped in 1879 by an important paper of G. Frobenius and L. Stickelberger entitled "On groups of commuting elements." Although Frobenius and Stickelberger built on Kronecker's work, they used the concept of an abelian group explicitly and, moreover, made the important advance of recognizing that the abstract group concept embraces congruences and Gauss' composition of forms as well as the substitution groups of Galois. (They also mention, in footnotes, groups of infinite order, namely groups of units of number fields and the group of all roots of unity.) One of their main results is a proof of the basis theorem for finite abelian groups, including a proof of the uniqueness of decomposition. It is interesting to compare their explicit, "modern," formulation of the theorem to that of Kronecker ((iv) above):

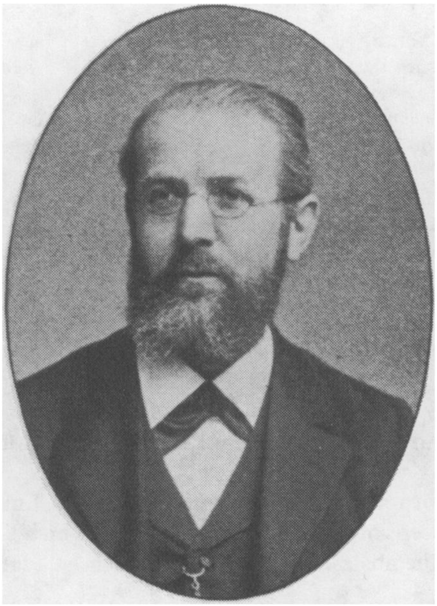
> A group that is not irreducible [indecomposable] can be decomposed into purely irreducible factors. As a rule, such a decomposition can be accomplished in many ways. However, regardless of the way in which it is carried out, the number of irreducible factors is always the same and the factors in the two decompositions can be so paired off that the corresponding factors have the same order [**33**, p. 235].

They go on to identify the "irreducible factors" as cyclic groups of prime power orders. They then apply their results to groups of integers modulo m, binary quadratic forms, and ideal classes in algebraic number fields.

The paper by Frobenius and Stickelberger is "a remarkable piece of work, building up an independent theory of finite abelian groups on its own foundation in a way close to modern views" (Fuchs, [**30**]). For details on this section (b), see [**5**], [**9**], [**24**[, [**30**], [**33**].

(c) *Transformation Groups*

As in number theory, so in geometry and analysis, group-theoretic ideas remained implicit until the last third of the 19th century. Moreover, Klein's (and Lie's) explicit use of groups in geometry influenced conceptually rather than technically the evolution of group theory, for it signified a



Georg Frobenius

genuine shift in the development of that theory from a preoccupation with permutation groups to the study of groups of transformations. (That is not to imply, of course, that permutation groups were no longer studied.) This transition was also notable in that it pointed to a turn from finite groups to infinite groups.

Klein noted the connection of his work with permutation groups but also realized the departure he was making. He stated that what Galois theory and his own program have in common is the investigation of "groups of changes," but added that "to be sure, the objects the changes apply to are different: there [Galois theory] one deals with a finite number of discrete elements, whereas here one deals with an infinite number of elements of a continuous manifold" [**33**, p. 191]. To continue the analogy, Klein notes that just as there is a theory of permutation groups, "we insist on a *theory of transformations*, a study of groups generated by transformations of a given type" [**33**, p. 191].

Klein shunned the abstract point of view in group theory, and even his technical definition of a (transformation) group is deficient: "Now let there be given a sequence of transformations $A, B, C, \ldots$ If this sequence has the property that the composite of any two of its transformations yields a transformation that again belongs to the sequence, then the latter will be called a group of transformations" [**33**, p. 185]. His work, however, broadened considerably the conception of a group and its applicability in other fields of mathematics. Klein did much to promote the view that group-theoretic ideas are fundamental in mathematics: "Group theory appears as a distinct discipline throughout the whole of modern mathematics. It permeates the most varied areas as an ordering and classifying principle" [33, p. 228].

There was another context in which groups were associated with geometry, namely, "motion-geometry;" i.e., the use of motions or transformations of geometric objects as group elements. Already in 1856 W. R. Hamilton considered (implicitly) "groups" of the regular solids. Jordan, in 1868, dealt with the classification of all subgroups of the group of motions of Euclidean 3-space. And Klein in his *Lectures on the Icosahedron* of 1884 "solved" the quintic equation by means of the symmetry group of the icosahedron. He thus discovered a deep connection between the groups of rotations of the regular solids, polynomial equations, and complex function theory. (In these *Lectures* there also appears the "Klein 4-group".)

Already in the late 1860's Klein and Lie had undertaken, jointly, "to investigate geometric or analytic objects that are transformed into themselves by *groups of changes*." (This is Klein's retrospective description, in 1894, of their program.) While Klein concentrated on discrete groups, Lie studied continuous transformation groups. Lie realized that the theory of continuous transformation groups was a very powerful tool in geometry and differential equations and he set himself the task of "determining all groups of...[continuous] transformations" [33, p. 214]. He achieved his objective by the early 1880's with the classification of these groups (see [33] for details). A classification of discontinuous transformation groups was obtained by Poincaré and Klein a few years earlier.

Beyond the technical accomplishments in the areas of discontinuous and continuous transformation groups (extensive theories developed in both areas and both are still nowadays active fields of research), what is important for us in the founding of these theories is that

(i) They provided a major extension of the scope of the concept of a group—from permutation groups and abelian groups to transformation groups;
(ii) They provided important examples of infinite groups—previously the only objects of study were finite groups;
(iii) They greatly extended the range of applications of the group concept to include number theory, the theory of algebraic equations, geometry, the theory of differential equations (both ordinary and partial), and function theory (automorphic functions, complex functions).

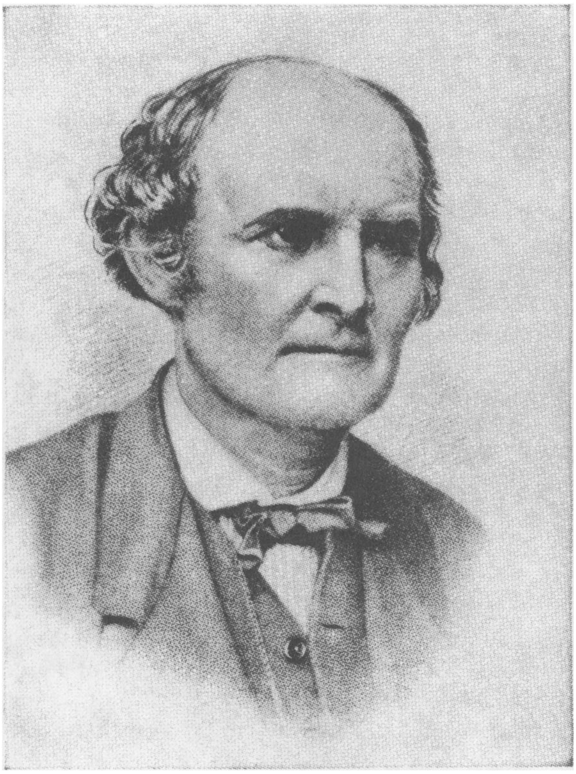
All this occurred prior to the emergence of the abstract group concept. In fact, these developments were instrumental in the emergence of the concept of an abstract group, which we describe next. For further details on this section (c), see [5], [7], [9], [17], [18], [20], [24], [29], [33].

3. Emergence of abstraction in group theory

The abstract point of view in group theory emerged slowly. It took over one hundred years from the time of Lagrange's implicit group-theoretic work of 1770 for the abstract group concept to evolve. E. T. Bell discerns several stages in this process of evolution towards abstraction and axiomatization:

> The entire development required about a century. Its progress is typical of the evolution of any major mathematical discipline of the recent period; first, the discovery of isolated phenomena, then the recognition of certain features common to all, next the search for further instances, their detailed calculation and classification; then the emergence of general principles making further calculations, unless needed for some definite application, superfluous; and last, the formulation of postulates crystallizing in abstract form the structure of the system investigated [2].

Although somewhat oversimplified (as all such generalizations tend to be), this is nevertheless a useful framework. Indeed, in the case of group theory, first came the "isolated phenomena"—e.g., permutations, binary quadratic forms, roots of unity; then the recognition of "common features" —the concept of a finite group, encompassing both permutation groups and finite abelian groups (cf. the paper of Frobenius and Stickelberger cited in section 2(b)); next the search for "other instances"—in our case transformation groups (see section 2(c)); and finally the formulation of "postulates"—in this case the postulates of a group, encompassing both the finite and infinite



Arthur Cayley

cases. We now consider when and how the intermediate and final stages of abstraction occurred.

In 1854 Cayley, in a paper entitled "On the theory of groups, as depending on the symbolic equation $\theta^n = 1$," gave the first abstract definition of a finite group. (In 1858 R. Dedekind, in lectures on Galois theory at Göttingen, gave another.) Here is Cayley's definition:

> A set of symbols $1, \alpha, \beta, \ldots$ all of them different, and such that the product of any two of them (no matter in what order), or the product of any one of them into itself, belongs to the set, is said to be a *group*.

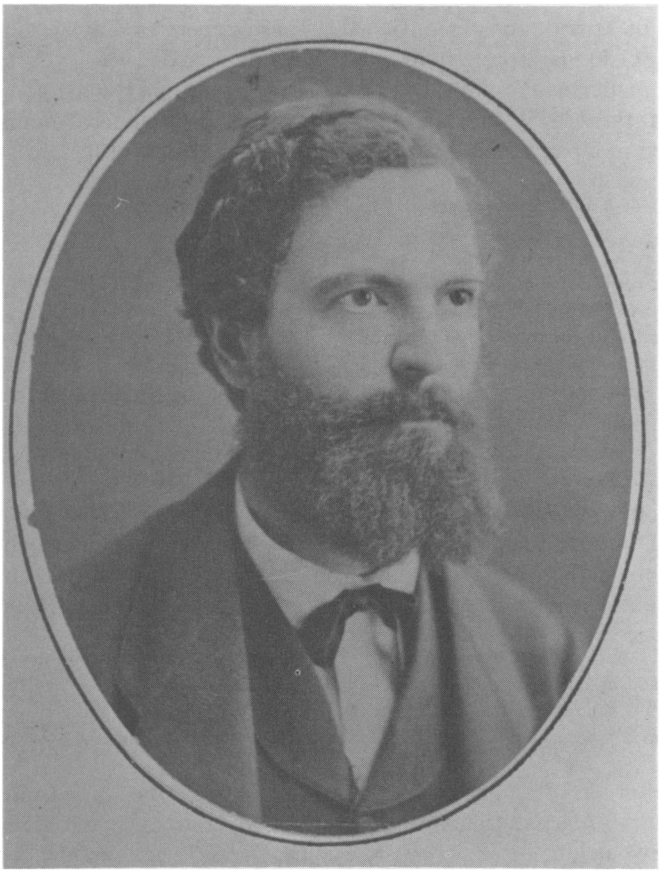
Cayley goes on to say that

> These symbols are not in general convertible [commutative], but are associative,

and

> it follows that if the entire group is multiplied by any one of the symbols, either as further or nearer factor [i.e., on the left or on the right], the effect is simply to reproduce the group.

Cayley then presents several examples of groups, such as the quaternions (under addition), invertible matrices (under multiplication), permutations, Gauss' quadratic forms, and groups arising in elliptic function theory. Next he shows that every abstract group is (in our terminology) isomorphic to a permutation group, a result now known as "Cayley's theorem." He seems to have been well aware of the concept of isomorphic groups, although he does not define it explicitly. He introduces, however, the multiplication table of a (finite) group and asserts that an abstract group



Heinrich Weber

is determined by its multiplication table. He then goes on to determine all the groups of orders four and six, showing there are two of each by displaying multiplication tables. Moreoever, he notes that the cyclic group of order n "is in every respect analogous to the system of the roots of the ordinary equation $x^n - 1 = 0$," and that there exists only one group of a given prime order.

Cayley's orientation towards an abstract view of groups—a remarkable accomplishment at this time of the evolution of group theory—was due, at least in part, to his contact with the abstract work of G. Boole. The concern with the abstract foundations of mathematics was characteristic of the circles around Boole, Cayley, and Sylvester already in the 1840's. Cayley's achievement was, however, only a personal triumph. His abstract definition of a group attracted no attention at the time, even though Cayley was already well known. The mathematical community was apparently not ready for such abstraction: permutation groups were the only groups under serious investigation, and more generally, the formal approach to mathematics was still in its infancy. As M. Kline put it in his inimitable way **[21]**: "Premature abstraction falls on deaf ears, whether they belong to mathematicians or to students." For details see **[22]**, **[23]**, **[24]**, **[25]**, **[29]**, **[33]**.

It was only a quarter of a century later that the abstract group concept began to take hold. And it was Cayley again who in four short papers on group theory written in 1878 returned to the abstract point of view he adopted in 1854. Here he stated the general problem of finding all groups of a given order and showed that any (finite) group is isomorphic to a group of permutations. But, as he remarked, this "does not in any wise show that the best or easiest mode of treating the general problem is thus to regard it as a problem of substitutions; and it seems clear that the better course is to consider the general problem in itself, and to deduce from it the theory of groups of substitutions" [**22**, p. 141]. These papers of Cayley, unlike those of 1854, inspired a number of fundamental group-theoretic works.

Another mathematician who advanced the abstract point of view in group theory (and more generally in algebra) was H. Weber. It is of interest to see his "modern" definition of an abstract (finite) group given in a paper of 1882 on quadratic forms [**23**, p. 113]:

> A system G of h arbitrary elements $\theta_1, \theta_2, \dots, \theta_h$ is called a group of degree h if it satisfies the following conditions:
> I. By some rule which is designated as composition or multiplication, from any two elements of the same system one derives a new element of the same system. In symbols $\theta_r\theta_s = \theta_t$.
> II. It is always true that $(\theta_r\theta_s)\theta_t = \theta_r(\theta_s\theta_t) = \theta_r\theta_s\theta_t$.
> III. From $\theta\theta_r = \theta\theta_s$ or from $\theta_r\theta = \theta_s\theta$ it follows that $\theta_r = \theta_s$.

Weber's and other definitions of abstract groups given at the time applied to *finite* groups only. They thus encompassed the two theories of permutation groups and (finite) abelian groups, which derived from the two sources of classical algebra (polynomial equations) and number theory, respectively. Infinite groups, which derived from the theories of (discontinuous and continuous) transformation groups, were not subsumed under those definitions. It was W. von Dyck who, in an important and influential paper in 1882 entitled "Group-theoretic studies," consciously included and combined, for the first time, all of the major historical roots of abstract group theory —the algebraic, number theoretic, geometric, and analytic. In von Dyck's own words:

> The aim of the following investigations is to continue the study of the properties of a group in its abstract formulation. In particular, this will pose the question of the extent to which these properties have an invariant character present in all the different realizations of the group, and the question of what leads to the exact determination of their essential group-theoretic content.

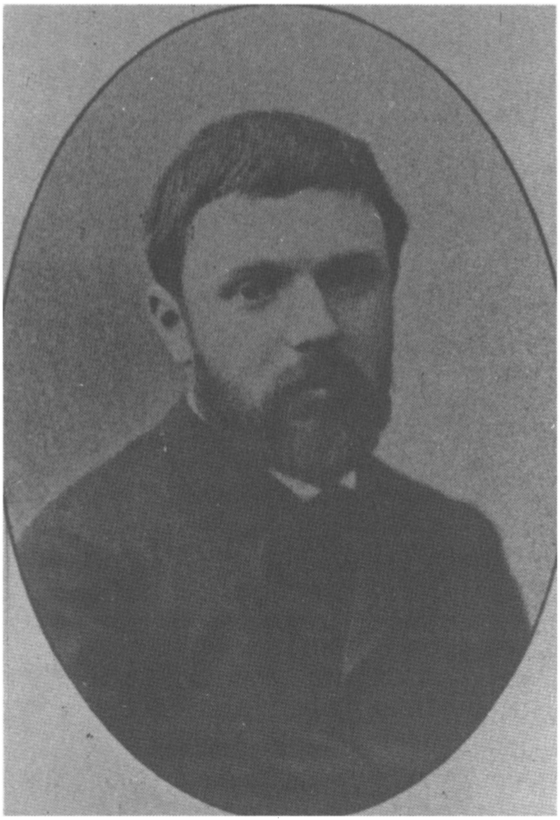
Von Dyck's definition of an abstract group, which included both the finite and infinite cases, was given in terms of generators (he calls them "operations") and defining relations (the definition is somewhat long—see [**7**, pp. 5, 6]). He stresses that "in this way all... *isomorphic* groups are included in *a single* group," and that "the *essence* of a group is no longer expressed by a particular presentation form of its operations but rather by their mutual relations." He then goes on to construct the free group on n generators, and shows (essentially, without using the terminology) that every finitely generated group is a quotient group of a free group of finite rank. What is important from the point of view of postulates for group theory is that von Dyck was the first to require explicitly the existence of an inverse in his definition of a group: "We require for our considerations that a group which contains the operation T_k must also contain its inverse T_k^{-1}." In a second paper (in 1883) von Dyck applied his abstract development of group theory to permutation groups, finite rotation groups (symmetries of polyhedra), number theoretic groups, and transformation groups.

Although various postulates for groups appeared in the mathematical literature for the next twenty years, the abstract point of view in group theory was not universally applauded. In particular, Klein, one of the major contributors to the development of group theory, thought that the "abstract formulation is excellent for the working out of proofs but it does not help one find new ideas and methods," adding that "in general, the disadvantage of the [abstract] method is that it fails to encourage thought" [**33**, p. 228].

Despite Klein's reservations, the mathematical community was at this time (early 1880's) receptive to the abstract formulations (cf. the response to Cayley's definition of 1854). The major reasons for this receptivity were:

(i) There were now several major "concrete" theories of groups—permutation groups, abelian groups, discontinuous transformation groups (the finite and infinite cases), and continuous transformation groups, and this warranted abstracting their essential features.

(ii) Groups came to play a central role in diverse fields of mathematics, such as different parts of algebra, geometry, number theory and several areas of analysis, and the abstract view of



Henri Poincaré

groups was thought to clarify what was essential for such applications and to offer opportunities for further applications.

(iii) The formal approach, aided by the penetration into mathematics of set theory and mathematical logic, became prevalent in other fields of mathematics, for example, various areas of geometry and analysis.

In the next section we will follow, very briefly, the evolution of that abstract point of view in group theory.

4. Consolidation of the abstract group *concept*; dawn of abstract group *theory*

The abstract group concept spread rapidly during the 1880's and 1890's, although there still appeared a great many papers in the areas of permutation and transformation groups. The abstract viewpoint was manifested in two ways:

(a) Concepts and results introduced and proved in the setting of "concrete" groups were now reformulated and reproved in an abstract setting;

(b) Studies originating in, and based on, an abstract setting began to appear.

An interesting example of the former case is a reproving by Frobenius, in an abstract setting, of Sylow's theorem, which was proved by Sylow in 1872 for permutation groups. This was done in 1887, in a paper entitled "Neuer Beweis Sylowschen Satzes." Although Frobenius admits that the fact that every finite group can be represented by a group of permutations proves that Sylow's theorem must hold for all finite groups, he nevertheless wishes to establish the theorem abstractly:

"Since the symmetric group, which is introduced in all these proofs, is totally alien to the context of Sylow's theorem, I have tried to find a new derivation of it... ." (For a case study of the evolution of abstraction in group theory in connection with Sylow's theorem see [**28**] and [**32**].)

Hölder was an important contributor to abstract group theory, and was responsible for introducing a number of group-theoretic concepts abstractly. For example, in 1889 he introduced the abstract notion of a quotient group (the "quotient group" was first seen as the Galois group of the "auxiliary equation", later as a homomorphic image and only in Hölder's time as a group of cosets), and "completed" the proof of the Jordan-Hölder theorem, namely, that the quotient groups in a composition series are invariant up to isomorphism (see section 2(a) for Jordan's contribution). In 1893, in a paper on groups of order p^3, pq^2, pqr, and p^4, he introduced abstractly the concept of an automorphism of a group. Hölder was also the first to study simple groups abstractly. (Previously they were considered in concrete cases—as permutation groups, transformation groups, and so on.) As he says [**29**, p. 338]. "It would be of the greatest interest if a survey of all simple groups with a finite number of operations could be known." (By "operations" Hölder meant elements.) He then goes on to determine the simple groups of order up to 200.

Other typical examples of studies in an abstract setting are the papers by Dedekind and G. A. Miller in 1897/1898 on Hamiltonian groups–i.e., nonabelian groups in which all subgroups are normal. They (independently) characterize such groups abstractly, and introduce in the process the notions of the commutator of two elements and the commutator subgroup (Jordan had previously introduced the notion of commutator of two permutations).

The theory of group characters and the representation theory for finite groups (created at the end of the 19th century by Frobenius and Burnside/Frobenius/Molien, respectively) also belong to the area of abstract group theory, as they were used to prove important results about abstract groups. See [**17**] for details.

Although the abstract group *concept* was well established by the end of the 19th century, "this was not accompanied by a general acceptance of the associated method of presentation in papers, textbooks, monographs, and lectures. Group-theoretic monographs based on the abstract group concept did not appear until the beginning of the 20th century. Their appearance marked the birth of abstract group *theory*" (Wussing, [**33**]).

The earliest monograph devoted entirely to abstract group theory was the book by J. A. de Séguier of 1904 entitled *Elements of the Theory of Abstract Groups* [**27**]. At the very beginning of the book there is a set-theoretic introduction based on the work of Cantor: "De Séguier may have been the first algebraist to take note of Cantor's discovery of uncountable cardinalities" (B. Chandler and W. Magnus, [**7**]). Next is the introduction of the concept of a semigroup with two-sided cancellation law and a proof that a finite semigroup is a group. There is also a proof, by means of counterexamples, of the independence of the group postulates. De Séguier's book also includes a discussion of isomorphisms, homomorphisms, automorphisms, decomposition of groups into direct products, the Jordan-Hölder theorem, the first isomorphism theorem, abelian groups including the basis theorem, Hamiltonian groups, and finally, the theory of p-groups. All this is done in the abstract, with "concrete" groups relegated to an appendix. "The style of de Séguier is in sharp contrast to that of Dyck. There are no intuitive considerations... and there is a tendency to be as abstract and as general as possible..." (Chandler and Magnus, [**7**]).

De Séguier's book was devoted largely to finite groups. The first abstract monograph on group theory which dealt with groups in general, relegating finite groups to special chapters, was O. Schmidt's *Abstract Theory of Groups* of 1916 [**26**]. Schmidt, founder of the Russian school of group theory, devotes the first four chapters of his book to group properties common to finite and infinite groups. Discussion of finite groups is postponed to chapter 5, there being ten chapters in all. See [**7**], [**10**], [**33**].

5. Divergence of developments in group theory

Group theory evolved from several different sources, giving rise to various concrete theories. These theories developed independently, some for over one hundred years (beginning in 1770) before they converged (early 1880's) within the abstract group concept. Abstract group theory emerged and was consolidated in the next thirty to forty years. At the end of that period (around 1920) one can discern the divergence of group theory into several distinct "theories." Here is the barest indication of *some* of these advances and new directions in group theory, beginning in the 1920's (with contributors and approximate dates):

(a) Finite group theory. The major problem here, already formulated by Cayley (1870's) and studied by Jordan and Hölder, was to find all finite groups of a given order. The problem proved too difficult and mathematicians turned to special cases (suggested especially by Galois theory): to find all simple or all solvable groups (cf. the Feit-Thompson theorem of 1963, and the classification of all finite simple groups in 1981). See [**14**], [**15**], [**30**].

b) Extensions of certain results from finite group theory to infinite groups with finiteness conditions; e.g., O. J. Schmidt's proof, in 1928, of the Remak-Krull-Schmidt theorem. See [**5**].

c) Group presentations (Combinatorial Group Theory), begun by von Dyck in 1882, and continued in the 20th century by M. Dehn, H. Tietze, J. Nielsen, E. Artin, O. Schreier, et al. For a full account, see [**7**].

d) Infinite abelian group theory (H. Prüfer, R. Baer, H. Ulm et al.—1920's to 1930's). See [**30**].

e) Schreier's theory of group extensions (1926), leading later to the cohomology of groups.

f) Algebraic groups (A. Borel, C. Chevalley et al.—1940's).

g) Topological groups, including the extension of group representation theory to continuous groups (Schreier, E. Cartan. L. Pontrjagin, I. Gelfand, J. von Neumann et al.—1920's and 1930's). See [**4**].

FIGURE 1 gives a diagrammatic sketch of the evolution of group theory as outlined in the various sections and as summarized at the beginning of this section.

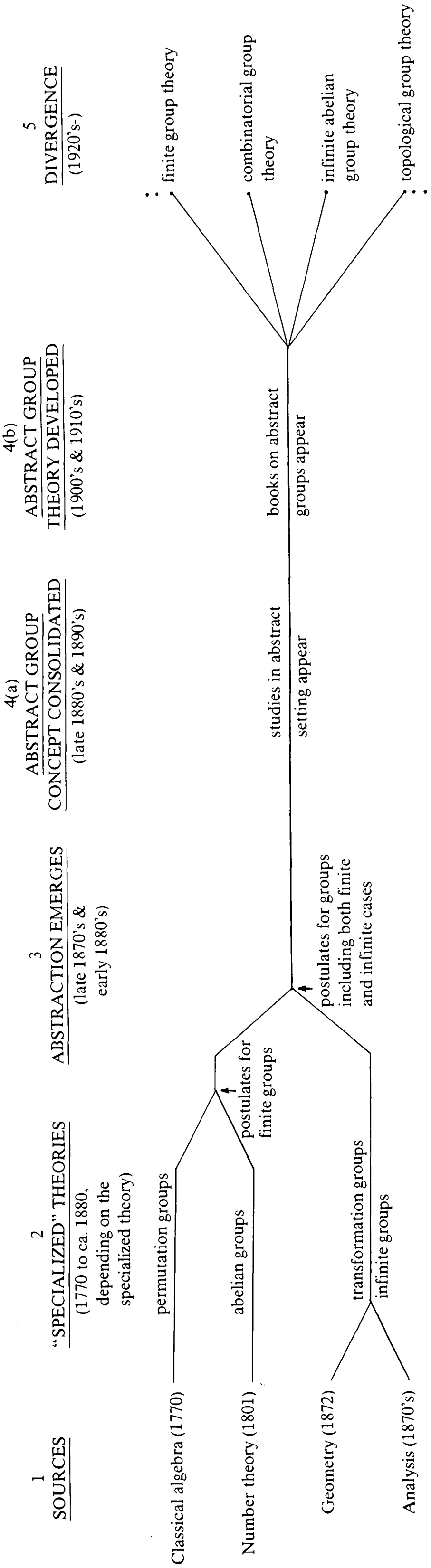

1
SOURCES
2
"SPECIALIZED" THEORIES
(1770 to ca. 1880, depending on the specialized theory)
3
ABSTRACTION EMERGES
(late 1870's & early 1880's)
4(a)
ABSTRACT GROUP CONCEPT CONSOLIDATED
(late 1880's & 1890's)
4(b)
ABSTRACT GROUP THEORY DEVELOPED
(1900's & 1910's)
5
DIVERGENCE
(1920's-)
Classical algebra (1770)
Number theory (1801)
Geometry (1872)
Analysis (1870's)
permutation groups
abelian groups
transformation groups
infinite groups
postulates for finite groups
postulates for groups including both finite and infinite cases
studies in abstract setting appear
books on abstract groups appear
finite group theory
combinatorial group theory
infinite abelian group theory
topological group theory

FIGURE 1

References

We give references here to *secondary* sources. Extensive references to *primary* sources, including works referred to in this article, may be found in [**25**] and [**33**].

[1] R. G. Ayoub, Paolo Ruffini's contributions to the quintic, Arch. Hist. Ex. Sc., 23 (1980) 253–277.

[2] E. T. Bell, The Development of Mathematics, McGraw Hill, 1945.

[3] G. Birkhoff, Current trends in algebra, Amer. Math. Monthly, 80 (1973) 760–782 and 81 (1974) 746.

[4] ______, The rise of modern algebra to 1936, in Men and Institutions in American Mathematics, eds. D. Tarwater, J. T. White and J. D. Miller, Texas Tech. Press, 1976, pp. 41–63.

[5] N. Bourbaki, Eléments d'Histoire des Mathématiques, Hermann, 1969.

[6] J. E. Burns, The foundation period in the history of group theory, Amer. Math. Monthly, 20 (1913) 141–148.

[7] B. Chandler and W. Magnus, The History of Combinatorial Group Theory: A Case Study in the History of Ideas, Springer-Verlag, 1982.

[8] A. Dahan, Les travaux de Cauchy sur les substitutions. Etude de son approche du concept de groupe, Arch. Hist. Ex. Sc., 23 (1980) 279–319.

[9] J. Dieudonné (ed.), Abrégé d'Histoire des Mathématiques, 1700–1900, 2 vols., Hermann, 1978.

[10] P. Dubreil, L'algèbre, en France, de 1900 à 1935, Cahiers du seminaire d'histoire des mathématiques, 3 (1981) 69–81.

[11] C. H. Edwards, The Historical Development of the Calculus, Springer-Verlag, 1979.

[12] H. M. Edwards, Galois Theory, Springer-Verlag, 1984.

[13] J. A. Gallian, The search for finite simple groups, this *Magazine*, 49 (1976) 163–179.

[14] D. Gorenstein, Finite Simple Groups: An Introduction to Their Classification, Plenum Press, 1982.

[15] ______, The Classification of Finite Simple Groups, Plenum Press, 1983.

[16] R. R. Hamburg, The theory of equations in the 18th century: The work of Joseph Lagrange, Arch. Hist. Ex. Sc., 16 (1976/77) 17–36.

[17] T. Hawkins, Hypercomplex numbers, Lie groups, and the creation of group representation theory, Arch. Hist. Ex. Sc., 8 (1971/72) 243–287.

[18] ______, The *Erlanger Programme* of Felix Klein: Reflections on its place in the history of mathematics, Hist. Math., 11 (1984) 442–470.

[19] B. M. Kiernan, The development of Galois theory from Lagrange to Artin, Arch. Hist. Ex. Sc., 8 (1971/72) 40–154.

[20] F. Klein, Development of Mathematics in the 19th Century (transl. from the 1928 German ed. by M. Ackerman), in Lie Groups: History, Frontiers and Applications, vol. IX, ed. R. Hermann, Math. Sci. Press, 1979, pp. 1–361.

[21] M. Kline, Mathematical Thought from Ancient to Modern Times, Oxford Univ. Press, 1972.

[22] D. R. Lichtenberg, The Emergence of Structure in Algebra, Doctoral Dissertation, Univ. of Wisconsin, 1966.

[23] U. Merzbach, Development of Modern Algebraic Concepts from Leibniz to Dedekind, Doctoral Dissertation, Harvard Univ., 1964.

[24] G. A. Miller, History of the theory of groups, Collected Works, 3 vols., pp. 427–467, pp. 1–18, and pp. 1–15, Univ. of Illinois Press, 1935, 1938, and 1946.

[25] L. Novy, Origins of Modern Algebra, Noordhoff, 1973.

[26] O. J. Schmidt, Abstract Theory of Groups, W. H. Freeman & Co., 1966. (Translation by F. Holling and J. B. Roberts of the 1916 Russian edition.)

[27] J.-A. de Séguier, Théorie des Groupes Finis. Eléments de la Théorie des Groupes Abstraits, Gauthier Villars, Paris, 1904.

[28] L. A. Shemetkov, Two directions in the development of the theory of non-simple finite groups, Russ. Math. Surv., 30 (1975) 185–206.

[29] R. Silvestri, Simple groups of finite order in the nineteenth century, Arch. Hist. Ex. Sc., 20 (1979) 313–356.

[30] J. Tarwater, J. T. White, C. Hall, and M. E. Moore (eds.), American Mathematical Heritage: Algebra and Applied Mathematics, Texas Tech. Press, 1981. Has articles by Feit, Fuchs, and MacLane on the history of finite groups, abelian groups, and abstract algebra, respectively.

[31] B. L. Van der Waerden, Die Algebra seit Galois, Jahresbericht der Deutsch. Math. Ver., 68 (1966) 155–165.

[32] W. C. Waterhouse, The early proofs of Sylow's theorem, Arch. Hist. Ex. Sc., 21 (1979/80) 279–290.

[33] H. Wussing, The Genesis of the Abstract Group Concept, M.I.T. Press, 1984. (Translation by A. Shenitzer of the 1969 German edition.)

NOTES

Reciprocity in Ramanujan's Sum

KENNETH R. JOHNSON
North Dakota State University
Fargo, ND 58105

Reciprocity laws have held a special place in the theory of numbers ever since Gauss, at the age of nineteen, proved the first number-theoretic reciprocity law, the law of quadratic reciprocity. Since that time, many reciprocity laws have been discovered. Following a general discussion of reciprocity laws, we present a recently discovered reciprocity law for a well-known function of number theory, Ramanujan's sum [5]. The proofs of the reciprocity law are elementary and use concepts normally presented in an undergraduate number theory course. This is but another example demonstrating that once a fact is known, it is often not too difficult to prove.

Reciprocity laws

Basically, a reciprocity law is a relationship involving a function of two variables which allows one to interchange the arguments of the variables. As a simple example, if

$$F(x,y)=\frac{x}{y},$$

a reciprocity law for F might be stated:

$$F(x,y)\,F(y,x)=1.$$

The law of quadratic reciprocity, Gauss' "gem of arithmetic," was known to Euler and Legendre although they were unable to provide a proof of it before Gauss' remarkable accomplishment. Gauss eventually provided six different proofs of the theorem and there are currently many known proofs as the title of [4] suggests. With a little background, the law is simple to state.

Let q be an odd prime. A number a is called a quadratic residue mod q if there exists an integer n such that

$$a\equiv n^2 \pmod q.$$

If no such integer exists, a is called a quadratic nonresidue. Thus, for example, 23 is a quadratic residue mod 29 since $23\equiv 9^2 \pmod{29}$, while 11 is a quadratic nonresidue mod 29 since there are no squares congruent to 11 (mod 29). As it happens, exactly half of the numbers $1,2,\dots,q-1$ are quadratic residues. The Legendre symbol is defined by:

DEFINITION. If q is an odd prime, then the symbol $\left(\frac{a}{q}\right)$ is given the value $+1$ if a is a quadratic residue mod q and $\left(\frac{a}{q}\right)$ is given the value of -1 if a is a quadratic nonresidue mod q. Thus, in our example $\left(\frac{23}{29}\right)=+1$ while $\left(\frac{11}{29}\right)=-1$.

For odd primes p and q, the law of quadratic reciprocity is an elegant relationship between the Legendre symbols $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$. It can be stated:

THEOREM. *For odd primes p and q,*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right)=(-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

Thus $\left(\frac{p}{q}\right)=\left(\frac{q}{p}\right)$ unless both $\frac{p-1}{2}$ and $\frac{q-1}{2}$ are odd, i.e., $p\equiv q\equiv 3 \pmod 4$. In the example above, since $\left(\frac{23}{29}\right)=+1$ and $29\equiv 1 \pmod 4$, we know that 29 is also a quadratic residue mod 23. This is confirmed by noting that $29\equiv 6\equiv 11^2 \pmod{23}$. Many reciprocity laws been discovered which relate the variables of other number theoretic functions. In TABLE 1, a list of several other reciprocity laws and appropriate references are provided. The function $\mathrm{Sa}(x)$ is the well-known sawtooth function of period 1,

$$\mathrm{Sa}(x)=\begin{cases} x-[x]-1/2 & \text{if } x \text{ is not an integer}\\ 0 & \text{if } x \text{ is an integer,}\end{cases}$$

with $[x]$ the greatest integer not exceeding x. The function $\zeta(s)$ is the famous Riemann zeta function, defined for all complex s with $|s|>1$ by $\zeta(s)=\sum_{n=1}^{\infty}n^{-s}$. The first two functions in the table arise in the study of theta-functions (as well as elsewhere) while the third function and its reciprocity law were recently discovered in work connected with the Riemann zeta function.

Function	Defined by	Reciprocity law	Reference
Dedekind Sums	$s(h,k)=\sum_{j=1}^{k} Sa\left(\frac{hj}{k}\right)Sa\left(\frac{j}{k}\right)$	$s(k,h)+s(h,k)=-\frac{1}{4}+\frac{1}{12}\left(\frac{h}{k}+\frac{1}{hk}+\frac{k}{h}\right)$	[6]
Gauss Sums	$G(h;k)=\sum_{j=1}^{k}\exp\left(\frac{2\pi ihj^2}{k}\right)$	$G(h;k)=\sqrt{\frac{k}{h}}\left(\frac{1+i}{2}\right)\left(1+\exp\left(\frac{-\pi ihk}{2}\right)\right)\overline{G(k:h)}$	[2]
—	$H(s;z)=\sum_{n=1}^{\infty}n^{-s}\sum_{m=1}^{n}m^{-z}$	$H(s;z)+H(z;s)=\zeta(s)\zeta(z)+\zeta(s+z)$	[3]

TABLE 1

Arithmetic functions

In this section we provide the background from the theory of arithmetic functions necessary to state and prove the reciprocity law for Ramanujan's sum.

An arithmetic function is simply a function which is defined on the integers. Examples of these functions with number-theoretic importance that we will use are given in TABLE 2. (By $\sum_{d|n}$ we denote that the sum is to be taken over the positive divisors of n.)

Function	Defined by	Arithmetic significance
$d(n)$	$\sum_{d\|n} 1$	The number of divisors of n
$\sigma(n)$	$\sum_{d\|n} d$	The sum of the divisors of n
$\sigma_s(n)$	$\sum_{d\|n} d^s$	The sum of the sth powers of the divisors of n
$\phi(n)$	$\sum_{\substack{j=1\\(j,n)=1}}^{n} 1$	The number of integers not exceeding n which are relatively prime to n

TABLE 2

Given two arithmetic functions f and g, we can define a new arithmetic function $f*g$ by

$$f*g(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right).$$

This new function is called the Dirichlet product of f and g. For example, if we take the Dirichlet product of the functions $i(n) = n$ and $1(n) = 1$ for all n, we have

$$\begin{aligned} i*1(n) &= \sum_{d|n} i(d)1\left(\frac{n}{d}\right) \\ &= \sum_{d|n} d \\ &= \sigma(n). \end{aligned}$$

With this product, the function

$$\delta(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases}$$

serves as the identity. That is, for any arithmetic function f, $f*\delta = f$.

If f and g are functions such that $f*g = \delta$, we say that g is f inverse (and vice versa), and write f^{-1}. Although it is perhaps not obvious, only arithmetic functions f such that $f(1) \neq 0$ have Dirichlet product inverses. One inverse function that is prominent in the theory of numbers is the function which is the inverse of the function $1(n) = 1$ for all n, mentioned above. The function 1^{-1} is called the Möbius function and is denoted by $\mu(n)$. Although it is not now easy to see what values μ assumes for arbitrary n, we can find its value on prime powers fairly easily, directly from its definition as 1^{-1}. Since μ is defined by $\mu*1(n) = \delta(n)$, we see that

$$\begin{aligned} 1 = \delta(1) = \mu*1(1) &= \sum_{d|1} \mu(d)1\left(\frac{1}{d}\right) \\ &= \mu(1). \end{aligned}$$

Now if p is prime,

$$\begin{aligned} 0 = \delta(p) = \mu*1(p) &= \sum_{d|p} \mu(d)1\left(\frac{p}{d}\right) \\ &= \mu(1) + \mu(p) \end{aligned}$$

so $\mu(p) = -1$. Continuing, we see that

$$\begin{aligned} 0 = \delta(p^2) = \mu*1(p^2) &= \sum_{d|p^2} \mu(d)1\left(\frac{p^2}{d}\right) \\ &= \mu(1) + \mu(p) + \mu(p^2), \end{aligned}$$

so $\mu(p^2) = 0$. Using induction, it is now easy to show that $\mu(p^a) = 0$ for all $a > 1$.

An arithmetic function f, not identically zero, is called multiplicative if $f(m)f(n) = f(mn)$ whenever m and n are relatively prime.

All of the functions which we have discussed in this section are multiplicative. There are many advantages gained in working with multiplicative functions. For example, although it would be cumbersome (at best) to sum all of the divisors of 630, we can quite simply compute $\sigma(630)$ using the fact that σ is multiplicative and $630 = 2 \cdot 3^2 \cdot 5 \cdot 7$. Since the prime factors are relatively prime,

$$\begin{aligned} \sigma(630) &= \sigma(2)\,\sigma(9)\,\sigma(5)\,\sigma(7) \\ &= 3 \cdot 13 \cdot 6 \cdot 8 \\ &= 1872. \end{aligned}$$

The Möbius function is also multiplicative and since we know that for any prime p,

$$\mu(p^a) = \begin{cases} -1 & \text{if } a = 1 \\ 0 & \text{if } a > 1 \end{cases}$$

we can see that for any n:

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n \text{ is divisible by a square} \\ (-1)^k & \text{where } k \text{ is the number of distinct prime factors of } n. \end{cases}$$

As you can see, the behavior of a multiplicative function is completely determined by its behavior when the argument is a prime power. It is a common technique to show that two functions are equal by first showing that each is multiplicative and then showing that they agree on prime powers. In this way, since the functions i, μ, and ϕ are all multiplicative, we can show that $\phi = i^*\mu$. Remembering that $\phi(p^a)$ is the number of integers not exceeding p^a (of which there are p^a) minus the number of these which are divisible by p (namely, $p, 2p, \ldots, p^2, \ldots, p^a$, of which there are p^{a-1}), we have for any prime p,

$$\phi(p^a) = p^a - p^{a-1}. \tag{1}$$

It is easy to check that $i^*\mu(p^a)$ assumes this same value and thus $\phi(n) = i^*\mu(n)$ for all n.

Ramanujan's sum

Ramanujan's sum, $c(n, m)$, is the sum of the nth powers of the mth primitive roots of unity. That is,

$$c(n, m) = \sum_{\substack{j=1 \\ (j,m)=1}}^{m} \exp\left(\frac{2\pi ijn}{m}\right).$$

This function has been studied extensively for the last 70 years. The function satisfies an extremely large number of elegant arithmetic identities such as Ramanujan's own

$$\sigma(n) = \frac{\pi^2 n}{6} \sum_{m=1}^{\infty} \frac{c(n, m)}{m^2},$$

$$\phi(n) = \frac{6n}{\pi^2} \sum_{m=1}^{\infty} \frac{c(n, m)\mu(m)}{\phi_2(m)},$$

and

$$0 = \sum_{m=1}^{\infty} \frac{c(n, m)}{m}. \quad [7]$$

In his first paper on the function [7], Ramanujan presented the, perhaps surprising, relation between $c(n, m)$ and two multiplicative functions of one variable. Namely,

$$c(n, m) = \sum_{d|(n,m)} d\mu\left(\frac{m}{d}\right). \tag{2}$$

In a manner analogous to the single variable case, we say that a function, f, of two variables is multiplicative if $f(n, m)f(j, k) = f(nj, mk)$ whenever $(nm, jk) = 1$. We can use (2) to show that $c(n, m)$ is an example of a multiplicative function of two variables. To this end, choose any integers n, m, j, and k with $(nm, jk) = 1$. Then

$$c(n, m)c(j, k) = \sum_{d|(n,m)} d\mu\left(\frac{m}{d}\right) \sum_{g|(j,k)} g\mu\left(\frac{k}{g}\right)$$

$$= \sum_{d\|(n,m)} \sum_{g\|(j,k)} dg\mu\left(\frac{m}{d}\right)\mu\left(\frac{k}{g}\right)$$

$$= \sum_{dg\|(n,m)(j,k)} (dg)\mu\left(\frac{mk}{dg}\right), \quad \text{since } \left(\frac{m}{d}, \frac{k}{g}\right) = 1$$

and μ is multiplicative. Setting $r = dg$ and noting that $(n, m)(j, k) = (nj, mk)$ since $(nm, jk) = 1$, this last sum becomes

$$= \sum_{r\|(nj,mk)} r\mu\left(\frac{mk}{r}\right)$$

$$= c(nj, mk).$$

Since $c(n, m)$ is a multiplicative function of two variables it suffices to determine its behavior when n and m are powers of the same prime. Using (2), we find

$$c(p^a, p^b) = \begin{cases} 0 & \text{if } b - a > 1 \\ -p^{b-1} & \text{if } b - a = 1 \\ p^b - p^{b-1} & \text{if } b - a < 1. \end{cases} \tag{3}$$

Using this information, we can show that Ramanujan's sum satisfies the Hölder relation:

$$c(n, m) = \frac{\phi(m)\mu(m/(n,m))}{\phi(m/(m,n))}. \tag{4}$$

We merely verify that the right-hand side of (4) is also a multiplicative function of two variables and then check that both sides agree when m and n are powers of the same prime.

With the following definition we may state and prove the reciprocity law for Ramanujan's sum.

DEFINITION. For any integer n, let $\bar{n}$ denote the largest square-free divisor (sometimes called the **core**) of n and let $n^* = n/\bar{n}$.

Using this notation, (1), and the multiplicativity of ϕ we see that

$$\phi(n) = n^*\phi(\bar{n}). \tag{5}$$

In light of (3), which says $c(p^a, p^b) = 0$ unless $p^{b-1} | p^a$, the only nonzero values of Ramanujan's sum will be of the form $c(nm^*, m)$.

The reciprocity law for Ramanujan's sum which allows us to exchange the arguments when both $c(n, m)$ and $c(m, n)$ are nonzero may be stated

$$\frac{\mu(\bar{m})c(nm^*, m)}{m^*} = \frac{\mu(\bar{n})c(mn^*, n)}{n^*}. \tag{6}$$

The two proofs of this statement using the standard techniques from the study of arithmetic functions are as follows.

Proof 1. Denote the left-hand side of (6) by $F(n, m)$ and suppose that $(nm, jk) = 1$, then

$$F(n,m)F(j,k) = \frac{\mu(\bar{m})c(nm^*, m)}{m^*}\,\frac{\mu(\bar{k})c(jk^*, k)}{k^*}$$

$$= \frac{\mu(\overline{mk})c(nm^*jk^*, mk)}{m^*k^*},$$

since these functions are all multiplicative, and

$$F(n,m)F(j,k) = \frac{\mu(\overline{mk})c(nj(mk)^*, mk)}{(mk)^*},$$

since $m^*k^* = (mk)^*$, so that

$$F(n,m)F(j,k) = F(nj, mk).$$

Thus F is multiplicative, and so it suffices to prove (6) (which says $F(n,m) = F(m,n)$) when n and m are powers of the same prime. To this end, let $n = p^a$ and $m = p^b$. It is easy to see that $F(p,1) = F(1,p) = 1$, so that we may assume that both a and b are positive integers, then

$$F(n,m) = F(p^a, p^b) = \frac{\mu(p)\,c(p^{a+b-1}, p^b)}{p^{b-1}}$$

$$= \frac{-1}{p^{b-1}}(p^b - p^{b-1}),$$

using (3). Thus $F(n,m) = 1-p$. Since $F(n,m) = 1-p$ is independent of a and b, the same argument shows that $F(m,n) = 1-p$ thus completing the proof.

The Hölder relation provides another simple proof of the reciprocity law with the aid of the following lemma.

LEMMA. *For all m and n,*

$$\mu(\bar{n})\,c(m,\bar{n}) = \mu(\bar{m})\,c(n,\bar{m}).$$

Proof. By (2),

$$\mu(\bar{n})\,c(m,\bar{n}) = \mu(\bar{n}) \sum_{d|(m,\bar{n})} d\mu(\bar{n}/d)$$

$$= \mu(\bar{m}) \sum_{d|(\bar{m},\bar{n})} d\mu(\bar{n}/d)\mu(\bar{n})\mu(\bar{m}),$$

since $\mu^2(\bar{m}) = 1$ and $(\bar{m},\bar{n}) = (m,\bar{n})$. But $(\bar{n}/d, \bar{m}) = 1$ and since μ is multiplicative, this becomes

$$\mu(\bar{m}) \sum_{d|(\bar{m},\bar{n})} d\mu(\bar{n}\bar{m}/d)\mu(\bar{n}).$$

Since $(\bar{m}/d, \bar{n}) = 1$ also, $(\bar{m},\bar{n}) = (\bar{m}, n)$, and $\mu^2(\bar{n}) = 1$, we see that this equals

$$\mu(\bar{m}) \sum_{d|(\bar{m},n)} d\mu(\bar{m}/d) = \mu(\bar{m})\,c(n,\bar{m}).$$

The second proof of the reciprocity law for Ramanujan's sum is now simple.

Proof 2. Beginning with the Hölder formula, we have

$$c(nm^*, m) = \frac{\phi(m)\,\mu(m/(m, nm^*))}{(m/(m,nm^*))}.$$

Since (m, nm^*) is equal to $m^*\ (\bar{m}, n)$ and $m/m^* = \bar{m}$, this becomes

$$\frac{\phi(m)\,\mu(\bar{m}/(\bar{m},n))}{\phi(\bar{m}/(\bar{m},n))}.$$

and after using (5) we see that this equals

$$\frac{m^*\phi(\bar{m})\,\mu(\bar{m}/(\bar{m},n))}{\phi(\bar{m}/(\bar{m},n))},$$

which we recognize by the Hölder relation as

$$m^*c(n,\bar{m}).$$

Thus,

$$\frac{\mu(\bar{m})\, c(nm^*, m)}{m^*} = \mu(\bar{m})\, c(n, \bar{m}).$$

In exactly the same manner we obtain

$$\frac{\mu(\bar{n})\, c(mn^*, n)}{n^*} = \mu(\bar{n})\, c(m, \bar{n}).$$

Upon application of the lemma, the proof is complete.

There are many generalizations of Ramanujan's sum, one of the most interesting is presented in a paper by Anderson and Apostol [**1**] which the reader may wish to explore on his or her own.

References

[1] D. Anderson and T. Apostol, The evaluation of Ramanujan's sum and generalizations, Duke Math. J., 20 (1952) 211–216.

[2] T. Apostol, Introduction to Analytic Number Theory, Springer-Verlag, New York, 1976.

[3] T. Apostol and T. Vu, Dirichlet series related to the Riemann zeta function, J. Number Theory, 19 (1984) 85–102.

[4] M. Gerstenhaber, The 152nd proof of the law of quadratic reciprocity, Amer. Math. Monthly, 70 (1963) 397–398.

[5] K. Johnson, A reciprocity law for Ramanujan sums, Pacific J. Math., 98 (1983) 99–105.

[6] H. Rademacher and E. Grosswald, Dedekind Sums, Mathematical Association of America, Washington, D.C., 1972.

[7] S. Ramanujan, On certain trigonometrical sums and their applications in the theory of numbers, Cambridge Phil. Trans., 22 (1918) 259-276.

Fields for Which the Principal Axis Theorem Is Valid

A. CHARNOW and E. CHARNOW

California State University—Hayward

Hayward, CA 94542

One version of the Principal Axis Theorem asserts that any symmetric matrix with entries in the field $\mathbb{R}$ of real numbers is similar over $\mathbb{R}$ to a diagonal matrix. Clearly this statement does not remain true if the field in question is the field $\mathbb{Q}$ of rational numbers; neither does it hold if the complex field $\mathbb{C}$ replaces $\mathbb{R}$: the matrix $\begin{pmatrix} 2 & i \\ i & 0 \end{pmatrix}$, for example, is not similar over $\mathbb{C}$ to a diagonal matrix. The primary purpose of this note is to describe the necessary and sufficient field properties of the real numbers that make this theorem valid.

We begin by assigning the name "principal axis field" to a field K if any symmetric matrix with entries in K is similar over K to a diagonal matrix. It will turn out that any principal axis field must be formally real; i.e., whenever $\sum_{i=1}^{n} a_i^2 = 0$, $a_i \in K$, then $a_i = 0$ for $i = 1, \ldots, n$. A formally real field K is given the additional label of real closed if no proper algebraic extension of K is formally real. The field of all real algebraic numbers is an example—less obvious than $\mathbb{R}$ itself—of a real closed field [**2**, Corollary, p. 278]. In [**5**, Remarks, p. 379] it is observed that the usual proof of the Principal Axis Theorem can be carried out in any real closed field, so a real

Thus,

$$\frac{\mu(\overline{m})\,c(nm^*,m)}{m^*} = \mu(\overline{m})\,c(n,\overline{m}).$$

In exactly the same manner we obtain

$$\frac{\mu(\bar{n})\,c(mn^*,n)}{n^*} = \mu(\bar{n})\,c(m,\bar{n}).$$

Upon application of the lemma, the proof is complete.

There are many generalizations of Ramanujan's sum, one of the most interesting is presented in a paper by Anderson and Apostol [**1**] which the reader may wish to explore on his or her own.

References

[1] D. Anderson and T. Apostol, The evaluation of Ramanujan's sum and generalizations, Duke Math. J., 20 (1952) 211–216.
[2] T. Apostol, Introduction to Analytic Number Theory, Springer-Verlag, New York, 1976.
[3] T. Apostol and T. Vu, Dirichlet series related to the Riemann zeta function, J. Number Theory, 19 (1984) 85–102.
[4] M. Gerstenhaber, The 152nd proof of the law of quadratic reciprocity, Amer. Math. Monthly, 70 (1963) 397–398.
[5] K. Johnson, A reciprocity law for Ramanujan sums, Pacific J. Math., 98 (1983) 99–105.
[6] H. Rademacher and E. Grosswald, Dedekind Sums, Mathematical Association of America, Washington, D.C., 1972.
[7] S. Ramanujan, On certain trigonometrical sums and their applications in the theory of numbers, Cambridge Phil. Trans., 22 (1918) 259-276.

Fields for Which the Principal Axis Theorem Is Valid

A. CHARNOW and E. CHARNOW
California State University—Hayward
Hayward, CA 94542

One version of the Principal Axis Theorem asserts that any symmetric matrix with entries in the field $\mathbb{R}$ of real numbers is similar over $\mathbb{R}$ to a diagonal matrix. Clearly this statement does not remain true if the field in question is the field $\mathbb{Q}$ of rational numbers; neither does it hold if the complex field $\mathbb{C}$ replaces $\mathbb{R}$: the matrix $\begin{pmatrix} 2 & i \\ i & 0 \end{pmatrix}$, for example, is not similar over $\mathbb{C}$ to a diagonal matrix. The primary purpose of this note is to describe the necessary and sufficient field properties of the real numbers that make this theorem valid.

We begin by assigning the name "principal axis field" to a field K if any symmetric matrix with entries in K is similar over K to a diagonal matrix. It will turn out that any principal axis field must be formally real; i.e., whenever $\Sigma_{i=1}^{n} a_i^2 = 0$, $a_i \in K$, then $a_i = 0$ for $i = 1, \ldots, n$. A formally real field K is given the additional label of real closed if no proper algebraic extension of K is formally real. The field of all real algebraic numbers is an example—less obvious than $\mathbb{R}$ itself—of a real closed field [**2**, Corollary, p. 278]. In [**5**, Remarks, p. 379] it is observed that the usual proof of the Principal Axis Theorem can be carried out in any real closed field, so a real closed field is always a principal axis field. (To obtain this result, one could also apply Tarski's meta-mathematical principle which says that any "elementary" statement of algebra valid in the real field is valid in every real closed field [**3**, pp. 323–324].) Is a principal axis field necessarily real closed? The answer is no, and the following description of a counterexample will shed light on the precise nature of principal axis fields.

With $\mathbb{Q}$ still denoting the rational numbers, let Ω be the algebraic closure of $\mathbb{Q}$; i.e., let Ω be the field of all algebraic numbers. Recall that two elements of Ω are called conjugate over $\mathbb{Q}$ if they are the roots of the same irreducible polynomial in $\mathbb{Q}[x]$. An element α of Ω is said to be totally real if α and all its conjugates (over $\mathbb{Q}$) are real. Thus, $\sqrt{2}$ is totally real while $\sqrt[3]{2}$ is not. Let S be the collection of all such totally real members of Ω. It is not hard to see that S is a subfield of Ω. The argument runs as follows. Assume α_1 and β_1 are totally real. Let f and g be the minimal polynomials (over $\mathbb{Q}$) of α_1 and β_1, respectively. Then $f(x)=\prod_{i=1}^{n}(x-\alpha_i)$ and $g(x)=\prod_{j=1}^{m}(x-\beta_j)$, where all the α_i and β_j are real. Clearly, $(-1)^n f(-x)$, with real roots $-\alpha_1,\dots,-\alpha_n$, is the minimal polynomial of $-\alpha_1$, so $-\alpha_1\in S$. If $\alpha_1\neq 0$, then all $\alpha_i\neq 0$, and the polynomial $x^n f(1/x)$ has rational coefficients as well as real roots $1/\alpha_1,\dots,1/\alpha_n$. Since the minimal polynomial of $1/\alpha_1$ must divide $x^n f(1/x)$, it follows that $1/\alpha_1\in S$. This same kind of argument, with the polynomials $h(x)=\prod_{j=1}^{m}\prod_{i=1}^{n}(x-\alpha_i-\beta_j)$ and $p(x)=\prod_{j=1}^{m}\prod_{i=1}^{n}(x-\alpha_i\beta_j)$, may be used to show that $\alpha_1+\beta_1$ and $\alpha_1\beta_1$ are totally real: both h and p have rational coefficients [**6**, Cor. 3.12, p. 39], and their roots are clearly all in $\mathbb{R}$. Therefore, this set S of totally real members of Ω is a field.

If S were real closed, then every polynomial of odd degree in $S[x]$ would have a root in S [**2**, Th. 4, p. 274]. Since the polynomial x^3-2 has no totally real root, the field S cannot be real closed.

What remains to be done is to verify that S is a principal axis field. This task will be facilitated by first establishing a sufficient condition for a field to have the principal axis property:

LEMMA. *Let Ω be the algebraic closure of a formally real field K, and let K_R be the intersection of all the real closed subfields of Ω that contain K. If $K=K_R$, then K is a principal axis field.*

Proof. Suppose A is a symmetric matrix with entries in K. Let f, in $K[x]$, be the minimal polynomial of A. There exists at least one real closed subfield F of Ω such that $K\subseteq F$ [**2**, Th. 3, p. 274]. Since F is a principal axis field, A is similar to a diagonal matrix over F. Thus, the minimal polynomial of A over F has the form $\prod_{i=1}^{n}(x-\lambda_i)$, where $\lambda_1,\dots,\lambda_n$ are distinct members of F [**1**, Th. 6, p. 204]. However, the minimal polynomial of A over F is f (see discussion in [**1**, p. 192]). Therefore, $f(x)=\prod_{i=1}^{n}(x-\lambda_i)$. To complete this proof, we need only show that the $\lambda_1,\dots,\lambda_n$ are in $K=K_R$. Let L be any real closed subfield of Ω containing K. Since A is similar to a diagonal matrix over L, any eigenvalue of A in Ω is already in L. From $\lambda_i\in L$, it follows that $\lambda_i\in K_R$, $i=1,\dots,n$.

Let us now apply this lemma to the field S. Clearly, this field of totally real numbers is formally real, and Ω (the algebraic closure of $\mathbb{Q}$) is an algebraic closure of S. We need to show that $S_R\subseteq S$, where S_R is the intersection of all the real closed subfields of Ω containing S. Assume $\alpha\in S_R$. Let L be the subfield of Ω consisting of all the real algebraic numbers. We know that L is real closed, and it is clear that $S\subseteq L$. Therefore, $S_R\subseteq L$, so α is real. Now let β (in Ω) be any conjugate (over $\mathbb{Q}$) of α. There is a $\mathbb{Q}$-isomorphism of $\mathbb{Q}(\alpha)$ onto $\mathbb{Q}(\beta)$ which sends α onto β. Since Ω is an algebraic closure of both $\mathbb{Q}(\alpha)$ and $\mathbb{Q}(\beta)$, this isomorphism can be extended to an automorphism Φ of Ω [**2**, Th. 2, p. 145]. Let $L'=\Phi^{-1}(L)=\{t\colon t\in\Omega, \Phi(t)\in L\}$. In order to complete this proof that $\alpha\in S$, we must show that β is real. This can be accomplished by showing $\beta \in L$—equivalently, by showing $\alpha \in L'$. Thus, the problem has been reduced to proving that L' is one of the subfields defining S_R. As L' is an isomorphic copy of the real closed subfield L, it is easy to see that L' is itself a real closed subfield of Ω. Moreover, if $t \in S$, then $\Phi(t)$ must be a conjugate (over $\mathbb{Q}$) of t. Therefore, $\Phi(t) \in S \subseteq L$; equivalently, $t \in L'$, so $S \subseteq L'$, and we are done.

The following theorem establishes the criteria of the lemma as necessary as well as sufficient:

THEOREM. *Let K be a field with algebraic closure Ω, and let K_R be the intersection of all the real closed subfields of Ω that contain K. Then K is a principal axis field if and only if K is formally real and $K = K_R$.*

Proof. In view of the lemma, we need only deal with the hypothesis that K is a principal axis field. First we shall prove that K is formally real.

Let $a_i \in K$, $i = 1, \dots, n$, and let A be the $n+1$ by $n+1$ symmetric matrix:

$$A = \begin{pmatrix} 0 & a_1 & \cdot & \cdot & \cdot & a_n \\ a_1 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_n & 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix}.$$

We first prove, by induction on n, that the characteristic polynomial of A is $x^{n+1} - (a_1^2 + a_2^2 + \cdots + a_n^2)x^{n-1}$. This statement is trivial if $n = 1$. Assuming the induction hypothesis, let B be the $n+2$ by $n+2$ symmetric matrix:

$$B = \begin{pmatrix} 0 & a_1 & \cdot & \cdot & \cdot & a_{n+1} \\ a_1 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n+1} & 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix}.$$

Then the characteristic polynomial of B is:

$$g(x) = \det \begin{pmatrix} x & -a_1 & \cdot & \cdot & \cdot & -a_{n+1} \\ -a_1 & x & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -a_{n+1} & 0 & \cdot & \cdot & \cdot & x \end{pmatrix}.$$

Expanding by cofactors of the second row, we obtain $g(x) = -a_1^2 x^n + x f(x)$, where f is the characteristic polynomial of:

$$C = \begin{pmatrix} 0 & a_2 & \cdot & \cdot & \cdot & a_{n+1} \\ a_2 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n+1} & 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix}.$$

Application of the induction hypothesis to C yields

$$g(x) = -a_1^2 x^n + x\left[x^{n+1} - \left(a_2^2 + \cdots + a_{n+1}^2\right)x^{n-1}\right] = x^{n+2} - \left(a_1^2 + \cdots + a_{n+1}^2\right)x^n,$$

as desired.

Now suppose $\sum_{i=1}^{n} a_i^2 = 0$. Then the characteristic polynomial of A is x^{n+1}, so the minimal polynomial of A is some power of x. As A is similar to a diagonal matrix over K, the minimal polynomial of A must be x [**1**, Th. 6, p. 204]. Therefore, $A = 0$, $a_i = 0$ for $i = 1, \ldots, n$, and K is formally real. (Note: since K is a principal axis field, the symmetric matrix A is similar to a diagonal matrix over K. Consequently, all the eigenvalues of A in Ω are in K, or $(a_1^2 + \cdots + a_n^2)^{1/2} \in K$. We shall later have reason to refer to this fact.)

It remains to be shown that $K_R \subseteq K$, so let $\lambda \in K_R$. Since K is also formally real, λ must be the eigenvalue of some symmetric matrix with entries in K [**4**, Satz 3.3, p. 231]. By hypothesis, such a matrix is similar to a diagonal matrix over K. Therefore, $\lambda \in K$.

We conclude with the observation that if K is a principal axis field, and if A is an n by n symmetric matrix with entries in K, then there exists an orthogonal matrix P (i.e., $P^{-1} = P^t$, the transpose of P) with entries in K such that $P^{-1}AP$ is diagonal. We know that K has to be formally real, so K must be an ordered field [**2**, Cor. 2, p. 274]. An inner product can be defined on K^n, in the usual way, and it will satisfy all the axioms of a real inner product. By the remark at the end of paragraph three in the proof of the theorem, square roots of sums of squares exist in K. Hence, we can define a norm on K^n in the standard way. The Gram-Schmidt process will produce an orthonormal basis of each eigenspace of A, and, since A is similar to a diagonal matrix over K, the union of these bases will provide an orthonormal basis for K^n [**1**, Th. 2, p. 187]. Then P may be chosen as the matrix whose columns are these basis vectors.

References

[1] K. Hoffman and R. Kunze, Linear Algebra, 2nd ed., Prentice Hall, Englewood Cliffs, 1971.
[2] N. Jacobson, Lectures in Abstract Algebra, Volume III, Van Nostrand, New York, 1964.
[3] ______, Basic Algebra I, W. H. Freeman, San Francisco, 1974.
[4] F. Krakowski, Eigenwerte und Minimalpolynome symmetrischer Matrizen in kommutativen Körpern, Comment. Math. Helv., 32 (1958), 224–240.
[5] S. Lang, Algebra, Addison-Wesley, Reading, Massachusetts, 1965.
[6] H. Pollard and H. Diamond, The Theory of Algebraic Numbers, 2nd ed., MAA, Washington, D.C., 1975.

On the Power Sums of the Roots of a Polynomial

MARTIN L. BROCK
2 Old Mamaroneck Rd, #4
White Plains, NY 10605

The sums of the kth powers of the roots of a polynomial are investigated here. Results on these are usually obtained from the theory of symmetric functions (van der Waerden [**3**]) or from Galois theory (Marcus [**2**]). Here, several results are obtained quickly and simply by the use of a theorem of linear algebra. That theorem describes the Jordan form of a matrix.

Let D be an integral domain and Q its quotient field. Let n be in $\mathbf{Z}^+$, f be in $D[x]$, f be monic and $\deg(f) = n$. So

$$f = x^n + a_{n-1}x^{n-1} + \cdots + a_0,$$

for some $a_0, \ldots, a_{n-1}$ in D. Let E be an extension field of the splitting field of f over Q. Let the n roots of f in E, respecting multiplicities, be $r_1, \ldots, r_n$. Let k be in $\mathbf{Z}^+$ and let $s_k(f)$ be the sum polynomial of A must be x [**1**, Th. 6, p. 204]. Therefore, $A=0$, $a_i=0$ for $i=1,\dots,n$, and K is formally real. (Note: since K is a principal axis field, the symmetric matrix A is similar to a diagonal matrix over K. Consequently, all the eigenvalues of A in Ω are in K, or $(a_1^2+\cdots+a_n^2)^{1/2}\in K$. We shall later have reason to refer to this fact.)

It remains to be shown that $K_R\subseteq K$, so let $\lambda\in K_R$. Since K is also formally real, λ must be the eigenvalue of some symmetric matrix with entries in K [**4**, Satz 3.3, p. 231]. By hypothesis, such a matrix is similar to a diagonal matrix over K. Therefore, $\lambda\in K$.

We conclude with the observation that if K is a principal axis field, and if A is an n by n symmetric matrix with entries in K, then there exists an orthogonal matrix P (i.e., $P^{-1}=P^t$, the transpose of P) with entries in K such that $P^{-1}AP$ is diagonal. We know that K has to be formally real, so K must be an ordered field [**2**, Cor. 2, p. 274]. An inner product can be defined on K^n, in the usual way, and it will satisfy all the axioms of a real inner product. By the remark at the end of paragraph three in the proof of the theorem, square roots of sums of squares exist in K. Hence, we can define a norm on K^n in the standard way. The Gram-Schmidt process will produce an orthonormal basis of each eigenspace of A, and, since A is similar to a diagonal matrix over K, the union of these bases will provide an orthonormal basis for K^n [**1**, Th. 2, p. 187]. Then P may be chosen as the matrix whose columns are these basis vectors.

References

[1] K. Hoffman and R. Kunze, Linear Algebra, 2nd ed., Prentice Hall, Englewood Cliffs, 1971.
[2] N. Jacobson, Lectures in Abstract Algebra, Volume III, Van Nostrand, New York, 1964.
[3] ______, Basic Algebra I, W. H. Freeman, San Francisco, 1974.
[4] F. Krakowski, Eigenwerte und Minimalpolynome symmetrischer Matrizen in kommutativen Körpern, Comment. Math. Helv., 32 (1958), 224–240.
[5] S. Lang, Algebra, Addison-Wesley, Reading, Massachusetts, 1965.
[6] H. Pollard and H. Diamond, The Theory of Algebraic Numbers, 2nd ed., MAA, Washington, D.C., 1975.

On the Power Sums of the Roots of a Polynomial

MARTIN L. BROCK
2 Old Mamaroneck Rd, #4
White Plains, NY 10605

The sums of the kth powers of the roots of a polynomial are investigated here. Results on these are usually obtained from the theory of symmetric functions (van der Waerden [**3**]) or from Galois theory (Marcus [**2**]). Here, several results are obtained quickly and simply by the use of a theorem of linear algebra. That theorem describes the Jordan form of a matrix.

Let D be an integral domain and Q its quotient field. Let n be in $\mathbf{Z}^+$, f be in $D[x]$, f be monic and $\deg(f)=n$. So

$$f=x^n+a_{n-1}x^{n-1}+\cdots+a_0,$$

for some $a_0,\dots,a_{n-1}$ in D. Let E be an extension field of the splitting field of f over Q. Let the n roots of f in E, respecting multiplicities, be $r_1,\dots,r_n$. Let k be in $\mathbf{Z}^+$ and let $s_k(f)$ be the sum of the kth powers of the roots of f. That is,

$$s_k(f) = r_1^k + \cdots + r_n^k.$$

The quantity $s_k(f)$ will be investigated by linear algebra.

Recall the **companion matrix** of f, denoted by C_f. This is the following $n \times n$ matrix over D:

$$C_f = \begin{bmatrix} 0 & 1 & & & \\ & 0 & & 0 & \\ & & \ddots & 1 & \\ & 0 & & 0 & \\ & & & & 1 \\ -a_0 & & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix}$$

The reason for introducing this matrix is the following standard (and easily proven) fact (see Herstein [1, p. 265]):

$$\text{characteristic polynomial of } C_f = f.$$

Since the characteristic polynomial of C_f (namely, f) splits completely over E (by construction), C_f can be brought to Jordan form by matrices over E (see Herstein [1, p. 258]). That is, there is an $n \times n$ invertible matrix, P, over E and an $n \times n$ matrix in Jordan form, J, over E such that

$$PC_fP^{-1} = J.$$

Among other things, J is upper triangular and its main diagonal consists of the roots of the characteristic polynomial, f, with multiplicities respected. That is, the main diagonal consists of $r_1, \ldots, r_n$.

Using the last equation, letting k be in $\mathbf{Z}^+$, and taking powers and matrix traces give

$$\begin{aligned} \operatorname{tr}(J^k) &= \operatorname{tr}\left(\left(PC_fP^{-1}\right)^k\right) \\ &= \operatorname{tr}\left(PC_f^kP^{-1}\right) \\ &= \operatorname{tr}\left(C_f^k\right). \end{aligned}$$

The previously mentioned properties about J give

$$\begin{aligned} \operatorname{tr}(J^k) &= r_1^k + \cdots + r_n^k \\ &= s_k(f). \end{aligned}$$

Combining these last two results gives the main result of this note, as stated in the following theorem.

THEOREM. *For k in Z^+, we have*

$$s_k(f) = \operatorname{tr}\left(C_f^k\right).$$

Remarks.

1) Since C_f is immediately gotten from f, this theorem gives a simple way to calculate $s_k(f)$, the sum of the kth powers of the roots of f. The speed of this calculation depends on how fast one can multiply matrices over D. This can be compared with the algorithms offered by the theory of symmetric functions (see van der Waerden [3]).
2) Since the entries of C_f are only $0, 1, -a_0, \ldots, -a_{n-1}$, this theorem shows at once that $s_k(f)$ is in $\mathbf{Z}[a_0, \ldots, a_{n-1}]$, the ring generated by $a_0, \ldots, a_{n-1}$. In particular, it shows that $s_k(f)$ is in D. This can be compared with the usual proofs obtained from the theory of symmetric functions (see van der Waerden [3]) or Galois theory with the theory of algebraic integers (see

Marcus [2]).

3) The theorem can be used to generate identities valid in *any* ring. For example,

$$r_1^2+\cdots+r_n^2=(r_1+\cdots+r_n)^2-2(r_1r_2+\cdots+r_{n-1}r_n);$$

$$r_1^3+\cdots+r_n^3=(r_1+\cdots+r_n)^3-3(r_1+\cdots+r_n)(r_1r_2+\cdots+r_{n-1}r_n)$$
$$+3(r_1r_2r_3+\cdots+r_{n-2}r_{n-1}r_n).$$

The above results can be extended to the case $k=0$ or k a negative integer, when a_0 is nonzero. This would involve taking the inverse of the companion matrix. (The latter is easily obtained, however, by any of several methods, e.g., matrix adjoints or the Cayley-Hamilton Theorem. Also, it is not too hard to see that this inverse closely resembles the original companion matrix.)

References

[1] I. N. Herstein, Topics in Algebra, Xerox College Publishing, 1964.
[2] D. A. Marcus, Number Fields, Springer-Verlag, 1977.
[3] B. L. van der Waerden, Frederick Ungar Publishing Company, 1953.

On the Laplacian

JAMES WARD BROWN
University of Michigan—Dearborn
Dearborn, MI 48128

FRANK FARRIS
Santa Clara University
Santa Clara, CA 95053

In various applied mathematics courses one appearance of the Laplacian operator ∇^2 is in the study of heat distributions. If u is a heat distribution in space, then $\nabla^2 u=0$ if and only if u is a steady-state distribution, one that could be maintained indefinitely inside a box with suitable boundary conditions.

In rectangular coordinates the Laplacian of u is

$$\nabla^2 u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}.$$

Careful use of the chain rule gives

$$\nabla^2 u=\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}+\frac{\partial^2 u}{\partial z^2}$$

and

$$\nabla^2 u=\frac{\partial^2 u}{\partial \rho^2}+\frac{2}{\rho}\frac{\partial u}{\partial \rho}+\frac{1}{\rho^2\sin^2\phi}\frac{\partial^2 u}{\partial \theta^2}+\frac{1}{\rho^2}\frac{\partial^2 u}{\partial \phi^2}+\frac{\cot\phi}{\rho^2}\frac{\partial u}{\partial \phi}$$

as the correct formulas for the Laplacian in cylindrical and spherical coordinates [1].

Marcus [2]).

3) The theorem can be used to generate identities valid in *any* ring. For example,

$$r_1^2+\cdots+r_n^2=(r_1+\cdots+r_n)^2-2(r_1r_2+\cdots+r_{n-1}r_n);$$

$$r_1^3+\cdots+r_n^3=(r_1+\cdots+r_n)^3-3(r_1+\cdots+r_n)(r_1r_2+\cdots+r_{n-1}r_n)$$
$$+3(r_1r_2r_3+\cdots+r_{n-2}r_{n-1}r_n).$$

The above results can be extended to the case $k=0$ or k a negative integer, when a_0 is nonzero. This would involve taking the inverse of the companion matrix. (The latter is easily obtained, however, by any of several methods, e.g., matrix adjoints or the Cayley-Hamilton Theorem. Also, it is not too hard to see that this inverse closely resembles the original companion matrix.)

References

[1] I. N. Herstein, Topics in Algebra, Xerox College Publishing, 1964.
[2] D. A. Marcus, Number Fields, Springer-Verlag, 1977.
[3] B. L. van der Waerden, Frederick Ungar Publishing Company, 1953.

On the Laplacian

JAMES WARD BROWN
University of Michigan—Dearborn
Dearborn, MI 48128

FRANK FARRIS
Santa Clara University
Santa Clara, CA 95053

In various applied mathematics courses one appearance of the Laplacian operator ∇^2 is in the study of heat distributions. If u is a heat distribution in space, then $\nabla^2 u=0$ if and only if u is a steady-state distribution, one that could be maintained indefinitely inside a box with suitable boundary conditions.

In rectangular coordinates the Laplacian of u is

$$\nabla^2 u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}.$$

Careful use of the chain rule gives

$$\nabla^2 u=\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}+\frac{\partial^2 u}{\partial z^2}$$

and

$$\nabla^2 u=\frac{\partial^2 u}{\partial \rho^2}+\frac{2}{\rho}\frac{\partial u}{\partial \rho}+\frac{1}{\rho^2\sin^2\phi}\frac{\partial^2 u}{\partial \theta^2}+\frac{1}{\rho^2}\frac{\partial^2 u}{\partial \phi^2}+\frac{\cot\phi}{\rho^2}\frac{\partial u}{\partial \phi}$$

as the correct formulas for the Laplacian in cylindrical and spherical coordinates [1].

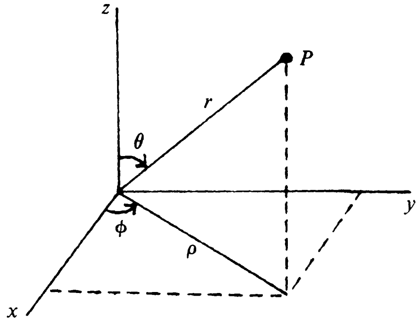


FIGURE 1

Since problems in two dimensions are often easier to solve than those in space, it is desirable to recover a two-dimensional Laplacian from the formulas above. In the first two instances this is straightforward: differentiations with respect to z may be ignored, leaving

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}$$

and

$$\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}$$

as the correct formulas for the Laplacian in rectangular and polar coordinates in the plane. Students readily accept the physical reason justifying this procedure: a steady-state heat distribution on a thin metal plate will also give a steady-state distribution when extended to space in such a way that it is independent of z. For example, $u(x, y)=xy$ could equally well describe the steady-state heat profile of a thin disk or of a cylinder.

A problem commonly arises, however, when students attempt the same reduction of the Laplacian in spherical coordinates. At first it seems reasonable to ignore differentiations with respect to ϕ, set $\phi=\pi/2$, and replace ρ with r. This gives

$$\frac{\partial^2 u}{\partial r^2}+\frac{2}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2},$$

apparently contradicting the formula above. Since the difference is a matter of a factor of 2, students may shrug away the contradiction as a misprint.

What is responsible for the incorrect coefficient of $\partial u/\partial r$? Students must examine more carefully what is entailed in ignoring differentiations with respect to ϕ. To do so is to pretend that one's two-dimensional heat distribution is actually a distribution in space which is independent of ϕ. The resulting heat distribution would be constant on the longitudes of the unit sphere and singular at the North Pole. Certainly the question of heat equilibrium for the two distributions is not the same. A steady-state distribution on a plate would yield quite an unsteady distribution when extended to space in this manner.

Having dispensed with the naive approach, we must turn to the chain rule for the correct computation. If $u(r, \theta)$ is a function in the plane, extend it to a function in space by supposing it to be independent of z. The Laplacian of the extended distribution will also be independent of z

and, when restricted to the plane, will give the correct planar Laplacian for the original function u. Applying the Laplacian to u amounts to computing the spherical partial derivatives of u in terms of the cylindrical derivatives. For instance, since $r = \rho \sin\phi$,

$$\frac{\partial u}{\partial \phi} = \frac{\partial u}{\partial r}\rho \cos\phi + \frac{\partial u}{\partial z}\frac{\partial z}{\partial \phi}$$

but $\partial u/\partial z$ is zero. Differentiating again shows that

$$\frac{\partial^2 u}{\partial \phi^2} = \frac{\partial^2 u}{\partial r^2}\rho^2\cos^2\phi - \frac{\partial u}{\partial r}\rho \sin\phi.$$

But when this is restricted to the plane where $r = \rho$ and $\phi = \pi/2$, we find that

$$\frac{\partial^2 u}{\partial \phi^2} = -\frac{\partial u}{\partial r}r.$$

This term corrects the coefficient of $\partial u/\partial r$ which was wrong in the first approach. The other terms are easy to check; for instance, $\partial^2 u/\partial \rho^2$ really does translate to $\partial^2 u/\partial r^2$ when $\phi = \pi/2$.

Another explanation of the error in the naive approach is possible for students who understand that the Laplacian of a function is the divergence of its gradient vector field. If a planar function whose Laplacian is zero is extended to be independent of z, then the gradient of the resulting function is also independent of z and its divergence is zero. On the other hand, the same function extended to be independent of ϕ will typically pick up nonzero divergence in its gradient. A one-dimensional analogue illustrates this phenomenon in a setting where the calculations are simple: the function $f(x) = x$ has the unit vector $\vec{i}$ as its gradient and the divergence of this is zero. However, if f is extended to the plane so that θ is constant, the resulting function $v(r, \theta)$ has the unit vector field

$$\frac{x}{r}\vec{i} + \frac{y}{r}\vec{j}$$

as its gradient. The divergence of this field is $1/r$. By extending the function in a strange way, extra divergence may be introduced in its gradient vector field.

At a more sophisticated level this phenomenon could be used to contrast the push-forward of the Laplacian under two different projections from space to the plane. However, typical students working through this dilemma may learn to appreciate the chain rule and the need for care in changing coordinates.

Reference

[1] R. V. Churchill and J. W. Brown, Fourier Series and Boundary Value Problems, 4th ed., McGraw-Hill, New York, 1987.

Does the Generalized Inverse of A Commute with A?

EDWARD T. WONG
Oberlin College
Oberlin, OH 44074

Every beginning linear algebra student is introduced to the concepts of an invertible matrix A, and its inverse A^{-1}. It is typically shown very early that A and A^{-1} commute, and somewhat later that A^{-1} is expressible as a polynomial in A. Recently the generalized inverse (Moore-Penrose inverse) A^+ of a matrix A has become a very important topic in theory as well as in applications of linear algebra [1], [2]. It is natural to ask if A^+ must commute with A. As we shall show, this is not generally the case. But we can still ask the question: when do A and A^+ commute? Can A^+ be expressed as a polynomial in A if they do commute? In this note, we shall answer these two questions. We pause to provide some background information about A^+ before proceeding to the questions raised above.

The generalized inverse or Moore-Penrose inverse of a matrix can be characterized algebraically or geometrically. The geometric view point involves a linear transformation

$$A: V \to W$$

where V and W are finite dimensional inner product spaces. We use the notations $(|)$ for the inner products and $\|\ \|$ for the norms of vectors in both V and W. Denote the nullspace and range of A by N_A and R_A, respectively. Also let $N_A^{\perp}$ and $R_A^{\perp}$ be their corresponding orthogonal complements. Then $V = N_A \oplus N_A^{\perp}$ and $W = R_A \oplus R_A^{\perp}$. For each $\beta \in W$, $\beta = \beta_1 + \beta_2$ uniquely, where $\beta_1 \in R_A$ and $\beta_2 \in R_A^{\perp}$. In other words, β_1 is the orthogonal projection of β on R_A. Since A maps $N_A^{\perp}$ onto R_A injectively, there exists a unique α_β in $N_A^{\perp}$ where $A(\alpha_\beta) = \beta_1$. The mapping A^+ from W into V defined by $A^+(\beta) = \alpha_\beta$ is the generalized inverse of A [1], [4]. Equivalently, $A^+ = \tilde{A}Q$ where Q is the orthogonal projection of W onto R_A, and $\tilde{A}$ is the inverse of the restriction of A to $N_A^{\perp}$. For each $\beta \in W$, one can easily verify that $\|A(A^+\beta) - \beta\| \leqslant \|A\alpha - \beta\|$ for all $\alpha \in V$. Furthermore, if $\alpha \in V$, $\|A\alpha - \beta\| = \|A(A^+\beta) - \beta\|$, then $\|A^+(\beta)\| \leqslant \|\alpha\|$ [1], [2], [4].

Algebraically, it can be shown that for any real or complex $m \times n$ matrix A, there exists a unique $n \times m$ matrix A^+ such that (1) $AA^+A = A$, $A^+AA^+ = A^+$, and (2) A^+A and AA^+ are self-adjoint (symmetric if A is real, conjugate symmetric if A is complex). Indeed, the unique A^+ is none other than the generalized inverse described geometrically above [1], [2], [3], [5]. Thus some authors take conditions (1) and (2) as a definition for the generalized inverse. We shall use both aspects of A^+ in the following discussion of conditions under which A and A^+ commute.

First, let us use the algebraic characterization. If A is not a square matrix, A^+A and AA^+ have different dimensions. Therefore, we may as well restrict our consideration to square matrices. Then if A and A^+ do commute, we may apply condition (1) immediately to obtain $A = A^2A^+$. Consequently, $A = A^k(A^+)^{k-1}$ for any k. Hence $A = 0$ if A is nilpotent. Thus for a nonzero nilpotent A, A and A^+ cannot commute. Very well, when do they commute? Our first theorem gives the answers.

THEOREM 1. *Let V be a finite dimensional inner product space and A be a linear transformation of V into V. Then the following are equivalent*:

(1) $AA^+ = A^+A$.

(2) $N_A^{\perp} = R_A$.

(3) $N_A = N_{A^*}$, *where A^* is the adjoint of A.*

(4) $A^* = PA$, *where P is an invertible linear transformation of V.*

In other words, A and A^* have the same row-reduced echelon form when they are considered as matrices.

Proof. We shall prove the equivalences of (1) and (2), (2) and (3), and (3) and (4).

Suppose (1) holds; then for any $\alpha \in N_A^\perp$, $\alpha = A^+(A\alpha) = A(A^+\alpha) \in R_A$. Also if $\alpha \in R_A$, then $\alpha = A(A^+\alpha) = A^+(A\alpha) \in N_A^\perp$. Hence $N_A^\perp = R_A$. Suppose (2) holds; then $R_A^\perp = N_A$. For any $\alpha \in V$, $\alpha = \alpha_1 + \alpha_2$ where $\alpha_1 \in N_A^\perp$ and $\alpha_2 \in N_A = R_A^\perp$. Hence $(A^+A)\alpha = A^+(A\alpha_1) = \alpha_1$ and $(AA^+)\alpha = A(A^+\alpha) = A(A^+\alpha_1) = \alpha_1$. Thus $AA^+ = A^+A$.

From the relation $(A\alpha|\beta) = (\alpha|A^*\beta)$ for all α, β in V, $N_A = R_{A^*}^\perp$ or equivalently $N_{A^*} = R_A^\perp$. If (2) holds then $N_A = R_A^\perp = N_{A^*}$. Conversely if $N_A = N_{A^*}$ then $N_A = R_A^\perp$ or $N_A^\perp = R_A$.

Suppose $N_A = N_{A^*}$. Let $\{\alpha_1, \ldots, \alpha_t, \alpha_{t+1}, \ldots, \alpha_n\}$ be a basis of V where $\{\alpha_1, \ldots, \alpha_t\}$ is a basis of $N_A = N_{A^*}$. The sets $\{A\alpha_{t+1}, \ldots, A\alpha_n\}$ and $\{A^*\alpha_{t+1}, \ldots, A^*\alpha_n\}$ are linearly independent. Let P be an invertible linear transformation of V such that $P(A^*\alpha_j) = A\alpha_j$, $j = t+1, \ldots, n$. Then $PA^*(\alpha_i) = A(\alpha_i)$ for all $i = 1, \ldots, n$. Hence, $PA^* = A$. The other direction is obvious.

Recall that a linear transformation is normal if it commutes with its adjoint. If A is normal then $(A\alpha|A\alpha) = (A^*\alpha|A^*\alpha)$ for all $\alpha \in V$. This implies $N_A = N_{A^*}$. On the other hand, the real matrix

$$A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is neither normal nor invertible, but $N_A = N_{A'}$. Thus, A commutes with A^+ for a larger set of matrices than normal matrices. Our next theorem shows that the matrices which commute with their generalized inverses are precisely those for which the generalized inverse is expressible as a polynomial in the original matrix.

THEOREM 2. *Let V be a finite dimensional inner product space and A be a linear transformation of V into V. Then the following are equivalent*:

(1) $AA^+ = A^+A$

(2) *A^+ can be expressed as a polynomial in A.*

Proof. Of course (2) implies (1). Suppose that A and A^+ commute. By Theorem 1, the subspace $N_A^\perp$ and R_A are identical—call this subspace W. The restriction of A to W is an isomorphism on W, and the restriction of A^+ to W is the inverse isomorphism. So relative to a basis for V obtained by extending a basis for W, the matrices of A and A^+ have the form

$$A = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \qquad A^+ = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix},$$

where D is an $r \times r$ invertible matrix.

Let $m(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1x + a_0$ be the minimum polynomial of D. Then $a_0 I_r = -(D^m + a_{m-1}D^{m-1} + \cdots + a_1 D) = -D(D^{m-1} + a_{m-1}D^{m-2} + \cdots + a_1 I_r)$. Since D is invertible and $m(D) = 0$, $a_0 \neq 0$. Hence

$$\begin{aligned} D^{-1} &= -\frac{1}{a_0}\left(D^{m-1} + a_{m-1}D^{m-2} + \cdots + a_2 D + a_1 I_r\right) \\ &= -\frac{1}{a_0^2}\left(a_0 D^{m-1} + a_{m-1}a_0 D^{m-2} + \cdots + a_2 a_0 D + a_1 a_0 I_r\right) \\ &= \frac{1}{a_0^2}\left[a_1 D^m + (a_1 a_{m-1} - a_0) D^{m-1} + \cdots + (a_1^2 - a_2 a_0) D\right] \\ &= p(D), \quad \text{a polynomial in } D \text{ with zero constant term.} \end{aligned}$$

Thus

$$A^+ = \left[\begin{array}{c|c} D^{-1} & 0 \\ \hline 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} P(D) & 0 \\ \hline 0 & 0 \end{array}\right] = P\left(\left[\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right]\right) = P(A).$$

Theorem 2 is proved.

References

[1] S. L. Campbell and C. D. Meyer Jr., Generalized Inverses of Linear Transformations, Pitman, London and San Francisco, 1979.

[2] R. E. Cline, Elements of the Theory of Generalized Inverses for Matrices, EDC/UMAP, Newton, Mass., 1979.

[3] R. A. Penrose, A generalized inverse for matrices, Proc. Cambridge Phil. Soc., 51 (1955) 406–413.

[4] Edward T. Wong, Generalized inverses as linear transformations. Math. Gaz., 63 (1979) 176–181.

[5] ______, Involutory functions and Moore-Penrose inverses of matrices in an arbitrary field, Linear Algebra and Its Applications, 48 (1982) 283–291.

William LeRoy Hart 1892-1984

G. Baley Price
The University of Kansas
Lawrence, KS 66045

William L. Hart, born in Chicago about the same time that Eliakim Hastings Moore was appointed to head the Department of Mathematics in the newly established University of Chicago, later received three degrees from the university, including a Ph.D. for a thesis supervised by Moore. A brilliant record launched him on a long career of research, teaching, and the writing of many elementary textbooks. But let us begin at the beginning.

W. L. Hart was born on October 12, 1892. He attended the South Chicago High School, which had a superior curriculum, principal, and staff of teachers, including Mabel Sykes, a distinguished mathematics teacher to whom Hart gave credit for his success in the field. Hart's good record brought accelerated promotions which enabled him to graduate in 1909—a year earlier than normal. President of his senior class and one of two valedictorians, Hart's record secured for him a scholarship from the University of Chicago, which he entered in 1909 with the firm determination to major in mathematics. Limited family resources made university attendance difficult (his father died after his sophomore year), but Hart held honor scholarships for four years as an undergraduate. Just before Christmas of his senior year, Forest Ray Moulton, a member of the National Academy of Sciences and head of the Department of Mathematical Astronomy, offered Hart a position as the Department's computer (human at that time). This position required 120 hours of work per month, but it paid $50 per month and was permanent, and it guaranteed that Hart would be able to attend graduate school. He was elected to Phi Beta Kappa in his junior year, and he received his bachelor's degree in 1913 with a major in mathematics and minors in physics and astronomy.

In 1914, at the end of his first year of graduate study, Hart was elected to Sigma Xi and received an M.S. degree in mathematics; the degree required both a thesis and an oral examination. Hart's thesis was directed by Gilbert Ames Bliss, a noted expositor and a mathematician who was a member of the National Academy of Sciences. As training for later mathematical exposition, Bliss made Hart rewrite his thesis again and again to improve the exposition.

E. H. Moore was the great mathematical genius at Chicago, and Hart met him for the first time in the summer of 1913. Hart decided to work with Moore for his Ph.D. degree and thesis; he continued with Moore for the three years he was a graduate student. In his second year he was one of only two students in Moore's advanced seminar, and in his last year Hart was the only

References

[1] S. L. Campbell and C. D. Meyer Jr., Generalized Inverses of Linear Transformations, Pitman, London and San Francisco, 1979.

[2] R. E. Cline, Elements of the Theory of Generalized Inverses for Matrices, EDC/UMAP, Newton, Mass., 1979.

[3] R. A. Penrose, A generalized inverse for matrices, Proc. Cambridge Phil. Soc., 51 (1955) 406–413.

[4] Edward T. Wong, Generalized inverses as linear transformations. Math. Gaz., 63 (1979) 176–181.

[5] ______, Involutory functions and Moore-Penrose inverses of matrices in an arbitrary field, Linear Algebra and Its Applications, 48 (1982) 283–291.

William LeRoy Hart 1892-1984

G. Baley Price

The University of Kansas

Lawrence, KS 66045

William L. Hart, born in Chicago about the same time that Eliakim Hastings Moore was appointed to head the Department of Mathematics in the newly established University of Chicago, later received three degrees from the university, including a Ph.D. for a thesis supervised by Moore. A brilliant record launched him on a long career of research, teaching, and the writing of many elementary textbooks. But let us begin at the beginning.

W. L. Hart was born on October 12, 1892. He attended the South Chicago High School, which had a superior curriculum, principal, and staff of teachers, including Mabel Sykes, a distinguished mathematics teacher to whom Hart gave credit for his success in the field. Hart's good record brought accelerated promotions which enabled him to graduate in 1909—a year earlier than normal. President of his senior class and one of two valedictorians, Hart's record secured for him a scholarship from the University of Chicago, which he entered in 1909 with the firm determination to major in mathematics. Limited family resources made university attendance difficult (his father died after his sophomore year), but Hart held honor scholarships for four years as an undergraduate. Just before Christmas of his senior year, Forest Ray Moulton, a member of the National Academy of Sciences and head of the Department of Mathematical Astronomy, offered Hart a position as the Department's computer (human at that time). This position required 120 hours of work per month, but it paid $50 per month and was permanent, and it guaranteed that Hart would be able to attend graduate school. He was elected to Phi Beta Kappa in his junior year, and he received his bachelor's degree in 1913 with a major in mathematics and minors in physics and astronomy.

In 1914, at the end of his first year of graduate study, Hart was elected to Sigma Xi and received an M.S. degree in mathematics; the degree required both a thesis and an oral examination. Hart's thesis was directed by Gilbert Ames Bliss, a noted expositor and a mathematician who was a member of the National Academy of Sciences. As training for later mathematical exposition, Bliss made Hart rewrite his thesis again and again to improve the exposition.

E. H. Moore was the great mathematical genius at Chicago, and Hart met him for the first time in the summer of 1913. Hart decided to work with Moore for his Ph.D. degree and thesis; he continued with Moore for the three years he was a graduate student. In his second year he was one of only two students in Moore's advanced seminar, and in his last year Hart was the only

W. L. Hart

student. Hart found him to be a poor teacher but an inspiring mathematician. Although Hart, in 1915, asked Moore and Moulton to be joint directors for his Ph.D., he had already selected "Differential Equations and Implicit Functions in Infinitely Many Variables" [**26**], [**28**] as his topic and he had almost finished the thesis without any help from them. This choice of topic is not surprising since both Moore and Moulton had written on the subject [**29**, p. 408], [**1**, pp. 355-375]. In 1916 Hart received his Ph.D. degree *summa cum laude* with a major in mathematics and a minor in mathematical and practical astronomy.

Hart's record at Chicago brought him an appointment as Benjamin Peirce instructor at Harvard for 1916–1917. There he met Dunham Jackson [**39**], later a colleague at Minnesota, and he published two research papers [**28**], [**29**]. Marston Morse was a student in a course which Hart taught at Harvard. In January, 1917, a letter from E. H. Moore offered Hart an instructorship at Chicago starting in September, 1917. He quickly accepted the offer, but ten days later he received a second letter from Moore stating that, because of the probability that the United States would be drawn into World War I soon, the University of Chicago was cancelling all of its new appointments for 1917–1918. Hart thought that the offer would be renewed after the war, but it was not. Hart described the loss of the position as the greatest disappointment of his life as a mathematician.

The United States was drawn into World War I, and Hart applied for officer's training camp and went to Fort Sheridan in April, 1917. In August, 1917, he was commissioned a 2d Lieutenant in the regular Army and finally he was sent to the Coast Artillery Training Center at Fortress Monroe, Virginia. Before he could be sent overseas, he was assigned to a mathematical bureau, the Ballistic Unit of the Ordnance Department, in Washington, D.C. By that time a 1st Lieutenant, he requested overseas service and was assigned to the 49th Coast Artillery, a regiment firing eight-inch guns, which was training in Virginia. In that regiment, he rose to the rank of Captain and, in the summer of 1918 when he was 25 years old, to the rank of Major. The regiment went overseas late in the summer of 1918, but it did not see combat service. Hart was kept in France on various detached assignments, including four months as a mathematics teacher at a university for Army personnel at Beaune, France. He returned to the U.S. in June, 1919, and again was assigned to duty at the Ballistic Unit of the Ordnance Department in Washington. In September, 1919, he resigned his commission and accepted an Assistant Professorship in mathematics at the University of Minnesota.

Hart made rapid progress at Minnesota. In 1920, at the end of only one year, he was promoted to Associate Professor, and in 1924 he was promoted to Professor. He was Chairman of the Department for five years from 1924 to 1928, and also for four years from 1935 to 1939. At first Hart continued his research program at Minnesota as is shown by the list of his published papers [**30**]–[**36**], but he began to respond to the other influence which he had encountered at the University of Chicago. Chicago was a world-renowned center of research in mathematics, but it was also distinguished for its interest in mathematics education. E. H. Moore, Hart's idol, was famous as a research mathematician, but his retiring address as President of the American Mathematical Society [**23**], [**24**, p. 16] dealt with mathematics education, and there were many other evidences of his interest in the subject. Furthermore, Herbert Ellsworth Slaught, who was one of the leaders in the founding of the Mathematical Association of America in 1915 (Hart was then in his last year of graduate study at Chicago), was a member of the faculty of the University of Chicago [**24**, pp. 16–23]. Finally, at Minnesota, administrative duties, teaching responsibilities, and—to some extent—his own interests crowded out his research, and he published no research between 1930 and 1956 [**36**], [**37**]. The other Chicago influence became dominant, and he became identified as a curriculum expert and teaching authority in the college field.

As a curriculum expert, Hart gave his support to at least two new subjects. The first of these was the mathematics of investment. In the discussion which followed a talk entitled "Mathematics for Commerce Students" by E. B. Skinner of the University of Wisconsin, Hart said that Professor Skinner developed for the first time in the United States a course in the mathematics of investment, and he described Skinner's success in the selection of material for the detailed work of the course [**3**]. Skinner published a textbook for the course he designed; the opening paragraph of P. R. Rider's review [**2**] of the revised edition of it reads as follows:

> The first edition of Professor Skinner's book, which was published in 1913, was a pioneer American textbook in the field of mathematics of investment. It was a good book and stood alone for some years. Recently, however, a number of other textbooks of a similar nature have appeared....

Hart was one of those who wrote a similar textbook; published in 1924, it was entitled *The Mathematics of Investment*. The subject appears to have held a special fascination for him; the fifth and last edition was published in 1975 [**40**]. He taught the course (to large sections) frequently at Minnesota, and he wrote an article, "The Mathematics of Investment," for the Sixth Yearbook [**4**] of the National Council of Teachers of Mathematics. The volumes of the Monthly record Hart's continuing promotion of the subject through the notes and reviews he wrote and talks he gave. The present overshadowing of the mathematics of investment in Schools of Business by linear programming, operations research, statistics, and computer applications serves to emphasize the changes which have occurred in mathematics and its applications since Hart wrote his first book in 1924.

Statistics was another new subject in the curriculum which Hart supported. He wrote a paper on a problem in statistics [**34**] and another on an application [**5**], and he gave a talk which dealt with the teaching of statistics [**9**]. Professor Hart taught mathematical statistics, and he served for five years as chairman of the graduate committee on statistics at Minnesota. It is significant that a number of his students became statisticians and especially applied statisticians in the field of public health.

Professor Hart contributed to teacher education. During nearly all of the time he was at Minnesota, he acted as liaison member of the Department of Mathematics with the College of Education for the training of teachers of secondary mathematics. At first this activity was informal, but it became increasingly formal as the demand for teachers expanded. During his last fifteen years at Minnesota, he served as the representative of the Department of Mathematics on the faculty of the College of Education and was the principal adviser for students in that College majoring in mathematics. A number of his talks and published papers attest to his deep interest in the problems of secondary and college education [**6**]–[**11**].

Because of his Army experience in World War I and his reputation as a curriculum expert and teaching authority, Hart became a leading adviser on educational and training matters concerning mathematics in World War II. In 1940 he was appointed chairman of the Subcommittee on Education for Service of the War Preparedness Committee of the AMS and the MAA [**12**]. In January, 1942, the War Department appointed a committee consisting of W. L. Hart, W. M. Whyburn, and C. C. Wylie "to make a survey of the ground school courses offered in pilot and non-pilot courses in the Air Corps Flying Training System, with a view to outlining preparatory courses to be given in colleges and universities"; the committee's recommendations were published in the Monthly [**18**], [**20**]. Professor Hart was the academic director during the organization of the curriculum and early months of activity of the Pre-meteorological Training Program of the Air Force [**21**]. He was a member of the committee appointed by the mathematical societies to prepare accreditation examinations in college mathematics for men of the armed forces who offered, for college credit, work done under the program of the Armed Forces Institute (a correspondence school in Madison, Wisconsin) [**16**]. There were other committee assignments and reports [**22**]. Hart's reports show that he gave strong support to various war activities which involved mathematics [**12**]–[**22**].

Professor Hart will be best remembered by many as a writer of textbooks for mathematics courses at the freshman and sophomore level [**40**]–[**70**]. His *Mathematics of Investment* in 1924 has been mentioned already; in all, there were five editions, the last appearing in 1975. He published six editions of his *College Algebra*, the first in 1926 and the last in 1978. After the publication of his first textbook, there followed a steady stream of new titles and new editions of old titles. Two of William L. Hart's books [**48**], [**49**] were written with his brother Walter W. Hart, whom some have considered the leading author of high school mathematics textbooks. (In 1936 W. W. Hart resigned from his position at the University of Wisconsin so that he could devote his full time "to writing texts for and pertaining to secondary and elementary mathematics" [Amer. Math. Monthly, 43 (1936) 198].) Reviews of many of W. L. Hart's textbooks can be found in the volumes of the American Mathematical Monthly. The following report describes the sale of W. L. Hart's textbooks by D. C. Heath:

> Up to June 1, 1964, Heath had sold 1,590,895 of his texts in hard cover books. Also, during World War II, the United States Armed Forces Institute bought 230,200 paper-bound copies of his *College Algebra*, *Revised Edition*, and his *Mathematics of Investment*, *2nd Edition*. His *College Algebra*, *4th Edition*, has sold 202,858 copies up to June 1, 1964,. . . .

W. L. Hart was a loyal and active member of the MAA. He was elected to membership in his first year at Minnesota [Amer. Math. Monthly, 27 (1920) 108]. He was twice elected a member of the MAA Board of Trustees [Amer. Math. Monthly, 37 (1930) 114 and 40 (1933) 136]. He was a Vice President of the MAA in 1939 [Amer. Math. Monthly, 46 (1939) 130]. Before World War II the Editor of the Monthly thanked Hart, in volume after volume, for refereeing papers and other assistance. Reports in the Monthly show that he attended, and participated in, many national meetings and also meetings of the Minnesota Section and other Sections. For at least six years during the 1970's Hart performed valuable service for the Association as a member of its Committee on Special Funds.

Professor Hart was considered an outstanding teacher by his students. He was dynamic and lucid as a teacher; he was tremendously energetic, with an enthusiasm for mathematics that was catching—some called him "a supersalesman." Hart was a pioneer in teaching mathematics to large sections (100–120 students) by the lecture-problem-solving method with the help of one assistant (in the lecture). He developed new types of questions to facilitate the ease and accuracy of grading tests and examinations. He was demanding and upheld high standards, but his students worked hard and produced excellent results.

One former colleague wrote, "Bill Hart was one of the most helpful and likeable persons I have ever met in the academic field." He helped his students find jobs so they could stay in school. He played tennis with them—and held his own with persons 20 to 30 years younger than himself; later he played much golf. He kept generous office hours to answer his students' questions; he invited colleagues to Sunday dinner. One of his former students wrote: "Professor Hart never ceased being interested in his former students and became a good friend to most of us"; and also, "I did not hear from him this Christmas for the first time in 30 years."

Professor Hart retired from the University of Minnesota in 1962 and moved to California. He retained his youthful appearance and enthusiasm to the end of his life. He died at his home in La Jolla on December 16, 1984.

On January 25, 1985 Professor Byron William Brown, Jr., Head of the Division of Biostatistics in Stanford University, wrote the following letter to Mrs. Hart; he received degrees in mathematics, statistics, and biostatistics at Minnesota.

> Word reached me only now of Professor Hart's death. I write to tell you that I am very sorry for that, and to tell you that there must be dozens, perhaps hundreds across the country, like me, who feel a personal loss in his death. He was my adviser at Minnesota, and only for a brief time, but his enthusiasm, his warm personal interest, his sound advice and continued interest thru the years have been a great and good inspiration and influence in my career. You can be sure that he will continue to live in the hearts of all of us.

I, G. B. Price, obtained the information for this article on the life and work of Professor William LeRoy Hart from the following sources: (a) my personal acquaintance with him over a period of more than 45 years; (b) autobiographical and other information supplied by Professor Hart's family; (c) Hart's biographies in *Who's Who in America*, *American Men of Science*, and *Leaders in American Education*; (d) the published record of his activities, especially as it appears in the volumes of the *American Mathematical Monthly*; (e) a conversation (my last one with him) on the occasion of a visit in May, 1984; and (f) letters from a number of Professor Hart's former students and colleagues.

Notes and References

[1] Forest Ray Moulton, Differential Equations, New York, The Macmillan Company, 1930, xv + 395 pages.

[2] American Mathematical Monthly, 32 (1925) 476–477, review by P. R. Rider of The Mathematical Theory of Investment, by E. B. Skinner, revised edition, New York, Ginn and Company, 1924, xi + 269 pages.

[3] Mathematics for Commerce Students, by E. B. Skinner, discussion by W. L. Hart, American Mathematical Monthly, 34 (1927) 450–451.

[4] The Mathematics of Investment, by W. L. Hart. An article in NCTM's Sixth Yearbook (1931) entitled Mathematics in Modern Life.

[5] The probability curve, by W. L. Hart, Journal of the American Water Works Association, November, 1920.

[6] Student placement in secondary mathematics, by W. L. Hart, American Mathematical Monthly, 41 (1934) 611–619.

[7] The need for a reorientation of secondary mathematics from the college viewpoint, by W. L. Hart, The Mathematics Teacher, February, 1935. See also American Mathematical Monthly, 42 (1935) 130.

[8] Report on the training of teachers of mathematics, by W. L. Hart, committee chairman, American Mathematical Monthly, 42 (1935) 263–277.

[9] Mathematics preparation of students in pre-business statistics, by W. L. Hart, American Mathematical Monthly, 43 (1936) 66, 68.

[10] The nation calls for mathematics, by W. L. Hart, School, Science, and Mathematics, February, 1943.

[11] Short term and long term effects of the war on the secondary curriculum in mathematics, by W. L. Hart, School, Science, and Mathematics, June, 1943.

[12] Report of the war preparedness committee, American Mathematical Monthly, 47 (1940) 500–502.

[13] Specialized training for flying cadets, by W. L. Hart, American Mathematical Monthly, 48 (1941) 221–222.

[14] Mathematics in the defense program, by Marston Morse and William L. Hart, American Mathematical Monthly, 48 (1941) 293–302.

[15] On education for service, by William L. Hart, American Mathematical Monthly, 48 (1941) 353–362.

[16] Certification of mathematical training for military service, by W. L. Hart, American Mathematical Monthly, 48 (1941) 492.

[17] Distribution of the report on mathematical education for defense, by W. L. Hart, American Mathematical Monthly, 48 (1941) 570.

[18] Pre-training of aviation cadets, by W. L. Hart, American Mathematical Monthly, 49 (1942) 274–276.

[19] Pre-induction training, by W. L. Hart, American Mathematical Monthly, 49 (1942) 351–352.

[20] Report to the chief of the Air Force concerning pilot and navigator training, by W. L. Hart, 1942. (Not for publication)

[21] The Pre-meteorological training program of the Air Force, by W. L. Hart, 1943.

[22] Universal military service in peace time, a report prepared by W. L. Hart as chairman of a subcommittee of the War Policy Committee, American Mathematical Monthly, 53 (1946) 49–58.

[23] On the foundations of mathematics, by E. H. Moore, Bulletin of the American Mathematical Society, 9 (1903) 402–424. See also [**24**, p. 16].

[24] The Mathematical Association of America: Its First Fifty Years. Kenneth O. May, Editor. Washington, The Mathematical Association of America, 1972. vii + 172 pages.

William L. Hart's Research Papers

[25] (Abstract) Determination of the coefficients of a trigonometric Series, Bulletin of the American Mathematical Society, 20 (1914) 455–456.

[26] (Abstract of thesis) Differential equations and implicit functions in infinitely many variables, Proceedings of the National Academy of Sciences, June, 1916.

[27] On trigonometric series, Annals of Mathematics, 18 (1916–1917) 99–104. (Abstract in **25** above)

[28] Differential equations and implicit functions in infinitely many variables, Transactions of the American Mathematical Society, 18 (1917) 125–160. (Abstract in **26** above)

[29] Linear differential equations in infinitely many variables, American Journal of Mathematics, 39 (1917) 407–424.

[30] The Cauchy-Lipschitz method for infinite systems of differential equations, American Journal of Mathematics, 43 (1921) 226–231.

[31] The pseudo-derivative of a summable function, Bulletin of the American Mathematical Society, 27 (1920–1921) 202–211.

[32] Functions of infinitely many variables in Hilbert space, Transactions of the American Mathematical Society, 23 (1922) 30–50.

[33] Summable infinite determinants, Bulletin of the American Mathematical Society, 28 (1922) 171–178.

[34] The method of monthly means for determination of a seasonal variation, quarterly publication of the American Statistical Association, September, 1922. An abstract is in the Bulletin of the American Mathematical Society, 28 (1922) 301.

[35] Functionals of summable functions, Annals of Mathematics, 24 (1923) 23–38.

[36] Limited tri-linear forms in Hilbert space, American Journal of Mathematics, 52 (1930) 563–570.

[37] (With Theodore S. Motzkin) A composite Newton-Raphson gradient method for the solution of systems of equations, Pacific Journal of Mathematics, 6 (1956) 691–707.

[38] (With Theodore S. Motzkin) Proof of the fundamental theorem on implicit functions by use of composite gradient corrections, Pacific Journal of Mathematics, 8 (1958) 429–436.

[39] Dunham Jackson, a biographical memoir, Reprinted from volume XXXIII of Biographical Memoirs of the National Academy of Sciences of the United States, 1959. See also Dunham Jackson: 1888–1946, Bulletin of the American Mathematical Society, 54 (1948) 847–860.

William L. Hart's Textbooks

Published by D. C. Heath and Company

[40] Mathematics of Investment, editions in 1924, 1930, 1946, 1957, and 1975.
[41] College Algebra, editions in 1926, 1931, 1937, 1948, 1953, and (with Bert K. Waits) 1978.
[42] Brief College Algebra, editions in 1932 and 1947.
[43] Plane Trigonometry, 1933.
[44] Plane and Spherical Trigonometry, 1934.
[45] Introduction to Mathematics of Business, 1936.
[46] Introduction to College Algebra, editions in 1938 and 1947.
[47] Plane Trigonometry with Applications, 1942.
[48] (With Walter W. Hart) Solid Geometry and Spherical Trigonometry with Applications, 1942.
[49] (With Walter W. Hart) Plane Trigonometry, Solid Geometry, and Spherical Trigonometry with Applications, 1942.
[50] Plane and Spherical Trigonometry with Applications, 1943.
[51] (With W. A. Wilson and J. I. Tracey) First Year College Mathematics, 1943.
[52] First Year College Mathematics, brief edition, 1943.
[53] Essentials of College Algebra and Mathematics of Investment, 1946.
[54] Intermediate College Algebra, 1947.
[55] Elements of Analytic Geometry, 1950.
[56] College Trigonometry, 1951.
[57] Trigonometry, 1954.
[58] Calculus, 1955.
[59] Analytic Geometry and Calculus, editions in 1957 and 1963.
[60] College Algebra and Trigonometry, editions in 1959 and (with Bert K. Waits) 1978.
[61] Modern Plane Trigonometry, 1961.
[62] Algebra, Elementary Functions, and Probability, 1965.
[63] Basic College Algebra, 1972.

Published by Prindle, Weber & Schmidt

[64] Intermediate Mathematics for Business, 1967.
[65] Intermediate Algebra, 1968.
[66] Mathematics for Managerial and Social Sciences, 1970.
[67] Algebra, A College Course, 1975.

Published by Page-Ficklin

[68] Calculus with Applications, 1973.
[69] Preparation for Calculus, A Brief Course, 1973.

Published by Goodyear

[70] Algebra and the Elementary Functions, 1974.

PROBLEMS

LOREN C. LARSON, Editor
BRUCE HANSON, Associate Editor
St. Olaf College

LEROY F. MEYERS, Past Editor
The Ohio State University

Proposals

To be considered for publication, solutions should be received by March 1, 1987.

1247. *Proposed by Constantin Gonciulea, Traian College, Drobeta Turnu-Severin, Romania.*

Let a and b be positive integers. Suppose that A and B are finite and disjoint sets of integers such that if $i \in A \cup B$ then $i+a \in A$ or $i-b \in B$. Show that $a|A| = b|B|$. (If X is a finite set, $|X|$ denotes the number of elements in X.)

1248. *Proposed by Edward Kitchen, Santa Monica, California.*

From $1/7 = .1428571\ldots$, we get the points (1,4), (4,2), (2,8), (8,5), (5,7), and (7,1). Show that they lie on an ellipse.

1249. *Proposed by Víctor Hernandez, Universidad Autónoma de Madrid, Spain.*

a. Find all continuous functions f that satisfy:

$$f(f(x)) = f(x), \quad \text{for all real } x. \tag{1}$$

b. Find all differentiable functions that satisfy (1).

1250. *Proposed by J. Metzger, University of North Dakota, and S. Kaler, Honeywell Corporation.*

The Pellian sequence $(x_n)_{n=1}^{\infty}$ is defined as follows: x_n is the smallest positive integer x for which nx^2+1 is a perfect square; if no such x exists, x_n is defined to be 0. The sequence begins $0, 2, 1, 0, 4, 2, 3, \ldots$.

a. Show that every positive integer occurs infinitely often in the Pellian sequence.

b. Determine all occurrences of p^k, p a prime, $k > 0$, in the Pellian sequence.

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

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

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

1251. *Proposed by Harry Ruderman, Lehman College, The Bronx, New York.*

Show that

$$\frac{(3n)!}{n!(n+1)!(n+2)!}$$

is an integer for $n \geqq 3$.

Quickies

Q713. *Submitted by M. S. Klamkin, University of Alberta.*

If one of the arcs joining the midpoints of the sides of a spherical triangle is 90°, show that the other two such arcs are also 90°.

Q714. *Proposed by Norman Schaumberger, Bronx Community College.*

If a circle meets each side of a 99-sided polygon in precisely one point and does not pass through any vertex of the polygon then it must be tangent to at least one side of the polygon.

Solutions

Subgroup Generated by k-Cycles **September 1985**

1221. *Proposed by William A. McWorter, Jr., The Ohio State University.*

Let k and n be fixed integers with $1 < k < n$. Which subgroup of the symmetric group S_n is generated by the k-cycles $(x, x+1, \ldots, x+k-1)$ for $x = 1, 2, \ldots, n-k+1$?

Solution by William P. Wardlaw, U.S. Naval Academy.

We will show that the given k-cycles generate the whole symmetric group S_n when k is even, and the alternating group A_n when k is odd.

We will use the following well-known properties of S_n and A_n:

(i) S_n is generated by the transpositions $(1,2), (2,3), \ldots, (n-1, n)$;

(ii) A_n is generated by the 3-cycles $(1, 2, y)$, $y = 3, 4, \ldots, n$.

Let $G(k)$ denote the subgroup of S_n generated by the k-cycles $(x, x+1, \ldots, x+k-1)$, $x = 1, 2, \ldots, n-k+1$.

By property (i) above, $G(2) = S_n$. Also, $(1,2,4) = (2,3,4)\,(1,2,3)^{-1}(2,3,4)^{-1}$, and for $4 < y \leqslant n$, $(1,2,y) = (y-2, y-1, y)^{-1}(1,2,y-1)(y-2, y-1, y)$. Thus, by induction, $(1,2,y)$ belongs to $G(3)$ for $y = 3, 4, \ldots, n$. Now property (ii) above shows that $G(3) \supseteq A_n$.

Suppose that $3 < k < n$. Let $\alpha = (1,2,\dots,k)$ and $\beta = (2,3,\dots,k+1)$. Then $\alpha\beta^{-1} = (1,k+1,k)$, $(1,3,2) = \beta^{-2}(1,k+1,k)\beta^2$, and $(1,2,3) = (1,3,2)^2$ are in $G(k)$. Suppose that $1 < x \leqslant n-2$. There is a generating k-cycle $\gamma = (y, y+1,\dots, y+k-1)$ which moves each of the marks $x-1, x, x+1, x+2$ (because $k > 3$). Thus, $(x, x+1, x+2) = \gamma^{-1}(x-1, x, x+1)\gamma$. It follows by induction that each of the 3-cycles $(x, x+1, x+2)$, $1 \leqslant x \leqslant n-2$ is in $G(k)$. Thus, $G(3) \subseteq G(k)$, and therefore, from the previous reasoning, $A_n \subseteq G(k)$.

If k is even, the odd permutation $(1,2,\dots,k)$ is in $G(k)$, so $G(k) = S_n$. If k is odd, $G(k)$ is generated by even permutations, and this implies that $G(k) \subseteq A_n$, and, therefore, $G(k) = A_n$. This completes the proof.

Also solved by Duane Broline, Tatiana Deretsky, Karen R. Fawcett, Robert Gilmer and Budh Nashier, Bjorn Poonen (student), Gary L. Walls, and the proposer.

Fawcett notes that the result follows immediately from a theorem of C. Jordan (1873): *If G is a doubly transitive permutation group of degree n and G contains a 3-cycle, then G is either an alternating or a symmetric group.* She also points out that the same subgroup can be generated by the $\left(\left[\frac{n-2}{k-1}\right]+1\right)$ k-cycles $(1,2,\dots,k)$, $(k, k+1,\dots, 2k-1)$, $(2k-1,\dots,3k-2)$, $\dots$, $(n-k+1,\dots,n)$ and this is the minimal number of k-cycles which are required to generate this subgroup. Details are in her master's thesis, *Basic Properties of the Alternating and Symmetric Groups*, University of Southern Mississippi, 1985.

Odd Binomial Coefficients **September 1985**

1222. *Proposed by Marta Sved, The University of Adelaide, Australia.*

Given the nonnegative integer k, find the number of odd binomial coefficients $\binom{m}{r}$ with $0 \leqslant m < 2^k$.

I. *Solution by Roger B. Nelsen, Lewis and Clark College.*

Let A_k denote the number of such coefficients. We claim that $A_k = 3^k$. We first note that the number of odd binomial coefficients $\binom{m}{r}$ with $0 \leqslant r \leqslant m$ is $2^{s(m)}$, where $s(m)$ is the number of 1's in the binary representation of m (Problem E1288, *American Mathematical Monthly*, 65 (1958), pp. 368–9). Thus we have

$$A_{k+1} = A_k + \sum_{m=2^k}^{2^{k+1}-1} 2^{s(m)} = A_k + \sum_{m=0}^{2^k-1} 2^{s(m+2^k)}.$$

But $s(m+2^k) = 1 + s(m)$ when $0 \leqslant m < 2^k$, so

$$A_{k+1} = A_k + 2\sum_{m=0}^{2^k-1} 2^{s(m)} = 3A_k.$$

Induction on k (with $A_0 = 1$) establishes the result.

II. *Solution by Edward T. H. Wang, Wilfrid Laurier University, Canada.*

Since the binary representation of each number in $\{0,1,\dots,2^k-1\}$ is equivalent to a string of 0's and 1's of length k, we see that exactly $\binom{k}{i}$ of these numbers have i 1's in their binary representation, $i = 0,1,\dots,k$. Therefore, using the notation of the previous solution,

$$A_k = \sum_{m=0}^{2^k-1} 2^{s(m)} = \sum_{i=0}^{k} \binom{k}{i} 2^i = 3^k.$$

III. *Solution by Bjorn Poonen, student, Harvard College.*

We claim more generally that the number of binomial coefficients $\binom{m}{r}$ not divisible by a prime p with $0 \leqslant m < p^k$ is $(p(p+1)/2)^k$.

In the polynomial ring $(Z/pZ)[x]$, $(1+x)^p = 1 + x^p$, since $\binom{p}{r} = p!/r!(p-r)!$ is divisible by p for $1 \leqslant r \leqslant p-1$. Hence by induction we have $(1+x)^{p^q} = 1 + x^{p^q}$. Thus, if $m = \sum_{i=0}^{k-1} a_i p^i$, $0 \leqslant a_i \leqslant p-1$, then

$$(1+x)^m = (1+x)^{a_0}(1+x^p)^{a_1} \cdots (1+x^{p^{k-1}})^{a_{k-1}}$$
$$= \sum_{b_0=0}^{a_0} \sum_{b_1=0}^{a_1} \cdots \sum_{b_{k-1}=0}^{a_{k-1}} \binom{a_0}{b_0}\binom{a_1}{b_1} \cdots \binom{a_{k-1}}{b_{k-1}} x^{b_0+b_1p+\cdots+b_{k-1}p^{k-1}}.$$

Each term in this sum has a different power of x and $p \nmid \binom{a_i}{b_i}$. Thus, there are $\prod_{i=0}^{k-1}(a_i+1)$ binomial coefficients $\binom{m}{r}$ not divisible by p for that fixed m.

Let $S(k) = \sum_{m=0}^{p^k-1} \prod_{i=0}^{k-1}(a_i+1)$, be the total number of $\binom{m}{r}$ not divisible by p with $0 \leqslant m < p^k$. Clearly, $S(0) = 1$. If now $0 \leqq m < p^{k+1}$, then we may write

$$m = \sum_{i=0}^{k} a_i p^i = a_k p^k + n \quad \text{with} \quad 0 \leqq a_i \leqq p-1$$

and $0 \leqq n < p^k$, so that

$$S(k+1) = \sum_{m=0}^{p^{k+1}-1} \prod_{i=0}^{k} (a_i+1)$$
$$= \sum_{a_k=0}^{p-1} (a_k+1) \sum_{n=0}^{p^k-1} \prod_{i=0}^{k-1} (a_i+1)$$
$$= \frac{p(p+1)}{2} S(k).$$

By induction, $S(k) = \left(\frac{p(p+1)}{2}\right)^k$.

Also solved by Robert Bernstein, Paul Bracken, Duane Broline, Kenneth A. Brown, Jr., Patrick Costello, Alberto Facchini (Italy), David C. Flaspohler, G. A. Heuer, L. Kuipers (Switzerland), Kee-wai Lau (Hong Kong), Syrous Marivani, Oxford Running Club (University of Mississippi), Robert Patenaude, Charles Peltier, Christopher Peterson, Daniel M. Rosenblum, Harry D. Ruderman, Allen J. Schwenk, Robert E. Shafer, Jan Söderkvist (Sweden), Michael Vowe (Switzerland), Western Maryland College Problems Group, William P. Wardlaw, and the proposer. There were three incomplete solutions.

Schwenk showed that the number of binomial coefficients $\binom{m}{r}$ not divisible by p with $m < n$ is

$$\sum_{i=0}^{k}\left(\prod_{j=i+1}^{k}(1+a_j)\right)\binom{1+a_i}{2}\binom{p+1}{2}^i.$$

An Erdős Problem — September 1985

1223. *Proposed by Paul Erdős, Hungarian Academy of Sciences.*

Let $(a_k)_{k=1}^{\infty}$ be a strictly increasing sequence of positive integers such that $\sum_{k=1}^{\infty}(1/a_k)$ converges.

(a) For each positive integer n, set

$$f(n) = \sum_{\substack{k=1 \\ a_k \neq n}}^{\infty} \frac{1}{|n-a_k|}.$$

Prove that $\liminf_{n\to\infty} f(n) = 0$. Give an example where $\lim_{n\to\infty} f(n) = 0$ fails.

(b) Prove that

$$\liminf_{l\to\infty} \sum_{k=1}^{l-1} \frac{1}{a_l - a_k} = 0 \quad \text{and} \quad \liminf_{l\to\infty} \sum_{k=l+1}^{\infty} \frac{1}{a_k - a_l} = 0.$$

Give an example where

$$\liminf_{l\to\infty} \sum_{\substack{k=1\\k\neq l}}^{\infty} \frac{1}{|a_k - a_l|} = 0$$

fails.

Solution by Bjorn Poonen, student, Harvard College.

(a) If $\liminf_{n\to\infty} f(n)$ were not zero, there would exist $\varepsilon > 0$ and a positive integer M such that for all $n \geqslant M$, $f(n) \geqslant \varepsilon$. We shall show that this cannot be by finding a positive divergent series $\sum_{n=1}^{\infty} g(n)$ such that $\sum_{n=1}^{\infty} f(n)g(n)$ converges.

Let $S_0 = \left\{ n > 1 : \sum_{a_k \geqslant n} \frac{1}{a_k} > 1 \right\}$ and let $S_i = \left\{ n > 1 : 2^{-i} < \sum_{a_k \geqslant n} \frac{1}{a_k} \leqslant 2^{-i+1} \right\}$ for $i \geqslant 1$. Since $\sum_{k=1}^{\infty} \frac{1}{a_k}$ converges, each S_i is finite, and the S_i and $\{1\}$ partition the set of positive integers. Next, let $c_i = \sum_{n\in S_i} \frac{1}{n \log n}$. Finally, let $g(1) = 1$, and for $n \in S_i$, let

$$g(n) = \begin{cases} \dfrac{1}{n \log n} & \text{if } c_i \leqslant 1 \\ \dfrac{1}{c_i n \log n} & \text{if } c_i > 1. \end{cases}$$

We claim that $g(n)$ has the following properties:

(1) $0 < g(n) \leqslant \dfrac{1}{n \log n}$ for $n > 1$,

(2) $\sum_{n=1}^{\infty} g(n)$ diverges,

(3) $\sum_{n=1}^{\infty} g(n) \sum_{a_k \geqslant n} \dfrac{1}{a_k}$ converges.

Now (1) is obvious from the definitions. To prove (2), observe that if only finitely many c_i are greater than 1, then for some M, $g(n) = \frac{1}{n \log n}$ for all $n \geqslant M$. But $\int_M^{\infty} \frac{dx}{x \log x}$ diverges so $\sum_{n=1}^{\infty} g(n)$ diverges by the integral test in this case. On the other hand, whenever $c_i > 1$, $\sum_{n \in S_i} g(n) = 1$, so if there are infinitely many such c_i, $\sum_{n=1}^{\infty} g(n)$ diverges.

Next we have

$$\sum_{n=1}^{\infty} g(n) \sum_{a_k \geqslant n} \frac{1}{a_k} = \sum_{k=1}^{\infty} \frac{1}{a_k} + \sum_{n\in S_0} g(n) \sum_{a_k \geqslant n} \frac{1}{a_k} + \sum_{i=1}^{\infty} \sum_{n \in S_i} g(n) \sum_{a_k \geqslant n} \frac{1}{a_k}.$$

The first two sums converge since $\sum_{k=1}^{\infty} \frac{1}{a_k}$ converges and S_0 is finite. Also,

$$\sum_{i=1}^{\infty} \sum_{n\in S_i} g(n) \sum_{a_k \geqslant n} \frac{1}{a_k} \leqslant \sum_{i=1}^{\infty} \sum_{n\in S_i} g(n) 2^{-i+1} \leqslant \sum_{i=1}^{\infty} 2^{-i+1} = 2,$$

since $\sum_{n\in S_i} g(n) \leqslant 1$, for every i. This proves (3).

Now

$$\sum_{n=1}^{\infty} f(n) g(n) = \sum_{n=1}^{\infty} \sum_{\substack{k=1\\a_k \neq n}} \frac{g(n)}{|n - a_k|}$$

$$= \sum_{k=1}^{\infty} \sum_{n=1}^{\left[\frac{a_k}{2}\right]} \frac{g(n)}{a_k - n} + \sum_{k=1}^{\infty} \sum_{n=[a_k/2]+1}^{a_k - 1} \frac{g(n)}{a_k - n} + \sum_{k=1}^{\infty} \sum_{n=a_k+1}^{\infty} \frac{g(n)}{n - a_k}.$$

(Here and in the following, $[x]$ is the greatest integer $\leqslant x$.)

We shall show that each of these three sums converges. First,

$$\sum_{k=1}^{\infty} \sum_{n=1}^{[a_k/2]} \frac{g(n)}{a_k - n} \leqslant \sum_{k=1}^{\infty} \sum_{n=1}^{[a_k/2]} \frac{g(n)}{a_k/2}$$
$$\leqslant \sum_{k=1}^{\infty} \sum_{n=1}^{a_k} \frac{2g(n)}{a_k}$$
$$= 2 \sum_{n=1}^{\infty} g(n) \sum_{a_k \geqslant n} \frac{1}{a_k},$$

which converges by (3).

In the second sum the terms with $a_k \leqslant 2$ drop out:

$$\sum_{k=1}^{\infty} \sum_{n=[a_k/2]+1}^{a_k - 1} \frac{g(n)}{a_k - n} \leqslant \sum_{\substack{k=1 \\ a_k > 2}}^{\infty} \frac{1}{\frac{a_k}{2} \log\left(\frac{a_k}{2}\right)} \sum_{n=[a_k/2]+1}^{a_k - 1} \frac{1}{a_k - n}$$
$$= 2 \sum_{\substack{k=1 \\ a_k > 2}}^{\infty} \frac{1}{a_k \log(a_k/2)} \sum_{m=1}^{[(a_k - 1)/2]} \frac{1}{m}$$
$$\leqslant 2 \sum_{\substack{k=1 \\ a_k > 2}}^{\infty} \frac{\log[(a_k - 1)/2] + 1}{\log(a_k/2)} \left(\frac{1}{a_k}\right)$$

(since $\sum_{m=1}^{N} \frac{1}{m} \leqslant 1 + \int_1^N \frac{dx}{x}$). This last sum converges since

$$\lim_{k \to \infty} \frac{\log[(a_k - 1)/2] + 1}{\log(a_k/2)} = 1.$$

Finally,

$$\sum_{k=1}^{\infty} \sum_{n=a_k+1}^{\infty} \frac{g(n)}{n - a_k} \leqslant \sum_{k=1}^{\infty} \sum_{n=a_k+1}^{\infty} \frac{1}{(n - a_k) n \log(a_k + 1)}$$
$$= \sum_{k=1}^{\infty} \frac{1}{a_k \log(a_k + 1)} \sum_{n=a_k+1}^{\infty} \left(\frac{1}{n - a_k} - \frac{1}{n}\right),$$

which telescopes to

$$\sum_{k=1}^{\infty} \frac{1}{a_k \log(a_k + 1)} \sum_{m=1}^{a_k} \frac{1}{m} \leqslant \sum_{k=1}^{\infty} \frac{1 + \log a_k}{\log(a_k + 1)} \left(\frac{1}{a_k}\right).$$

But $\lim_{k \to \infty}(1 + \log a_k)/\log(a_k + 1) = 1$, so this sum converges also, and we are done.

However, $\lim_{n \to \infty} f(n) = 0$ always fails since $f(a_l + 1) \geqslant 1$ for all l.

(b) For $n > a_1$ let $p(n) = \sum_{a_k < n} \frac{1}{n - a_k}$ and $q(n) = \sum_{a_k > n} \frac{1}{a_k - n}$, and choose the greatest $r(n)$ and smallest $s(n)$ such that $a_{r(n)} \leqslant n \leqslant a_{s(n)}$. Since $0 \leqslant p(n) \leqslant f(n)$ and $0 \leqslant q(n) \leqslant f(n)$, $\liminf_{n \to \infty} p(n) = 0$ and $\liminf_{n \to \infty} q(n) = 0$, by (a). Thus, given $l, \varepsilon > 0$ we can pick $m > a_l$ and

$n > a_l$ such that $p(m) < \varepsilon$ and $q(n) < \varepsilon$. Then by comparing the sums termwise, we see that

$$p(a_{s(m)}) \leqslant p(m) < \varepsilon,$$

and

$$q(a_{r(n)}) \leqslant q(n) < \varepsilon.$$

But $s(m) \geqslant l$ and $r(n) \geqslant l$, so $\liminf_{l \to \infty} p(a_l) = \liminf_{l \to \infty} q(a_l) = 0$, as desired.

However, if we choose $a_{2k-1} = 2^k$ and $a_{2k} = 2^k + 1$, then

$$\sum_{k=1}^{\infty} \frac{1}{a_k} \leqslant 2 \sum_{k=1}^{\infty} 2^{-k} = 2,$$

but

$$\liminf_{l \to \infty} \sum_{\substack{k=1 \\ k \neq l}}^{\infty} \frac{1}{|a_k - a_l|} = 0$$

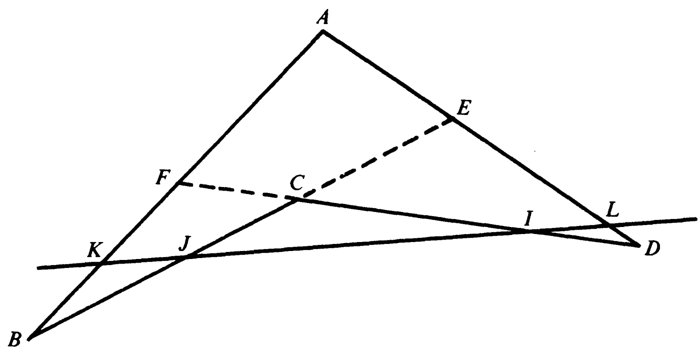
fails since for any l there exists k such that $|a_k - a_l| = 1$.

Also solved by the proposer.

A Quadrilateral Problem — September 1985

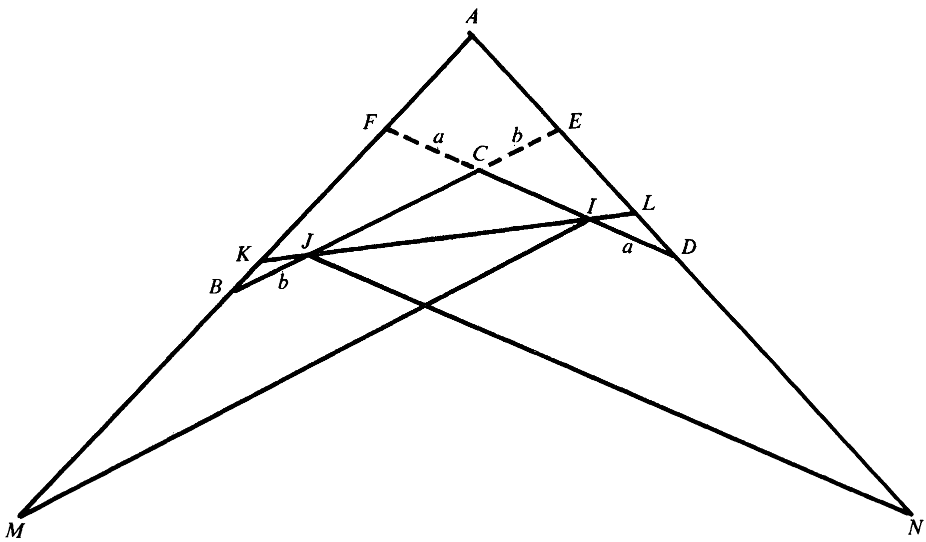
1224. *Proposed by Aristomenis Siskakis, Purdue University at Indianapolis.*

Consider the nonconvex quadrilateral $ABCD$ as in the FIGURE. Let I and J be chosen so that $DI = CF$ and $BJ = CE$. Let K and L be the points where the line IJ intersects AB and AD, respectively. Show that $KJ = IL$.



I. *Solution by Bjorn Poonen, student, Harvard College.*

Choose M on line AB and N on line AD such that MI is parallel to BE and NJ is parallel to DF. Let $a = DI = CF$ and $b = BJ = CE$.



Then we have

$$\triangle KJB \sim \triangle KIM, \qquad \frac{KJ}{b} = \frac{KJ + JI}{MI},$$

$$\triangle LJN \sim \triangle LID, \qquad \frac{JI + IL}{NJ} = \frac{IL}{a},$$

$$\triangle NJE \sim \triangle DCE, \qquad \frac{NJ}{b + CJ} = \frac{a + IC}{b},$$

$$\triangle BCF \sim \triangle MIF, \qquad \frac{b + CJ}{a} = \frac{MI}{a + IC}.$$

Multiplying the four equations we get

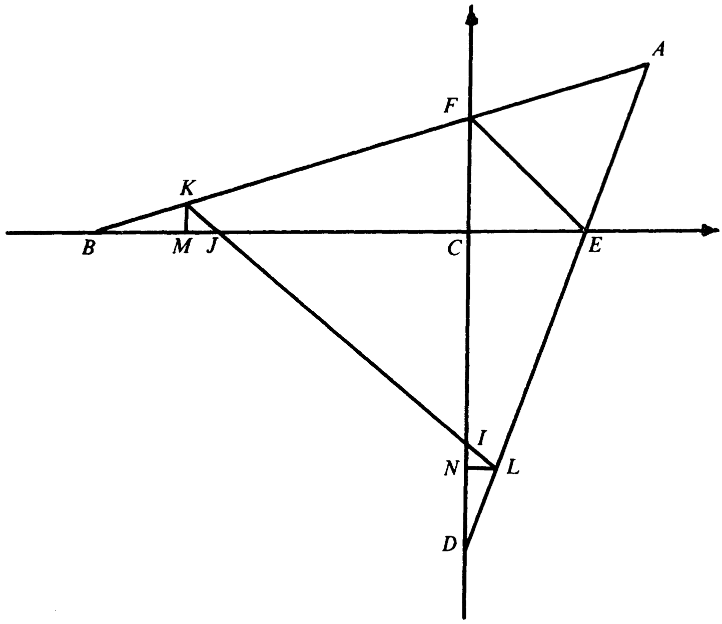
$$\frac{(KJ)(JI) + (KJ)(IL)}{ab} = \frac{(IL)(JI) + (KJ)(IL)}{ab}$$

which easily yields $KJ = IL$.

II. *Solution by Howard Eves, University of Maine.*

Since equality of segments on any given line is preserved under orthogonal projection, we shall simplify the figure by an appropriate such projection. Now any triangle can be orthogonally projected into a right isosceles triangle with the right angle at any chosen vertex of the triangle. Let us, then, orthogonally project the figure of the theorem into one in which triangle FCE

becomes a right isosceles triangle with right angle at vertex C (see FIGURE). We now show that in this projected figure $KJ = IL$. Toward this end we superimpose a rectangular Cartesian coordinate system on the figure with the positive x-axis along CE and the positive y-axis along CF, and choose $CE = CF = 1$. Then, designating BC by a and DC by b, we have the following coordinates for some of the points of the figure:



$$C\colon (0,0),\ E\colon (1,0),\ F\colon (0,1),\ B\colon (-a,0),\ D\colon (0,-b),$$
$$J\colon (1-a,0),\ I\colon (0,1-b).$$

Using these coordinates we find, by the intercept form for the equation of a line, the following equations of lines BF, DE, IJ:

$$BF\colon\ x - ay = -a,$$
$$DE\colon\ -bx + y = -b,$$
$$IJ\colon\ (1-b)x + (1-a)y = (1-a)(1-b).$$

From the equations of BF and IJ we find the y-coordinate of K to be

$$MK = (1-b)/(1-ab).$$

From the equations of DE and IJ we find the y-coordinate of L to be

$$CN = ab(1-b)/(ab-1).$$

We now have

$$NI = NC - IC = ab(1-b)/(1-ab) + (1-b) = (1-b)/(1-ab) = MK.$$

It follows that the two similar right triangles KMJ and INL are congruent, whence $KJ = IL$.

Also solved by Frank B. Allen, Jordi Dou (Spain), Howard Eves (second solution), Clifford S. Gardner, P. L. Hon (Hong Kong), H. Koter (Japan), L. Kuipers (Switzerland), Mou-liang Kung, Syrous Marivani, Christopher

Peterson, Daniel M. Rosenblum, Harry D. Ruderman, Paulette Tino, and the proposer.

There was a wide variety of approaches to this problem: synthetic, analytic, trigonometric, projective, vector. Most solutions were based on two applications of Menelaus' Theorem (to triangle *CJI*, first with transversal *BKF* and then with transversal *DLE*). *Rosenblum* showed that the result holds for *any* quadrilateral (nonconvex, convex, or self-intersecting).

A Series Inequality **September 1985**

1225. *Proposed by Robert E. Shafer, Berkeley, California.*

Show that

$$\sum_{n=1}^{\infty} \log^2\left(1+\frac{x}{n}\right) > 4\left(\log\Gamma(1+x) - 2\,\Gamma\left(1+\frac{x}{2}\right)\right)$$

for real $x > 0$.

Solution by Kee-wai Lau, Hong Kong.

By using the facts that $\Gamma(1+x) = x\Gamma(x)$ and

$$\Gamma(x) = x^{-1}\prod_{n=1}^{\infty}\left(1+\frac{1}{n}\right)^x\left(1+\frac{x}{n}\right)^{-1},$$

we find that the right-hand side of the desired inequality is equal to

$$4\left(\sum_{n=1}^{\infty}\left(x\log\left(1+\frac{1}{n}\right) - \log\left(1+\frac{x}{n}\right)\right) - 2\sum_{n=1}^{\infty}\left(\frac{x}{2}\log\left(1+\frac{1}{n}\right) - \log\left(1+\frac{x}{2n}\right)\right)\right)$$

$$= 4\sum_{n=1}^{\infty}\left(2\log\left(1+\frac{x}{2n}\right) - \log\left(1+\frac{x}{n}\right)\right).$$

Thus it suffices to show that

$$\sum_{n=1}^{\infty}\left[\log^2\left(1+\frac{x}{n}\right) - 8\log\left(1+\frac{x}{2n}\right) + 4\log\left(1+\frac{x}{n}\right)\right] > 0 \quad \text{for } x>0. \tag{1}$$

For $y \geqslant 0$, let $f(y) = \log^2(1+2y) - 8\log(1+y) + 4\log(1+2y)$ and $g(y) = \log(1+2y) - 2y/(1+y)$. We have

$$\frac{df(y)}{dy} = \frac{4g(y)}{1+2y}, \text{ and } \frac{dg(y)}{dy} = \frac{2y^2}{(1+2y)(1+y)^2} > 0.$$

This together with $f(0) = g(0) = 0$ implies that for $y > 0$, $g(y) > 0$, $df(y)/dy > 0$, and $f(y) > 0$. It follows that $\sum_{n=1}^{\infty} f(x/2n) > 0$ for $x > 0$. This proves inequality (1) and the proof is complete.

Also solved by Paul Bracken, Chico Problem Group, William A. Newcomb, Bjorn Poonen (student), Volkhard Schindler (East Germany), Robert J. Wagner, and the proposer.

Answers

A713. If $\mathbf{A}, \mathbf{B}, \mathbf{C}$ denote vectors from the center of the sphere to the vertices of the spherical triangle, we have to show equivalently that $|\mathbf{A}| = |\mathbf{B}| = |\mathbf{C}|$ and $(\mathbf{B}+\mathbf{C})\cdot(\mathbf{C}+\mathbf{A}) = 0$ imply that $(\mathbf{B}+\mathbf{C})\cdot(\mathbf{A}+\mathbf{B}) = 0$ and $(\mathbf{A}+\mathbf{B})\cdot(\mathbf{A}+\mathbf{C}) = 0$.

Since $(\mathbf{B}+\mathbf{C})\cdot(\mathbf{C}+\mathbf{A}) - (\mathbf{B}+\mathbf{C})\cdot(\mathbf{A}+\mathbf{B}) = (\mathbf{B}+\mathbf{C})\cdot(\mathbf{C}-\mathbf{B}) = \mathbf{C}^2 - \mathbf{B}^2 = 0$, and $(\mathbf{B}+\mathbf{C})\cdot(\mathbf{C}+\mathbf{A}) - (\mathbf{A}+\mathbf{B})\cdot(\mathbf{A}+\mathbf{C}) = (\mathbf{C}-\mathbf{A})\cdot(\mathbf{A}+\mathbf{C}) = 0$, the result is now immediate.

A714. If we call the 99-sided polygon $A_1A_2\ldots A_{99}$ and agree that A_{99} and A_1 are consecutive vertices, then any selection of 50 or more vertices must include at least two that are consecutive. We could not have 50 or more vertices inside the circle since the side joining a pair of consecutive vertices inside the circle would also be inside the circle. Hence, 50 or more vertices must be outside the circle, and the side joining a pair of consecutive vertices among them will be tangent to the circle.

REVIEWS

PAUL J. CAMPBELL, Editor
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles and books are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of the mathematics literature. Readers are invited to suggest items for review to the editors.

Euler, Leonhard, *Elements of Algebra*, translated by John Hewlett, Springer-Verlag, 1984; xxxix + 592 pp.

Reprint of an 1840 translation, with a biography by C. Truesdell in 1972 (Fig. 22, mentioned on p. xxxiv, was a casualty of the reprinting). The writing is remarkably clear, as one might expect of a work that after Euclid is "the most widely read of all books on mathematics, having been printed at least thirty times..." (Truesdell, xxxiii). Unfortunately, the readership includes few in this century; perhaps the new printing will renew interest in this, one of the very few of Euler's works available in English.

O'Donnell, Sean, *William Rowan Hamilton: Portrait of a Prodigy*, Boole Pr, 1983; xvi + 224 pp.

Only in 1980, a hundred years after the first, did a second biography of Hamilton appear (T. L. Hankins, *Sir William Rowan Hamilton*, Johns Hopkins U Pr, 1980); here we have a third. O'Donnell says Hankins "has clearly produced the definitive biography of Hamilton in the context of the history of mathematics." So--why O'Donnell's book? "It is an attempt to come to grips with Hamilton's formidable personality rather than his achievements to further an understanding of the individual rather than what he did." O'Donnell's general aim is to understand genius better; his psychological portrait provides fresh perspective and may bring Hamilton to the attention of non-mathematicians.

Schroeder, M. R., *Number Theory in Science and Communication, with Applications in Cryptography, Physics, Biology, Digital Information, and Computing*, Springer-Verlag, 1984; ix + 324 pp.

"...[T]he theory of integers can provide totally unexpected answers to real-world problems...." Among the applications are design of concert halls, random number generation, public-key encryption, error-correcting codes, the creation of artistic graphics, and the confirmation of slowing of electromagnetic waves in gravitational fields (as predicted by general relativity). Although the book is billed as an introduction for non-mathematicians, its level presumes a mathematically literate reader.

van der Waerden, B. L., *A History of Algebra, from al-Khwārizmī to Emmy Noether*, Springer-Verlag, 1985; xi + 271 pp, $34.50.

Succinct history of algebra, concentrating on algebraic equations, groups, and algebras. (Mathematical trivia question: What do the initials "B.L." stand for?)

Cullen, M. R., *Linear Models in Biology: Linear Systems Analysis with Biological Applications*, Wiley, 1985; 213 pp, $19.95 (P).

Beginning with compartment models, proceeding via eigenvalue analysis and computer implementation, and concluding with sensitivity analysis, this excellent text demonstrates the usefulness in biology of linear algebra, differential and difference equations, and linear programming. The modeling process itself is not discussed ("such discourses are poorly understood"). Among applications discussed: diffusion, population projection, optimal harvesting, tracer methods, and systems ecology.

Stein, Dorothy, *Ada: A Life and a Legacy*, MIT Press, 1985; xix + 321 pp, $19.95.

Thoroughly revisionist biography of Ada Lovelace, solidly based on copious quotations from the Lovelace Papers and other original sources. According to Stein, not only was Ada's true achievement much less than is popularly acclaimed: she was in fact "unable to assimilate the symbolic processes" (p. 84) of mathematics, despite ten years of tutelage by Somerville, Babbage, DeMorgan, and others. Stein finds that Lovelace lost a fortune gambling but not for the sake of funding Babbage's Engines, nor was Babbage involved in the gambling. Stein also asserts that Lovelace's "Notes" on Babbage's Analytical Engine are read anachronistically by modern interests, since the ideas, themes, and presentation were much more inspired by the metaphor of the calculating machine as harbinger of economic progress than by the concept of a programmable machine (p. 93). Here Stein--a psychologist--shows her preference in discerning motives; most of the book deals with the intricacies of schemings in Lovelace's family's affairs, in which psychological explanations serve admirably.

Bentley, Jon, *Programming Pearls*, Addison-Wesley, 1986; viii + 195 pp (P).

Substantially-revised versions of Bentley's tremendous columns from *Communications of the Association for Computing Machinery*. He treats programming fundamentals (problem definition, algorithms, data structures, verification), efficiency (5 columns), and applications of the enunciated principles (4 more). "The book covers much of the material in a college algorithms course, but the emphasis is more on applications and coding than on mathematical analysis." The basic theme: "thinking hard about programming can be both useful and fun." Bentley even advises how to read the book: "... don't go too fast. Read them well, one per sitting. Try the problems..."

Board on Mathematical Sciences. *Mathematical Sciences: A Unifying and Dynamic Resource*. State of the Art Reviews, National Academy Press, 1986; ix + 35 pp, (P). (Available from Board on Mathematical Sciences, 2101 Constitution Ave. N.W., Washington, DC 20418). Reprinted in *Notices of the American Mathematical Society* 33:3 (June 1986) 462-479.

At the request of NSF, in 1985 the National Research Council initiated state-of-the-art reviews of rapidly-evolving areas of science and technology. Mathematics, cell biology, and materials science were chosen for review. This report emphasizes potential erosion of the leading role of the U.S. in mathematical research, internal unification of mathematics, flourishing applications, and symbiosis with computing. Six short vignettes are offered, on D-modules, computational complexity, nonlinear hyperbolic conservation laws, the Yang-Mills equations, operator algebras, and survival analysis. The discussions of these are relatively non-technical but disconcertingly impersonal: one reads of discoveries by "a Taiwanese mathematician now living in the United States," "an American mathematician, born in New Zealand and educated in Switzerland"; N. Karmarkar becomes "a young Indian mathematician in the United States," and the formulators of NP-completeness are just "a Canadian and a former Soviet citizen who now resides in the United States." No doubt NSF--not to mention us and our students--would have found a few identifying footnotes valuable, instead of this game of "Guess the Greats."

Harte, John, *Consider a Spherical Cow: A Course in Environmental Problem Solving*, William Kaufmann, 1985; xvi + 283 pp, $24.95, $12.95 (P).

How many cobblers are there in the U.S.? How long will it take to use up the world's petroleum? What fraction of the earth's plant growth is eaten by humans? What is the residence time of water vapor in the atmosphere? How much electricity could be produced by burning everybody's junk mail? How much warmer is a city than the countryside? This marvelous book takes up these problems, along with acid rain, deforestation/desertification, the greenhouse effect, nuclear winter, and a couple dozen others, in an effort to teach quantitative problem-solving. Harte promotes the "art of the reasonable guess" and the "back of the envelope" calculation (no need for computers). High-school mathematics suffices for 80% of the problems; a few others use differential equations. Students will need to become comfortable with a variety of scientific units; concepts from physics and chemistry demand at least a high-school background in each. Included in the book are exercises (plus answers), an appendix of "useful numbers," a glossary, and a bibliography.

Schoenfeld, Alan H., *Mathematical Problem Solving*, Academic, 1985; xvi + 409 pp, $58.00, $29.95 (P).

"...What does it mean to 'think mathematically'? and How can we help students to do it? Those are the issues at the core of this book..." The book features a framework for problem-solving behavior, comprising four aspects: cognitive resources, heuristics, control, and belief systems. The book includes a series of empirical studies that document productive and unproductive student problem-solving behavior. This is a major new study on problem-solving; it demonstrates that problem-solving can be learned when teaching focuses on the higher-order skills necessary for using mathematics.

Kilpatrick, Jeremy, *Academic Preparation in Mathematics*, The College Board, 1985; vi + 86 pp, $6.95 (P).

In 1983 the College Board published *Academic Preparation for College: What Students Need to Know and Be Able to Do*, "an agenda for high schools to pursue." Subsequent comparison volumes go into greater detail, as the book at hand does on mathematics topics. In particular, it offers suggestions on how to achieve the recommended outcomes, including a three-year course sequence of basic topics in computing, statistics, algebra, geometry, and functions.

Guile, Bruce R. (ed.), *Information Technologies and Social Transformation*, National Academy Press, 1985; viii + 173 pp (P).

Most anthologies on computers and society, whatever their intended audience, pre-vision futures that are extrapolations of the present, usually clothed in shining faith in technology. Insight is rare. Three essays in this collection fit the mold; the other three merit reading and reflection. M. Kranzberg compares the Information Age to the Industrial Revolution and suggests we are observing change evolutionary in timescale (the benefits will not appear overnight) but revolutionary in effects. (Kranzberg, founder and president of the Society for the History of Technology, is one of the few writers on computers and society who allude to earlier literature on technology and culture, by W. F. Ogburn, Lewis Mumford, and others.) H. Cleveland discusses how the information society is eroding hierarchies based on control, secrecy, ownership, access, and geography. Finally, A. W. Branscomb offers a masterful discussion of property rights in information.

Beasley, John D., *The Ins and Outs of Peg Solitaire*, Oxford U Pr, 1985; xii + 275 pp, $16.95.

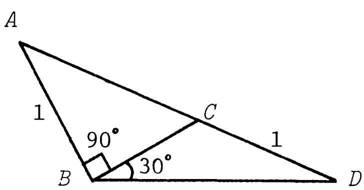
The authoritative book on 33-hole peg solitaire, including mathematical analysis and extension to boards of other sizes and dimensions.

NEWS & LETTERS

EIGHTEENTH CANADIAN MATHEMATICS OLYMPIAD

The problems of the Eighteenth Canadian Mathematics Olympiad, held May 7, 1986, have been provided by M.S. Klamkin and are reprinted here courtesy of the Crux Mathematicorum.

1.



In the diagram AB and CD are of length 1 while angles ABC and CBD are 90° and 30° respectively. Find AC.

2. A Mathlon is a competition in which there are M athletic events. Such a competition was held in which only A, B and C participated. In each event p_1 points were awarded for first place, p_2 for second and p_3 for third where $p_1 > p_2 > p_3 > 0$ and p_1, p_2, p_3 are integers. The final score for A was 22, for B was 9 and for C was also 9. B won the 100 metres. What is the value of M and who was second in the high jump?

3. A chord ST of constant length slides around a semicircle with diameter AB. M is the mid-point of ST and P is the foot of the perpendicular from S to AB. Prove that angle SPM is constant for all positions of ST.

4. For positive integers n and k, define $F(n,k) = \sum_{r=1}^{n} r^{2k-1}$. Prove that $F(n,1)$ divides $F(n,k)$.

5. Let $u_1, u_2, u_3, \ldots$ be a sequence of integers satisfying the recurrence relation $u_{n+2} = u_{n+1}^2 - u_n$. Suppose $u_1 = 39$ and $u_2 = 45$. Prove that 1986 divides infinitely many terms of the sequence.

FIFTEENTH U.S.A. MATHEMATICAL OLYMPIAD

The problems of the 15th U.S.A. Mathematical Olympiad have been provided by M.S. Klamkin and Walter Mientka.

1. Part a. Do there exist 14 consecutive positive integers each of which is divisible by one or more primes p from the interval $2 \leq p \leq 11$?

Part b. Do there exist 21 consecutive positive integers each of which is divisible by one or more primes p from the interval $1 \leq p \leq 13$?

2. During a certain lecture, each of five mathematicians fell asleep exactly twice. For each pair of these mathematicians, there was some moment when both were sleeping simultaneously. Prove that, at some moment, some three of them were sleeping simultaneously.

3. What is the smallest integer n, greater than one, for which the root-mean-square of the first n positive integers is an integer?

Note. The root-mean square of n numbers $a_1, \ldots, a_n$ is defined to be $[(a_1^2 + \ldots + a_n^2)/n]^{1/2}$.

4. Two distinct circles K_1 and K_2 are drawn in the plane. They intersect at points A and B, where AB is a diameter of K_1. A point P on K_2 and inside of K_1 is also given.

Using only a "T-square" (i.e. an instrument which can produce the straight line joining two points and the perpendicular to a line through a point on or off the line), find an explicit construction for:

two points C and D on K_1 such that CD is perpendicular to AB and CPD is a right angle.

5. By a <u>partition</u> π of an integer $n \geq 1$, we mean a representation of n as a sum of one or more positive integers, where the summands must be put in nondecreasing order. (E.g. if $n = 4$, then the partitions π are $1 + 1 + 1 + 1$, $1 + 1 + 2$, $1 + 3$, $2 + 2$, and 4.)

For any partition π, define $A(\pi)$ to be the number of 1's which appear in π, and define $B(\pi)$ to be the number of distinct integers which appear in π. (E.g. if $n = 13$ and π is the partition $1 + 1 + 2 + 2 + 2 + 5$, then $A(\pi) = 2$ and $B(\pi) = 3$.)

Prove that, for any fixed n, the sum of $A(\pi)$ over all partitions π of n is equal to the sum of $B(\pi)$ over all partitions π of n.

The U.S.A.M.O. was set by J. Konhauser, A. Liu, G. Patruno, and I. Richards, chairman.

REPORT ON THE 27th INTERNATIONAL MATHEMATICAL OLYMPIAD

The following report on the 27th IMO is excerpted from an article by M.S. Klamkin in *Crux Mathematicorum*.

The Twenty-Seventh International Mathematical Olympiad (IMO) was held this year in Warsaw, Poland from July 4 to July 15. Teams from 37 countries took part in the competition, one less than the record number of 38 countries of last year. The maximum team size for each country was 6 students, the same as for the last three years. The 1987 and 1988 IMO's are to be held in Cuba and Australia, respectively.

The six problems of the competition were assigned equal weights of 7 points each (the same as in the last five IMO's) for a maximum possible individual score of 42 and a maximum possible team score of 252.

First, Second, and Third Prizes were awarded to students with scores in the respective intervals [34-42], [26-33], and [17-25]. Congratulations to the following 18 First-Prize winners:

Name	Country	Score
Kos Geza	Hungary	42
Vladimir Roganov	USSR	42
Stanislav Smirnov	USSR	42
Fang Weimin	China	41
Jurg Jahnel	East Germany	41
Joseph Keane	USA	41
Marius Dabija	Rumania	40
Zhang Hao	China	39
Bruno Savalle	France	38
Ralph C. Teixera	Brazil	37
Li Ping Li	China	37
Martin A. Harterich	West Germany	36
David Grabiner	USA	36
Jeremy A. Kahn	USA	35
Iliya T. Kraichev	Bulgaria	34
Wieland E. Fischer	West Germany	34
Nicolae I. Beli	Rumania	34
Ha Anh Vu	Vietnam	34

As the IMO competition is an individual event, the results are announced officially only for individual team members. However, team standings are usually compiled unofficially by adding up the scores of the individual team members. The team results are given in the following table. Congratulations to the tying winning teams, the USA and the USSR:

Rank	Country	Score (max 252)	Total Prizes
1,2	USA	203	6
1,2	USSR	203	6
3	West Germany	196	6
4	China	177	5
5	East Germany	172	6
6	Rumania	171	5
7	Bulgaria	161	6
8	Hungary	151	5
9	Czechoslovakia	149	6
10	Vietnam	146	5
11	Great Britain	141	5
12	France	131	4
13	Austria	127	4
14	Israel	119	4
15	Australia	117	5
16	Canada	112	3
17	Poland	93	3
18	Morocco	90	3
19	Tunisia	85	1
20	Yugoslavia	84	2

It is of interest to note that (i) Joseph Keane (USA) also won the only special prize for his solution of problem #3, (ii) Terence Tao (Australia) was the youngest student (10 yrs.) of the competition and won a third prize with a score of 19, (iii) China which only entered the competition for the first time last year with two students did very well this time coming in 4th with a score of 177.

The problems of this year's competition are given below.

1. Let d be any positive integer not equal to 2, 5 or 13. Show that one can find distinct a,b in the set $\{2,5,13,d\}$ such that $ab-1$ is not a perfect square.

2. A triangle $A_1A_2A_3$ and a point P_0 are given in the plane. We define $A_s = A_{s-3}$ for all $s \geq 4$. We construct a sequence of points P_1, P_2, P_3, ... such that P_{k+1} is the image of P_k under rotation with center A_{k+1} through angle 120° clockwise (for $k = 0,1,2,\ldots$). Prove that if $P_{1986} = P_0$ then the triangle $A_1A_2A_3$ is equilateral.

3. To each vertex of a regular pentagon an integer is assigned in such a way that the sum of all the five numbers is positive. If three consecutive vertices are assigned the numbers x,y,z respectively and $y < 0$ then the following operation is allowed: the numbers x,y,z are replaced by $x+y$, $-y$, $z+y$ respectively. Such an operation is performed repeatedly as long as at least one of the five numbers is negative. Determine whether this procedure *necessarily* comes to an end after a finite number of steps.

4. Let A,B be adjacent vertices of a regular n-gon ($n \geq 5$) in the plane having center at O. A triangle XYZ, which is congruent to and initially coincides with OAB, moves in the plane in such a way that Y and Z each trace out the whole boundary of the polygon, X remaining inside the polygon. Find the locus of X.

5. Find all functions f, defined on the non-negative real numbers and taking non-negative real values, such that:

(i) $f[xf(y)]\cdot f(y) = f(x+y)$ for all $x,y \geq 0$,

(ii) $f(2) = 0$,

(iii) $f(x) \neq 0$ for $0 \leq x < 2$.

6. One is given a finite set of points in the plane, each point having integer coordinates. Is it always possible to color some of the points in the set red and the remaining points white in such a way that for any straight line L parallel to either one of the coordinate axes the difference (in absolute value) between the numbers of white points and red points on L is not greater that 1? Justify your answer.

NCTM SEEKS AUTHORS

The Educational Materials Committee of the National Council of Teachers of Mathematics announces that the 1989 Yearbook will be "Elementary School Mathematics: Issues and Directions," and will be edited by Professor Paul Trafton of National College of Education. The yearbook Advisory Panel is now seeking manuscripts for the yearbook. They are interested both in substantive papers addressing issues or directions in teaching of elementary school mathematics and in relatively short papers relating classroom practices to these issues and/or directions. Guidelines for the presentation of manuscripts are available from the General Editor, Albert P. Shulte, Oakland Schools, 2100 Pontiac Lake Road, Pontiac, MI 48054.

* *

LETTERS TO THE EDITOR

Dear Editor:

In the note "Sections of n-dimensional cones", this journal 42(1969) 80-83, there are rather inelegant linear algebra computational proofs that the intersection of a right circular cone and a plane is of 2nd degree (a conic) and its extension for sections of a spherical cone in E^n.

More generally and easy to show is that the intersection of a hyperplane H and a quadric Q, both in E^n, is, as expected, at most a quadric of dimension n-1. In special cases, the intersection can be a hyperplane. For example in E^3, the intersection of the parabolic cylinder $z = x^2$ with the plane $x = k$ is the line $x = k$, $z = k^2$. For a simple proof, consider a rotation of the coordinate system such that one of the coordinate planes becomes parallel to the hyperplane. Since the transformation equations are linear, the new equation of Q is still one of 2nd degree. The equation for the hyperplane is now $x_1 = k$. The intersection is now gotten by replacing x_1 in the equation for Q by k giving the desired result.

M.S. Klamkin
University of Alberta

Dear Editor:

With regard to the generalization of the Pythagorean Theorem in our joint paper [1986, p. 149], Richard Guy has pointed out that a somewhat earlier reference was overlooked. In addition to the two references given in the article, he notes that the result may be found in the commentary of T.L. Heath following Proposition 13 in Book II of Euclid's Elements. This appears on p. 407 of Vol. I of the Dover edition of 1956.

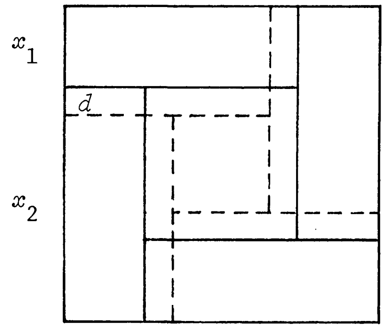
G.D. Chakerian
University of California
Davis

M.S. Klamkin
University of Alberta

Dear Editor:

In *Mathematics Magazine*, vol. 59, Nr. 1, p. 11 (Proof without words: ... by D. Schattschneider) a visualization is given for $4a\cdot b < (a+b)^2$. The same picture can induce a very simple proof of the arithmetic mean-geometric mean inequality in n (≥ 2) variables:

$$x_1,\ldots,x_n \in R^+ \Rightarrow \prod_{i=1}^{n} x_i \leq \left(\sum_{i=1}^{n} x_i/n\right)^n$$



Assume $x_1 < x_2$. If x_1 is increased and x_2 decreased by the same amount d, $0 < d < x_2-x_1$, then

$$4x_1\cdot x_2 < 4(x_1 + d)\cdot(x_2 - d).$$

Now we assume $n \geq 3$ and one of the numbers x_i to be less than $m = \sum_{i=1}^{n} x_i/n$, so that another x_j must be greater than m. Without loss of generality we can assume $x_1 < m < x_2$. Choosing $d = m - x_1$, we obtain

$$x_1\cdot\ldots\cdot x_n < m\cdot(x_2 - d)\cdot x_3\cdot\ldots\cdot x_n.$$

If $x_2 - d$, x_3, ..., x_n are not all equal, we proceed as before and after no more than n-1 steps we have

$$x_1\cdot\ldots\cdot x_n < \left(\frac{x_1 + \ldots + x_n}{n}\right)^n$$

The products on both sides are equal only if all x_i are equal.

Hermann Hering
Rheinisch-Westfälische
Technische Hochschule Aachen
Fachbereich 9
D-5100 Aachen, F.R. Germany

Mathematical Gems III

by Ross Honsberger

260 pp. Hardbound, 1985.

ISBN-0-88385-313-2

MAA Member: $21.00

List: $27.00

This book is packed with gems from elementary combinatorics, geometry, number theory, and probability. It will bring the enjoyment of good mathematics to the student as well as the teacher.

Glance at the Table of Contents and order now!

Order from:
The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, D.C. 20036

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, N.W.
Washington, DC 20036

MATHEMATICS MAGAZINE VOL. 59, NO. 4, June 1986