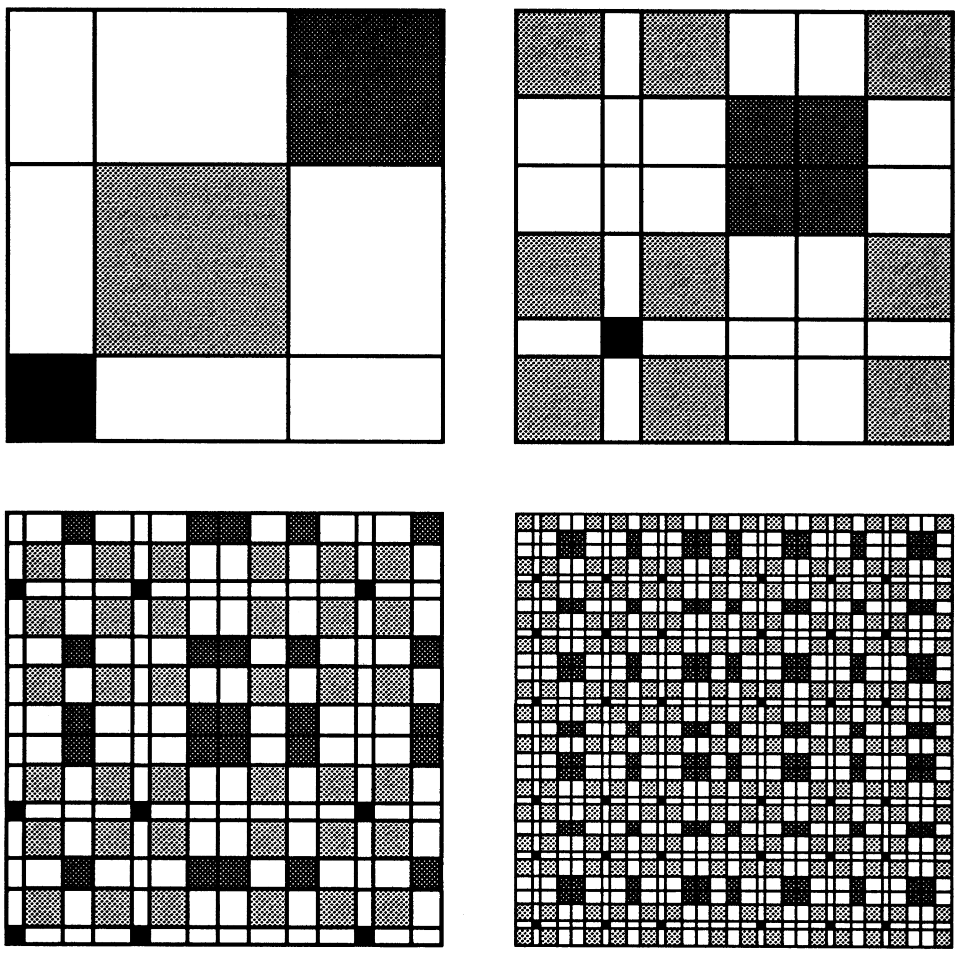
Vol. 70, No. 1 February 1997

MATHEMATICS MAGAZINE



- Combinatorial Group Theory and the Word Problem
- Moving into the Desert with Fibonacci
- Golden Fields: A Case for the Heptagon
- Where the Inflection Points of a Polynomial May Lie

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

EDITORIAL POLICY

Mathematics Magazine aims 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. Manuscripts on history are especially welcome, as are those showing relationships between various branches of mathematics and between mathematics and other disciplines.

A more detailed statement of author guidelines appears in this *Magazine*, Vol. 69, pp. 78–79, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Send new manuscripts to Paul Zorn, Editor, Department of Mathematics, St. Olaf College, 1520 St. Olaf Avenue, Northfield, MN 55057-1098. Manuscripts should be laser-printed, with wide line-spacing, and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit three copies and keep one copy. In addition, authors should supply the full five-symbol Mathematics Subject Classification number, as described in *Mathematical Reviews*, 1980 and later. Copies of figures should be supplied on separate sheets, both with and without lettering added.

Cover illustration: Aperiodic Tiling Sequence, by Peter Steinbach. One of a family of solutions to a problem in design economy posed by architect Le Corbusier: how to maximize repetition of similar figures in the least space.

AUTHORS

Martin Aigner received his Ph.D. from the University of Vienna in 1965. During the next five years spent in the United States he was introduced by C. C. Rota and the late R. C. Bose into the combinatorial world which he has never left since. He is currently a professor of mathematics at the Freie Universität Berlin.

Bruce Anderson received his Ph.D. from Cornell University in 1992, working under Moss Sweedler (whose inventive questions inspired this paper). He then taught a year at DePaul University before becoming an assistant professor at Kent State University, East Liverpool Campus.

His research veers toward searching for little uncovered mathematical gems accessible to undergraduates. He is particularly interested in polynomials and algorithms for working with polynomial systems. Outside of mathematics, he enjoys playing violin in chamber groups, and weaving web pages.

David Peifer received his Ph.D. from the University of Illinois in 1992. He taught for two years at St. Olaf College in Minnesota before joining the faculty at the University of North Carolina at Asheville. When he is not deep in contemplation of infinite groups, he likes to tend his garden, and hike with his family in the southern Appalachian Mountains.

Peter Steinbach, author of the book *Field Guide to Simple Graphs*, has been teaching secondary and undergraduate mathematics for ten years. He studied discrete mathematics with R. C. Entringer at the University of New Mexico and received his master's degree in 1992. Currently, he trains elementary teachers in mathematical form and practice at Albuquerque Technical-Vocational Institute. His interests include graph theory, combinatorial geometry, methods of teaching with manipulatives, and manipulative teaching methods. As a former industrial designer, he is especially motivated by intersections of mathematics with the Fine Arts.

Vol. 70, No. 1 February 1997

MATHEMATICS MAGAZINE

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

The annual subscription price for the MATHEMATICS MAGAZINE to an individual member of the Association is $16 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $64. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount; and new members receive a 40% dues discount for the first two years of membership.) The nonmember/library subscription price is $68 per year.

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

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

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

Postmaster: Send address changes to Mathematics Magazine, Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036-1385.

PRINTED IN THE UNITED STATES OF AMERICA

ARTICLES

An Introduction to Combinatorial Group Theory and the Word Problem

DAVID PEIFER
University of North Carolina at Asheville
Asheville, NC 28804

1. Introduction

As a result of the struggle to fit the essentials of the theory of abstract algebra into a one- or two-semester course, most undergraduate abstract algebra courses gloss over or entirely skip combinatorial group theory. The goal of this paper is to introduce some of the fundamental ideas and problems from this area of study at a level accessible to college mathematics students. Combinatorial group theory involves concepts from many different mathematical disciplines. The theory and techniques primarily involve a mix of algebra and logic, while the motivation and applications come from topology, geometry, and other areas of mathematics. This interplay among disciplines has attracted many mathematicians to work in this area.

Combinatorial group theory is the study of groups given in terms of *presentations*. A presentation is a set of group generators and a set of relations among the generators that completely describe the group. More details appear in Section 2. In many areas of mathematics, groups arise naturally in this form. In particular, presentations of groups appear frequently in topology. Much of the original motivation for the study of group presentations came from work in topology done around the turn of the century. In 1895, in order to better understand two-dimensional surfaces, H. Poincaré introduced the fundamental group of a surface. The fundamental group produced an important connection between topology and algebra. With it, mathematicians can use theorems and techniques from algebra to analyze questions from topology or geometry, and vice versa. For information on the fundamental group, see **[18]** and **[22]**. In 1905, working to distinguish and classify mathematical knots, W. Wirtinger first defined the knot groups. The *Scientific American* article **[15]** is a well-written introduction to knot groups for a general audience. In practice, both the fundamental group and the knot group arise naturally as group presentations. Today, combinatorial group theory still plays a significant role in a broad range of areas. Recent research in physics has uncovered important ties between physics and knot theory (see **[12]**). Group presentations also play fundamental roles in recent research in hyperbolic geometry and the topology of 3-manifolds (see **[6]** and **[9]**).

Section 2 begins with a brief introduction to free groups, leading to a precise definition of group presentations. In the next section Max Dehn's three original problems are stated; these problems signify the beginning of combinatorial group theory. A few of the most historically important theorems are stated. To illustrate the ideas, some solutions to Dehn's problems for specific groups are supplied. Combinatorial group theory is a large, rich area of study; this paper only touches on a small part of it. For more information, we recommend additional reading in **[2]**, **[10]**, **[13]**, **[14]**, and [22]. Although most of these books are intended for graduate students, [10] is written at the level of enthusiastic high school students.

2. Free Groups and Presentations

In combinatorial group theory, groups are typically given in terms of a group presentation. Let $X = \{x_1, x_2, \ldots, x_n\}$ be a set of distinct elements. For simplicity we will assume that X is finite and we will consider only finitely presented groups. Let $X^{-1} = \{x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\}$ be a set of elements, distinct from each other and from the elements of X. We call X the set of generators and X^{-1} the set of inverse generators. As the notation suggests, there is a one-to-one correspondence between the generators and inverse generators.

We can think of the set $X \cup X^{-1}$ as an alphabet, and each element as a letter. In this way, a string of elements from $X \cup X^{-1}$ is a word. For example, $x_1 x_1 x_2 x_2^{-1} x_2^{-1} x_1 x_2^{-1}$, is a word on $\{x_1, x_2\} \cup \{x_1^{-1}, x_2^{-1}\}$. It is customary to use exponentiation to denote a consecutive sequence of the same letter. Thus the word above could be written as $x_1^2 x_2 x_2^{-2} x_1 x_2^{-1}$. The *length* of a word is the number, counting multiplicity, of letters in the word. The notation $(X \cup X^{-1})^*$ denotes the set of all words of finite length on the alphabet $X \cup X^{-1}$. The word consisting of no letters, the *empty word*, is included in $(X \cup X^{-1})^*$. Words of the form $x_i x_i^{-1}$ or $x_i^{-1} x_i$, for any $i = 1, \ldots, n$, are called *inverse pairs*. We will form a group with elements from $(X \cup X^{-1})^*$. In order to satisfy the group axioms, we must restrict our attention to a subset of $(X \cup X^{-1})^*$. A word is said to be *reduced* if it contains no inverse pair subwords. The example given above is not a reduced word because it contains the subword, $x_2 x_2^{-1}$. The set of all reduced words in $(X \cup X^{-1})^*$ is denoted by F_n. We will show that, with the proper operation, this set forms a group known as the *free group on n generators.*

The most natural choice for the group operation on the set F_n is concatenation. That is, define the product of two words to be the result of tacking one word on the end of the other,

$$(x_{i1} x_{i2} \ldots x_{im})(x_{j1} x_{j2} \ldots x_{jl}) = (x_{i1} x_{i2} \ldots x_{im} x_{j1} x_{j2} \ldots x_{jl}).$$

However, this product may or may not result in a reduced word. An inverse pair may occur where the two words were joined. To avoid this problem, we define the product to be concatenation followed by deleting inverse pairs. Notice that a sequence of deletions may be required to reduce the product. With this convention, the product is closed on F_n.

One group axiom requires the existence of an identity element. By the definition of the product in F_n, the empty word is the identity. A second axiom requires that each element of the group have an inverse. We purposely defined the set X^{-1} so that each generator x_i has inverse x_i^{-1}. For an arbitrary word $\alpha \in F_n$, the inverse, α^{-1}, is obtained by reversing the order of the letters and reversing the exponents. For example, the inverse of $x_1 x_2^{-1} x_3$ is $x_3^{-1} x_2 x_1^{-1}$. The final axiom for groups states that the product must be associative. Because the definition of the product involves canceling inverse pairs, the proof of associativity requires checking a number of special cases. (We leave this as an exercise for the reader.) Since F_n satisfies these three axioms, F_n is a group.

Any group isomorphic to F_n, for n a positive integer, is called a *free group*. Many things are known about these groups. For example, it is easy to show that the free group on one generator, F_1, is the infinite cyclic group. It is also easy to show that for $n > 1$, F_n is a nonabelian infinite group. To see that F_n is not abelian, let x_1 and x_2 be two of the generators. Notice that $x_1 x_2 x_1^{-1} x_2^{-1}$ is reduced and not the empty word, so $x_1 x_2 x_1^{-1} x_2^{-1} \neq 1$. Multiplying by $x_2 x_1$ on both sides gives $x_1 x_2 \neq x_2 x_1$. Since x_1 and x_2 do not commute, F_n is not abelian. To show that F_n is infinite, notice that, for all integers m, x_1^m is reduced and not the empty word. Thus x_1 has infinite order, and F_n is infinite. The Nielsen–Schreier theorem gives a more surprising result: Every subgroup of a free group is again a free group. Proofs can be found in graduate-level textbooks (see **[20]** or **[21]**). For our purposes, the most important result is as follows:

THEOREM 1. *Every (finitely generated) group is the quotient group of a (finitely generated) free group by a normal subgroup.*

Here is a brief sketch of the proof. Let G be a finitely generated group; let $\{g_1, g_2, \dots, g_n\}$ be a generating set. Now consider the free group (with the same number of generators) F_n generated by $\{x_1, x_2, \dots, x_n\}$. There is a natural map ψ from F_n to G that maps the i^{th} generator of F_n to the i^{th} generator of G, and maps words in F_n as follows,

$$\psi(x_{i1} x_{i2} \dots x_{ir}) \to (g_{i1} g_{i2} \dots g_{ir}).$$

It is easy to show that ψ is a homomorphism, mapping F_n onto G. Therefore by the first isomorphism theorem for groups, G is isomorphic to the quotient of F_n by the normal subgroup $\ker(\psi)$. This proves the theorem. More thorough proofs of Theorem 1, as well as other basic material on free groups, can be found in undergraduate abstract algebra texts (see, e.g., **[7]** or **[8]**).

Theorem 1 illustrates the importance of normal subgroups of free groups. Here is an efficient way to describe a normal subgroup of F_n. Let $R = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$, where each α_i is an element of F_n. Define N to be the intersection of all normal subgroups of F_n that contain the elements of R. The set N is a normal subgroup of F_n, called the *normal closure* of R. Every element of N is a product of conjugates of elements of R. The subgroup N can be thought of as the smallest normal subgroup of F_n that contains R. If infinite sets are allowed for R, then any normal subgroup of F_n can be described in this manner. In general, one hopes to find the smallest possible set R to describe a given normal subgroup.

We can now define our main object of study. A *presentation* for a group G is a set of generators X of a free group F_n, and a set of relators R, words in F_n, such that the quotient group of F_n by the normal closure of R is isomorphic to G. We denote the presentation by $\langle X|R \rangle$, and write $G = \langle X|R \rangle$ to signify that G is isomorphic to the quotient group. It is clear from the definition that every presentation defines a group. By Theorem 1, every group has a presentation. In fact, any given group has infinitely many presentations.

Before we move on, we add a few remarks on presentations. Recall that a quotient group G/H is defined to be the group of cosets of H in G. In other words, we are defining an equivalence relation on G and assuming that every element of H is equivalent to the identity. What does this mean about the presentation $\langle X|R \rangle$?

To answer this question, return to the set of all words in $(X \cup X^{-1})^*$. To form the free group F_n, we considered only the reduced words, words that do not contain occurrences of $x_i x_i^{-1}$ or $x_i^{-1} x_i$, for any $i = 1, \dots, n$. Alternatively, we could consider the full set $(X \cup X^{-1})^*$, and define an equivalence relation by stating that two words are equivalent if one can be transformed into the other by a finite sequence of insertions or deletions of pairs, $x_i x_i^{-1}$ or $x_i^{-1} x_i$. It can be shown that this is indeed an equivalence relation, and that each equivalence class contains a unique reduced word (see [**14**]). This second approach also defines the free groups (up to isomorphism). The difference between the two approaches is that in the first, each group element has one representative, a reduced word; in the second, each group element is an entire equivalence class.

We can also think of the group G, given by a presentation $\langle X|R\rangle$, as a set of equivalence classes. Suppose, for instance, that $x_1 x_2^{-1} x_3$ is one of the relators in R. Then, in the group G, we want the relation $x_1 x_2^{-1} x_3 = 1$. So if any word contains the subword $x_1 x_2^{-1} x_3$ it is equivalent to the word with the occurrence of $x_1 x_2^{-1} x_3$ deleted. The relation $x_1 x_2^{-1} x_3 = 1$ also implies other relations, such as $x_1 = x_3^{-1} x_2$. Thus any word containing x_1 is equivalent to the same word with the occurrence of x_1 replaced by $x_3^{-1} x_2$. So we could have defined G as the set $(X \cup X^{-1})^*$ with equivalence classes given by insertions or deletions of pairs, $x_i x_i^{-1}$ or $x_i^{-1} x_i$, along with the equivalences defined by the relators.

Below are several examples of presentation of familiar groups.

- $\langle x_1 | x_1^n \rangle = \mathbb{Z}_n$, the cyclic groups of order n.
- $\langle x_1 \rangle = \mathbb{Z} = F_1$, the infinite cyclic group.
- $\langle x_1, x_2, \dots, x_n \rangle = F_n$, the free group on n generators.
- $\langle x_1, x_2 | x_1 x_2 x_1^{-1} x_2^{-1} \rangle = \mathbb{Z} \times \mathbb{Z}$, the free abelian group of rank 2.
- $\langle x_1, x_2 | x_1^n, x_2^2, x_1 x_2 x_1 x_2 \rangle = D_{2n}$, the dihedral group of order $2n$. Another presentation for this same group is $\langle x_1, x_2 | x_1^2, x_2^2, (x_1 x_2)^n \rangle$.

3. Dehn's Problems

The German-born mathematician Max Dehn lived from 1878 to 1952. Although Dehn's name is seldom mentioned in undergraduate courses, his work has influenced such research areas as geometry, topology, and group theory. Dehn was a student of David Hilbert, who posed twenty-three problems in his famous address to the Second Congress of Mathematics, held in Paris in 1900. Hilbert's problems were chosen to illustrate the kinds of work that would lead mathematics into the new century. Some of these problems still remain open after almost 100 years. In 1902, Dehn solved Hilbert's third problem, a question in geometry, involving the decomposition of tetrahedra. In 1907, Dehn worked with P. Heegaard to produce the first rigorous proof of the classification of two-dimensional surfaces [**5**]. In the years 1910–1914, Dehn published a series of papers on group theory and topology. The problems he posed and the techniques he developed shaped the study of combinatorial group theory and strongly influenced the study of low-dimensional topology up through the present day. Recently, these papers, along with some previously unpublished works of Dehn's, have been translated by John Stillwell and published as *Papers on Group Theory and Topology* [**4**].

Among other significant contributions to mathematics, Dehn posed three important problems, each one stemming from a specific topological question concerning fundamental groups of surfaces. For example, the word problem, stated below, corresponds to the topological problem of determining when a closed curve on a surface can be deformed in a continuous manner to a single point. Dehn posed these problems purely in terms of finite presentations. (A presentation is called *finite* if both the set of generators and the set of relators are finite.) Dehn's three problems are as follows:

The word problem *Let G be a group given by the finite presentation $\langle X|R\rangle$. Is there an algorithm that decides whether or not a given word is equivalent to the identity in G?*

The conjugacy problem *Let G be a group given by the finite presentation $\langle X|R\rangle$. Is there an algorithm that decides whether or not any pair of words, u and v, are conjugate in G? (Recall, u and v are conjugate if there exist $\rho \in G$ such that $u = \rho^{-1} v \rho$.)*

The isomorphism problem *Is there an algorithm that determines whether or not a pair of finite presentations define isomorphic groups?*

What exactly *is* an algorithm? A precise mathematical definition would involve a lengthy digression into language theory and mathematical logic. (Interested readers may see **[11]**.) Informally speaking, in Dehn's problems, "algorithm" means a procedure, given by a finite number of instructions, that produces an answer after a finite number of steps, without ever leaving any doubt as to the next step. If an algorithm exists for one of these problems, then we say the problem is *solvable*.

Example. The word problem for finitely generated free groups is solvable.

Proof. Let F_n be the free group given by the presentation $\langle x_1, x_2, \dots, x_n\rangle$. Let α be an element of $(X \cup X^{-1})^*$. It can be shown that to find the reduced element in the equivalence class of α requires only deleting inverse pairs, and that any order of deletions will result in the same reduced word (see **[14]**). Therefore we can find the reduced element equivalent to α by deleting the first inverse pair in α, then deleting the first inverse pair in the resulting word, and continuing until there are no more inverse pairs. Notice that this process must end in a finite number of steps because α has only a finite length and each deletion decreases the length by two. The word α is equivalent to the identity in G if and only if α reduces to the empty word.

Dehn's problems seem to be specific to the presentation given for the group. However, despite the fact that every group has an infinite number of presentations, having a solvable word problem is a group-theoretic property. That is to say, if a group has a solvable word problem for any particular finite presentation, then for *every* finite presentation of that group, there is an algorithm to solve the word problem. This is because there is a procedure, known as *Tietze transformation*, that gives an effective way to change between any two finite presentations (see **[14]**). Therefore we can say that the group has a solvable word problem, without reference to any specific presentation. Similarly, having a solvable conjugacy problem is a group-theoretic property.

Dehn himself solved these three problems for a specific set of groups. In particular, Dehn solved the word problem for the fundamental groups of closed, two-dimensional, orientable surfaces of genus $g \geq 2$. To get more general results for the word problem has turned out to be very difficult. During the first half of the century, many mathematicians worked to find general solutions to Dehn's problems. As with Dehn's results, most results in this area include strong restrictions on the groups and require ad hoc techniques that work only with specific sets of groups. In 1932, W. Magnus **[14]** proved one of the first general results.

THEOREM 2. *The word problem is solvable for any group that has a finite presentation with only one relator.*

There were many significant theorems and much work done in this area during the 1930's and 1940's. But few results were as general as Magnus's theorem. The reason for this became clear when, in 1954, P. S. Novikov **[17]** (and, independently, in 1958, W. W. Boone **[3]**) proved an astonishing result:

THEOREM 3. *There exists a finitely presented group with an unsolvable word problem.*

Novikov and Boone gave examples of finite presentations for which no algorithm can determine whether or not an arbitrary word is equivalent to the identity. A moment's reflection may help to explain the significance of Theorem 3. Determining whether a word is equivalent to the identity is perhaps the most basic question about a word that one could ask. There are more complicated questions. For example, does a given word represent an element of finite order? Does the word represent an element in the center of the group? Even more complicated are questions about the group itself. Is the group a p-group? Is the group abelian? Is the group trivial? Are there any subgroups of a specific kind? Is the group simple? Theorem 3 implies a dismal conclusion: Given an arbitrary finite presentation, there may be very little we can know about the group or even about any given element of the group.

The following example illustrates that even the trivial group, with one element, can be hard to recognize when given as a presentation.

Example. $H = \langle a, b \,|\, a^2ba^{-1}, ab^{-1}a^{-1} \rangle$ is a presentation of the trivial group.

Proof. If each generator is equivalent to the identity in H, then H must be the trivial group. Therefore, to prove the assertion it suffices to show that $a = b = 1$. The relators imply that

$$a^2ba^{-1} = 1 \quad \text{and} \quad ab^{-1}a^{-1} = 1.$$

Putting these two relations together gives

$$a^2ba^{-1} = 1 \Leftrightarrow a(aba^{-1}) = 1 \Leftrightarrow a(ab^{-1}a^{-1})^{-1} = 1 \Leftrightarrow a(1)^{-1} = 1 \Leftrightarrow a = 1.$$

Finally,

$$ab^{-1}a^{-1} = 1 \Leftrightarrow b^{-1} = 1 \Leftrightarrow b = 1.$$

Therefore H is the trivial group.

A set of *normal forms* for a group given by a specific presentation $\langle X|R \rangle$, is a set of words on $X \cup X^{-1}$ that includes one and only one representative for each group element. If there is an algorithm that can take any word and find its representative normal form, this provides a solution to the word problem. To determine if a word is equivalent to the identity, simply find its normal form and check if this is the normal form for the identity. The following example illustrates this technique.

Example. The word problem is solvable for the group given by the presentation $\langle a, b, c \,|\, aba^{-1}b^{-1}, bcb^{-1}c^{-1}, aca^{-1}c^{-1}, c^7 \rangle$.

Proof. This is a particularly nice presentation. The first three relators show that the generators all commute with each other, so the group is abelian. The only other relator shows that the third generator has order seven. Since there are no other relators, it is easy to see that this is the abelian group $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}_7$. With a little work it can be shown that

$$\{a^l b^m c^n \mid l, m, n \text{ are integers and } 0 \leq n \leq 6\}$$

is a set of normal forms for the group. There is a simple procedure to find the normal form for any word. Use the fact that the generators commute to gather all similar generators together. For example, by gathering similar generators, the word $b^2a^{-1}b^5c^{-1}a^{-2}c^{11}$ is transformed to $a^{-3}b^7c^{10}$. Now recall that c has order seven and reduce the exponent mod 7. The result is the normal form $a^{-3}b^7c^3$. Notice that the normal form for the identity is the empty word since each of the exponents must be zero. This gives a solution to the word problem for this group. Given a word, find the normal form. If the normal form is the empty word, then the word is equivalent to the identity; otherwise the word is not equivalent to the identity.

Theorem 3 implies that the word problem for finitely presented groups is unsolvable. That is, if we are given a finite presentation and no other information about the group, it is possible that there is no algorithm to solve the word problem for that group. Notice that Theorem 3 also implies a negative result for the conjugacy problem. If we cannot solve the word problem then we cannot determine whether a word is conjugate to the identity. In fact, Novikov **[16]** had already proved, a year before his 1955 paper, that the conjugacy problem for finitely presented groups is unsolvable. Using Theorem 3, in 1958 S. I. Adjan **[1]** and, independently, M. O. Rabin **[19]** showed that the isomorphism problem for finitely presented groups is also unsolvable.

Despite these negative results, the situation is not hopeless. As we can see from Theorem 2 and the examples above, the word problem can be solved in some cases. In fact, there are many positive results concerning finitely presented groups. The important thing is that we must have some extra knowledge about the group. For example, if we know that the group is abelian, then we can solve the word problem, conjugacy problem, and the isomorphism problem. (The proofs are too complicated to present here.) However, the main procedure is to reduce the presentation, using a technique called *Nielsen transformations*, down to a standard form. Once this is done, the structure of the abelian group can be seen from the new presentation. Therefore, by the fundamental theorem of finitely generated abelian groups we know exactly which abelian group is presented (see **[14]**). We can also answer other questions about abelian groups. For example, we can determine whether or not the group has any elements of finite order.

Theorem 3 was not the last word on Dehn's problems. There are still many open problems in this area. In the 1960's some of the original techniques developed by Dehn, in the early 1900's, were generalized. These techniques, called *small cancellation theory*, use graphs and two-dimensional diagrams to analyze the word and conjugacy problems (see **[13]**). Many other directions of study have also emerged. Mathematicians considered whether different types of algebraic combinations of groups would or would not preserve the property of having a solvable word or conjugacy problem. Other research considers complexity questions. For example, at what relative speeds do different algorithms solve the word problem? Studying groups given by presentations is still an important area. In the late 1970's the classification of mathematical knots was finally finished. The solution involved among other things, the solution to the word and conjugacy problem for the knot groups. In the late 1980s a number of new classes of groups were defined as tools to help understand 3-manifolds. The two most important are *automatic groups* (see **[6]**), and M. Gromov's *hyperbolic groups* (see **[9]**). These definitions are given in terms of group presentations. In recent years, many papers have been written analyzing these groups and determining which groups belong to one or another of these newly defined classes.

There is still work to be done. Open problems remain in all the different areas of combinatorial group theory. Some of these problems promise to be difficult and will require developing sophisticated new techniques. Other problems are just waiting for someone to discover the right trick.

REFERENCES

1. S. I. Adjan, On the algorithmic problem in effectively complete classes of groups, *Doklady Akad. Nauk SSSR* **123** (1958), 13–16. (Russian)
2. G. Baumslag, *Topics in Combinatorial Group Theory* (Lectures in Mathematics, ETH Zürich), Birkhäuser Verlag, 1993.
3. W. W. Boone, The word problem, *Ann. of Math.* **70** (1959), 207–265.
4. M. Dehn, *Papers on Group Theory and Topology*, Translated by John Stillwell, Springer-Verlag, New York, NY, 1987.
5. M. Dehn and P. Heegaard, *Analysis situs*, Encykl. Math. Wiss. vol III AB3 Leipzig (1907), 153–220.
6. D. B. A. Epstein, with J. W. Cannon, S. V. F. Levy, D. F. Holt, M. S. Paterson, W. P. Thurston, *Word Processing in Groups*, Jones and Bartlett, Boston, MA, 1992.
7. J. B. Fraleigh, *A First Course in Abstract Algebra*, fifth edition, Addison-Wesley, Reading, MA, 1994.
8. J. A. Gallian, *Contemporary Abstract Algebra*, third edition, D. C. Heath and Company, Boston, MA, 1994.
9. M. Gromov, *Hyperbolic groups*, Essays in Group Theory, ed. by S. M. Gersten, Springer-Verlag, New York, NY, 1987, pp. 75–264.
10. I. Grossman and W. Magnus, *Groups and Their Graphs*, New Mathematical Library no. 14, The Mathematical Association of America, 1964.
11. J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley Publishing Company, Reading, MA, 1979.
12. L. H. Kauffman, *Knots and Physics*, World Scientific, 1993.
13. R. Lyndon and P. E. Schupp, *Combinatorial Group Theory*, Springer-Verlag, New York, NY, 1977.
14. W. Magnus, A. Karrass and D. Solitar, *Combinatorial Group Theory*, Dover, New York, NY, 1966/1976.
15. L. Neuwirth, The theory of knots, *Scientific Amer.* **240** (1979), 110–124.
16. P. S. Novikov, Unsolvability of the conjugacy problem in the theory of groups, *Izv. Akad. Nauk. SSSR* **18** (1954), 485–524. (Russian)
17. P. S. Novikov, *On the Algorithmic Unsolvability of the Word Problem in Group Theory*, Trudy Mat. Inst. Skeklov no. 44, Izdat Akad. Nauk. SSSR, Moscow, 1955. (Russian)
18. C. W. Patty, *Foundations of Topology*, PWS-Kent, 1993.
19. M. O. Rabin, Recursive unsolvability of group theoretic problems, *Ann. of Math.* **67** (1958), 172–194.
20. D. J. S. Robinson, *A Course in the Theory of Groups*, Springer-Verlag, New York, NY, 1982.
21. J. J. Rotman, *An Introduction to the Theory of Groups*, fourth edition, Springer-Verlag, New York, NY, 1994.
22. J. Stillwell, *Classical Topology and Combinatorial Group Theory*, Springer-Verlag, New York, NY, 1993.

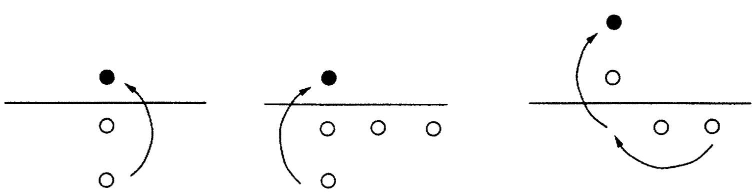
Moving into the Desert with Fibonacci

MARTIN AIGNER
Freie Universität Berlin
Arnimallee 3
14195 Berlin
Germany

The Solitaire Army

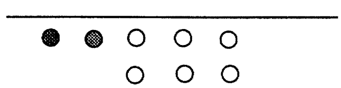
In Volume II of their admirable book *Winning Ways* [1] the authors consider the following solitaire game. Suppose we draw a horizontal line on an unbounded board consisting of squares. On one side of the line stands the (finite) solitaire army, with a square occupied by at most one soldier (peg). On the other side is the desert. A *move* consists of a jump of one peg over another onto a free place, in the horizontal or vertical direction, removing the peg that has been jumped over. How many soldiers do we need to move a scout one, two, three, ..., n steps into the desert?

This rather bland-looking game has a surprising answer. It is easy enough for a scout to venture one or two steps into the desert:



1 step 2 steps

To go 3 steps all we have to do is to reproduce the initial 2-step set-up, one line higher. No problem, the following army will do:



3 steps

The white pegs are used to form the top line, and the two grey pegs yield the additional peg in the bottom line. So we have needed 2, 4, and 8 pegs (and we will see soon that these are the minimal numbers) to move a scout 1, 2, and 3 steps into the desert. The pattern seems clear enough: we need 16 pegs to move up 4 steps, 32 pegs for 5 steps, and so on. To be sure, let us try 4 steps. Again, we want to reproduce the 3-step configuration, one line higher. Here is a possibility:

$$
\begin{array}{ccccccc}
\hline
\circ & \circ & \circ & \circ & \circ & \bullet & \bullet \\
\circ & \circ & \circ & \circ & \circ & & \\
\times & \times & \times & + & + & + & \\
 & & \times & + & & &
\end{array}
$$

The white pegs will bring the top five into line, the grey pegs yield the rightmost peg in the second line, and the $+$-pegs (resp. $\times$-pegs) produce the other two, considering the 2-step set-up from above. Our configuration, however, uses 20 pegs instead of the conjectured 16. Can we do better? No, 20 pegs is the minimal number, as we will see later on. It seems that we have to adjust our conjecture a bit. Quite a bit, in fact: no matter how many pegs we place on the board, the scout will *never* get as far as 5 steps. This surprising theorem is proved in *Winning Ways* by an ingenious argument, and it is our starting point.

THEOREM. *It is not possible to move a peg five steps into the desert.*

The key to the (otherwise simple) proof is a proper weighting of the board. Let σ be the positive root of the quadratic $x^2 + x = 1$, i.e., $\sigma = \dfrac{\sqrt{5}-1}{2} \approx 0.618$; σ is the inverse of the golden section $\tau = \dfrac{1+\sqrt{5}}{2}$. Hence $\sigma^2 + \sigma = 1$ and therefore

$$\sigma^k + \sigma^{k-1} = \sigma^{k-2} \quad \text{for all } k \in \mathbb{Z}. \tag{1}$$

Suppose, on the contrary, that we can move up 5 steps. We now assign weights to the positions of the board as in the following figure; weight 1 is assigned to the place below the final destination of the scout.

$$
\begin{array}{ccccccccc}
 & & & & \sigma^{-5} & & & & \\
 & & \cdots & \sigma^{-3} & \sigma^{-4} & \sigma^{-3} & \cdots & & \\
 & & \cdots & \sigma^{-2} & \sigma^{-3} & \sigma^{-2} & \cdots & & \\
 & & \cdots & \sigma^{-1} & \sigma^{-2} & \sigma^{-1} & \cdots & & \\
 & \cdots & \sigma & 1 & \sigma^{-1} & 1 & \sigma & \cdots & \\
\hline
\cdots & \sigma^3 & \sigma^2 & \sigma & 1 & \sigma & \sigma^2 & \sigma^3 & \cdots \\
\cdots & \sigma^4 & \sigma^3 & \sigma^2 & \sigma & \sigma^2 & \sigma^3 & \sigma^4 & \cdots \\
 & & \cdots & \sigma^3 & \sigma^2 & \sigma^3 & \cdots & & \\
 & & & \cdots & \sigma^3 & \cdots & & & \\
 & & & & \vdots & & & &
\end{array}
$$

Let W_0 be the sum of the weights (called the *value*) in the initial set-up of the pegs, and $W_1, W_2, \ldots$ the values after the first jump, after the second jump, and so on. Here is the crucial observation, which is directly implied by (1): The value *never increases* during the game:

$$W_0 \geq W_1 \geq W_2 \geq \ldots . \tag{2}$$

Note that only upward jumps or jumps toward the middle preserve the value; in all other cases, the value decreases.

Hence, in order to reach the 5th level we must have

$$W_0 \geq \sigma^{-5}.$$

But now let us compute the sum of *all* weights below the dividing line. We have

$$1 + \sigma + \sigma^2 + \sigma^3 + \cdots = \frac{1}{1-\sigma} = \frac{1}{\sigma^2} = \sigma^{-2}, \tag{3}$$

so

$$\sigma^k + \sigma^{k+1} + \sigma^{k+2} + \cdots = \sigma^{k-2} \quad \text{for all } k \in \mathbb{Z}. \tag{4}$$

Therefore, by (1) and (3), the weight-sum of the first horizontal line is

$$(1 + \sigma + \sigma^2 + \cdots) + (\sigma + \sigma^2 + \sigma^3 + \cdots) = \sigma^{-2} + \sigma^{-1} = \sigma^{-3};$$

hence the total sum is

$$\sigma^{-3} + \sigma^{-2} + \sigma^{-1} + 1 + \sigma + \cdots = \sigma^{-5}.$$

We conclude that *every finite* configuration has value $W_0 < \sigma^{-5}$, and hence that we can never move up 5 steps.

Another Proof. Weighting with powers of σ suffices for the impossibility proof, but it is unwieldy to deduce precise results about the minimal configurations. So let us re-prove the theorem in a finite version. Recursion (1) suggests that we use the Fibonacci numbers F_n for our weighting. Recall the definition of F_n; for convenience we start with $F_0 = 1$:

$$F_0 = 1, \quad F_1 = 1, \quad F_n = F_{n-1} + F_{n-2} \quad (n \geq 2).$$

The first few Fibonacci numbers are thus

n	0	1	2	3	4	5	6	7	8	9	10
F_n	1	1	2	3	5	8	13	21	34	55	89

We set $F_{-1} = 0$.

Now we weight the positions on the board as shown below, where n is large enough so that the initial configuration fits into the triangle.

$$\begin{array}{ccccccccc}
 & & & & F_{n+s} & & & & \\
 & & & & \vdots & & & & \\
 & & \cdots & F_{n+1} & F_{n+2} & F_{n+1} & \cdots & & \\
 & \cdots & F_{n-1} & F_n & F_{n+1} & F_n & F_{n-1} & \cdots & \\
\hline
F_1 & \cdots & F_{n-2} & F_{n-1} & F_n & F_{n-1} & F_{n-2} & \cdots & F_1 \\
 & F_1 & \cdots & F_{n-2} & F_{n-1} & F_{n-2} & \cdots & F_1 & \\
 & & F_1 & \cdots & F_{n-2} & \cdots & F_1 & & \\
 & & & \ddots & \vdots & \iddots & & & \\
 & & & & F_1 & & & &
\end{array}$$

Again, (2) holds: the value never increases during the game. All that remains is to evaluate the sum S_n of the Fibonacci numbers below the line.

First, we have the important relation

$$\sum_{k=1}^{n} F_k = F_{n+2} - 2 \qquad (n \geq 1), \tag{5}$$

which is easily verified by induction. Looking at the central vertical line of the triangle and the two halves left and right, we conclude, using (5), that

$$\begin{aligned}
S_n &= \sum_{k=1}^{n} F_k + 2\sum_{i=1}^{n-1} \sum_{k=1}^{i} F_k = (F_{n+2} - 2) + 2\sum_{i=1}^{n-1} (F_{i+2} - 2) \\
&= (F_{n+2} - 2) + 2\sum_{i=3}^{n+1} F_i - 4(n-1) \\
&= F_{n+2} - 2 + 2(F_{n+3} - 5) - 4(n-1) \quad \text{(the summation starts at } i = 3\text{)} \\
&= F_{n+4} + F_{n+3} - 4n - 8 = F_{n+5} - 4n - 8 \\
&< F_{n+5}.
\end{aligned}$$

We have proved again that 5 steps are impossible.

Minimal armies Let us now return to our previous assertions that 2, 4, 8, and 20 pegs are the minimal numbers to move a scout 1, 2, 3, and 4 steps into the desert. Trivially, we need two pegs to move one step up. Notice that in the Fibonacci triangle below the line there are $2k + 1$ positions with weight F_{n-k}.

The maximal weight-sum of 3 pegs (resp. 7 pegs) is therefore $F_n + 2F_{n-1}$ (resp. $F_n + 3F_{n-1} + 3F_{n-2}$). In the first case we obtain $F_n + 2F_{n-1} = F_{n+1} + F_{n-1} < F_{n+2}$. In the second case,

$$\begin{aligned}
F_n + 3F_{n-1} + 3F_{n-2} &= F_{n+1} + 2F_{n-1} + 3F_{n-2} = F_{n+1} + 2F_n + F_{n-2} \\
&= F_{n+2} + F_n + F_{n-2} < F_{n+2} + F_{n+1} = F_{n+3}.
\end{aligned}$$

Hence we need at least 4 pegs (resp. 8 pegs) to move up 2 (resp. 3) steps, and it is easily verified that the configurations we already know are the only ones up to reflection.

The proof that 20 pegs are minimal for a 4-step move is more involved. Suppose we place at most 18 pegs. Then the maximal value below the line is

$$W_0 = F_n + 3F_{n-1} + 5F_{n-2} + 7F_{n-3} + 2F_{n-4}.$$

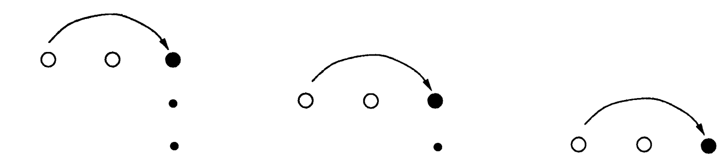
By the Fibonacci recursion it does not matter which $n \geq 5$ we use, so let us set $n = 5$. In this case

$$W_0 = 8 + 3\cdot 5 + 5\cdot 3 + 7\cdot 2 + 2\cdot 1 = 54 < 55 = F_9. \tag{6}$$

So, 18 pegs are not enough. Now suppose 19 pegs would suffice. In this case we have equality in (8): $W_0 = 55 = F_9$. This means, by a previous remark, that *all* jumps must be upward or toward the middle, but never across the middle line. Looking at the Fibonacci weighting we see that the initial set-up must be of the following form (with the weighting already written in):

$$
\begin{array}{ccccccccc}
 & & & & 55 & & & & \\
 & & & & 34 & & & & \\
 & & \cdots & 13 & 21 & 13 & \cdots & & \\
 & & \cdots & 8 & 13 & 8 & \cdots & & \\
\hline
\cdot & 2 & 3 & 5 & 8 & 5 & 3 & 2 & \cdot \\
 & \cdot & 2 & 3 & 5 & 3 & 2 & \cdot & \\
 & & \cdot & 2 & 3 & 2 & \cdot & & \\
 & & & \cdot & 2 & \cdot & & & \\
 & & & & \cdot & & & &
\end{array}
$$

We have 9 possible positions for the three 1's, so there are $\binom{9}{3} = 84$ possible configurations. Even taking symmetry into account, this seems too many to check. But we know that we may never jump across the middle, and this is the key to the proof. We will sketch two steps of the proof, and the reader can easily fill in the rest. First, we show that all moves within the desert must be upward, along the central line. If not, some peg must jump horizontally into the middle line. There are three possibilities:



The first case is immediately disposed of. We already know that it takes 8 pegs to move 3 steps up, so we need 16 pegs to produce two pegs that are 3 places up. To reach the target we would need another peg 2 places up, and hence an additional 4 pegs in our set-up: this gives 20 pegs altogether.

Let us look at the second case. Considering our weighting we have the following situation:

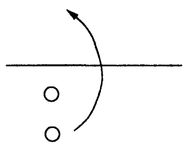
$$
\begin{array}{cccc}
 & & & \cdot \\
 & & & \cdot \\
 & 8 & 13 & \cdot \\
 & \cdot & \cdot & \cdot \\
\hline
2 & 3 & 5 & \cdot \\
 & 2 & 3 & \cdot \\
 & & 2 & \cdot
\end{array}
$$

Since we never jump across the middle line, only the pegs on the left-hand side can be used to create the two pegs with weights 8 and 13. Since

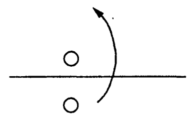
$$8 + 13 = (5 + 2 \cdot 3 + 3 \cdot 2) + 4 \cdot 1,$$

we would need four pegs of weight 1. But there are only three available, so the second case is impossible. The third case is dealt with in a similar fashion.

We have shown that for a peg to move into the desert there are two possibilities:



Type A



Type B

Climbing step by step up to our target position, we therefore need the following sequence of types:

$$A\,(\text{step }1),\ B\,(\text{step }2),\ A\,(\text{step }3),\ A \text{ and } B\,(\text{step }4).$$

This means that (below the line) we must be able to move 5 pegs into the position just below the desert line (for both types) and 3 pegs into the position one lower (type A). This requires 20 pegs, as the reader can easily check.

As an exercise the reader may show that up to reflection there are exactly four 20-peg configurations.

From the Plane to d Dimensions

Now that we have settled the problem in the plane, let us look at 3-space. We consider a 3-dimensional board and draw a dividing *plane*. How many steps can we move into the now 3-dimensional desert? The answer is 7, as we shall see. The stage for the general d-dimensional game is now set. The board is now $\mathbb{Z}^d$, the set of d-dimensional vectors with integral coordinates. The pegs are placed in the half-space $H_{\le 0} = \{(x_1,\dots,x_d) \in \mathbb{Z}^d : x_d \le 0\}$, and we ask for the highest point $(0,0,\dots,0,h_d)$ that can be reached. Hence $h_2 = 4$, and trivially $h_1 = 1$.

PROPOSITION 1. *In* $\mathbb{Z}^d$, $h_d \le 3d-2$ $(d \ge 1)$; *we can never move up* $3d-1$ *steps.*

Proof. We proceed along the lines familiar from the planar case. In $H_{\le 0}$ we assign weight 1 to the origin $(0,0,\dots,0)$, weight σ to $(0,\dots,\pm 1,\dots,0)$ and $(0,0,\dots,0,-1)$, and, in general, weight σ^k to $p=(p_1,\dots,p_d)$, $p_d \le 0$, if p has distance k from the origin, where distance is the so-called *Manhattan distance* $k = \sum_{i=1}^{d} |p_i|$. In the upper half-space $H_{\ge 0}$ we correspondingly choose weight σ^l with $l = \sum_{i=1}^{d-1}|p_i| - p_d$, $p_d \ge 0$. From the definition of the weighting it is again clear that our basic inequalities (2) hold. It remains to prove that the total weight-sum $A(d)$ of all points in $H_{\le 0}$ does not exceed $\sigma^{-(3d-1)}$.

To compute $A(d)$ it is convenient to also consider the related weight-sum $B(d)$ of all points in $\mathbb{Z}^d$, using weight σ^k for all points of distance k. The starting values are, by (4),

$$A(1) = \sigma^{-2} \quad \text{and} \quad B(1) = \sigma^{-3} \tag{7}$$

Classifying the points of $\mathbb{Z}^d$ according to the last coordinate $x_d \leq 0$ (resp. $x_d \geq 1$) gives

$$B(d) = A(d) + \sigma A(d) = \sigma^{-1}A(d). \tag{8}$$

On the other hand, by classifying the points of $H_{\leq 0}$ according to $x_d \leq 0$, we obtain as contributions to $A(d)$

class	contribution
$(\ldots, 0)$	$B(d-1)$
$(\ldots, -1)$	$\sigma B(d-1)$
$(\ldots, -2)$	$\sigma^2 B(d-1)$
$\vdots$	$\vdots$

and thus

$$A(d) = (1 + \sigma + \sigma^2 + \cdots)B(d-1) = \sigma^{-2}B(d-1). \tag{9}$$

From (8) and (9) we infer

$$A(d) = \sigma^{-3}A(d-1),$$

and the starting value (7) implies that

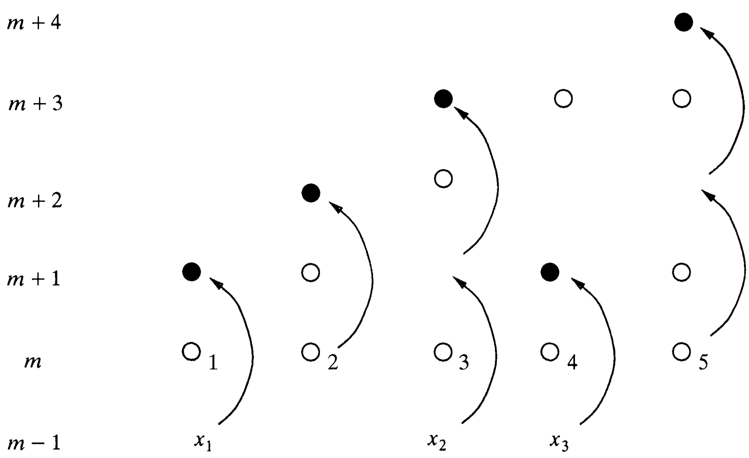
$$A(d) = \sigma^{-3(d-1)-2} = \sigma^{-(3d-1)}.$$

So, as in the plane, the weight-sum of *any* finite configuration remains below $\sigma^{-(3d-1)}$; this is what we wanted to prove. ∎

It is much harder to show that we can indeed move $3d-2$ steps up in dimension d. In fact, finding a successful configuration is beyond the scope of this paper, but here are the main ideas. The key to the proof is the observation we made in the plane: that we can move up 4 steps provided we are able to place in $H_{\leq 0}$ 5 successive pegs into position $(0,0)$ and 3 successive pegs into position $(0,-1)$. The astute reader may have noticed that $5 = F_4$ and $3 = F_3$; this is precisely what we want to generalize. To state the following basic Lemma we slightly generalize our solitaire game. We play the game on the vertical line $\mathbb{Z}$ whose positions are marked by $\ldots, -2, -1, 0, 1, 2, \ldots$. Having a *reservoir* R of size t at position $m \in \mathbb{Z}$ means that we can place a peg out of R into position m whenever we like, but only one at a time until the reservoir is exhausted. Apart from that, we play the game under the normal rules.

LEMMA. *Consider the line $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. Suppose we have a reservoir R_m of size F_k at position m and a reservoir R_{m-1} of size F_{k-1} at position $m-1$. Then we can move a peg up to position $m+k$.*

Proof. The result is trivial for $k = 0$ and also for $k = 1$ (by putting one peg on each of the positions m and $m-1$). By induction, we can move a peg A into position $m+k-1$ by using F_{k-1} pegs from R_m and F_{k-2} pegs from R_{m-1}. After that we use the remaining F_{k-2} pegs from R_m and F_{k-3} pegs from R_{m-1} to place a peg B into position $m+k-2$ (again by induction). Peg B jumping over A now reaches the desired position $m+k$. ∎



As an example consider again $k = 4$, $F_4 = 5$, $F_3 = 3$.

PROPOSITION 2. *Consider the board $\mathbb{Z}^d$, $d \geq 2$. Suppose we can successively move F_{3d-3} pegs into the origin of the board $\mathbb{Z}^{d-1}$ (where we allow any moves, not just within $H_{\leq 0}$). Then $h_d = 3d - 2$.*

Proof. For fixed $m \in \mathbb{Z}$, let $\mathbb{Z}_m^{d-1} = \{x \in \mathbb{Z}^d : x_d = m\}$, thus $\mathbb{Z}_m^{d-1}$ is a copy of $\mathbb{Z}^{d-1}$. By assumption, we can successively move F_{3d-3} pegs within $\mathbb{Z}_m^{d-1}$ into position $(0, \dots, 0, m)$, where, in fact, we just need $m = 0, -1, -2, -3, -4$. In the language of the previous Lemma, this means we have reservoirs R_m of size F_{3d-3} each on the positions $m = 0, -1, \dots, -4$. Now we make repeated use of the Lemma.

Step 1. Use F_{3d-3} pegs from R_0 and F_{3d-4} pegs from R_{-1} to put a peg A into $(0, \dots, 0, 3d-3)$. This exhausts R_0 and leaves F_{3d-5} pegs in R_{-1}.

Step 2. Use these F_{3d-5} pegs from R_{-1} and F_{3d-6} pegs from R_{-2} to move a peg B $3d-5$ steps up from $(0, \dots, 0, -1)$, i.e., into position $(0, \dots, 0, 3d-6)$. This exhausts R_{-1} and leaves $F_{3d-3} - F_{3d-6} = 2F_{3d-5}$ pegs in R_{-2}.

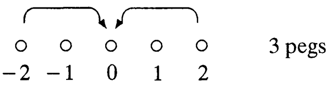
Step 3. Use F_{3d-5} pegs from R_{-2} and F_{3d-6} pegs from R_{-3} to move a peg C into $(0, \dots, 0, 3d-7)$. C can now jump over B and reaches position $(0, \dots, 0, 3d-5)$. This leaves F_{3d-5} pegs in R_{-2} and $F_{3d-3} - F_{3d-6} = 2F_{3d-5}$ pegs in R_{-3}.

Step 4. Use the remaining F_{3d-5} pegs from R_{-2} and F_{3d-6} from R_{-3} to place a peg D into $(0, \dots, 0, 3d-7)$. This exhausts R_{-2} and leaves $2F_{3d-5} - F_{3d-6} \geq F_{3d-5}$ pegs in R_{-3}.

Step 5. Use F_{3d-5} pegs from R_{-3} and F_{3d-6} pegs from R_{-4} to move a peg E into position $(0, \dots, 0, 3d-8)$. E can now jump over D into $(0, \dots, 0, 3d-6)$, then over C into $(0, \dots, 0, 3d-4)$, and finally over A into the desired position $(0, \dots, 0, 3d-2)$. ■

EXAMPLE. Let us see how the algorithm of Proposition 2 works for $d = 2$. On the horizontal line $\mathbb{Z}$ we can move successively 3 pegs into the origin as shown in the figure. One peg is already in the origin, and the other two jump in from left and right. Since $3d - 3 = 3$, we have $F_3 = 3$ as required, and we can now execute the moves of the game according to the Proposition. Note that we need 5 pegs on lines $0, -1, -2, -3$ (since the reservoirs are exhausted), but only 1 peg on line -4

according to step 5 of the Proposition. Our procedure uses therefore 21 pegs—not best possible as we know, but almost.



All that remains is to prove that we can indeed successively move F_{3d} pegs into the origin of $\mathbb{Z}^d$. For $d = 1$ this was demonstrated above, and the reader is invited to settle $d = 2$ ($F_6 = 13$ pegs must be moved into the origin). The rest is completed by induction and involves a lot of "Fibonacci-mathematics."

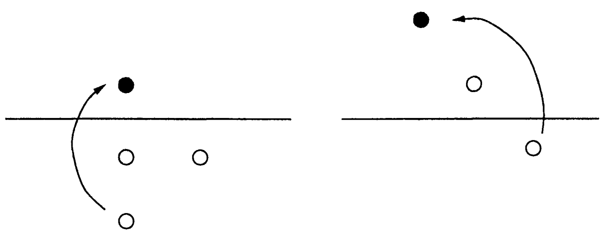
Thus we state our main result:

THEOREM. *We have $h_d = 3d - 2$ for all $d \geq 1$. In other words, on the d-dimensional board $\mathbb{Z}^d$ we can move $3d - 2$ steps into the desert and no further.*

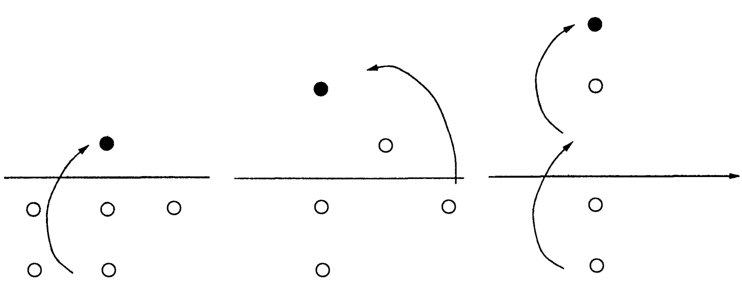
Allowing Diagonal Moves

Now that we have settled the ordinary game, let us go a little beyond. Consider again the plane, and assume we are allowed to move the pegs not only vertically and horizontally, but also along diagonals. How far can we now go into the desert?

Clearly, we need at least 3 pegs to move up two steps. In contrast to the ordinary game, 3 pegs also suffice, as the following figure shows:



To move up 3 steps we must produce two pegs in the desert, one peg two places up and one peg one place up. Hence we need at least 5 pegs to perform 3 steps. Again, this can be done:

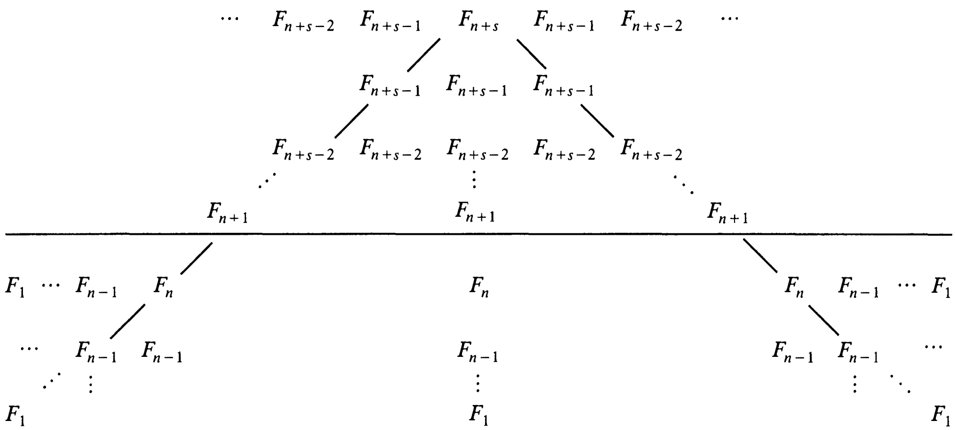


In general, we see by induction that we need at least F_{n+1} pegs for an n-step move. For $n = 4$, $F_5 = 8$ pegs again suffice, and it also works for $n = 5$ steps with F_{n+1} $F_6 = 13$ pegs. In the following figure, the group of pegs numbered i accounts for the move into the i-th position above the desert line. Readers can easily perform the actual solitaire for themselves:

4	4	·	3	1	2	5	5	5
	4	·	3	1	5		5	

For $n = 6$ the going gets heavier, and to move up 7 steps is a formidable task. Where does it end? Does it end at all? Well, let us try our weighting method again.

Suppose we can move up s steps. We use the following "triangular" weighting, adjusted to our diagonal game:



That is, we have three weights F_{n+s-1} below the target position F_{n+s}, then five weights F_{n+s-2}, and so on. Outside the triangle we proceed as in the usual weighting. Thus, below the desert line we have $2s+1$ F_n's, then $2s+3$ F_{n-1}'s, and so on.

Notice that, again, our weighting satisfies the crucial condition (2), that a move can never increase the weight-sum. Hence for a successful procedure we infer

$$W_0 \geq W_1 \geq W_2 \geq \cdots \geq W_t \geq F_{n+s}. \tag{10}$$

Now let us compute the weight-sum of all pegs below the desert line. First we have the $2s+1$ vertical lines starting with F_n, accounting for

$$(2s+1)(F_n + F_{n-1} + \cdots + F_1) = (2s+1) - (F_{n+2} - 2). \tag{11}$$

Next, look at the left-hand side of the triangle below the desert line and observe that there are two sums $F_{n-1} + F_{n-2} + \cdots + F_1$ starting with F_{n-1}, one horizontally and one vertically. Next there are two sums $F_{n-2} + \cdots + F_1$ starting with F_{n-2}, and so on. The remaining weights thus account for

$$\begin{aligned} 4\sum_{i=1}^{n-1}\sum_{k=1}^{i} F_k &= 4\sum_{i=1}^{n-1}(F_{i+2} - 2) = 4(F_{n+3} - 5 - 2(n-1)) \\ &= 4F_{n+3} - 8n - 12. \end{aligned} \tag{12}$$

Taking (12), (13) and (14) together, we conclude that at any rate (disregarding the negative terms)

$$F_{n+s} \leq (2s+1)F_{n+2} + 4F_{n+3}. \tag{13}$$

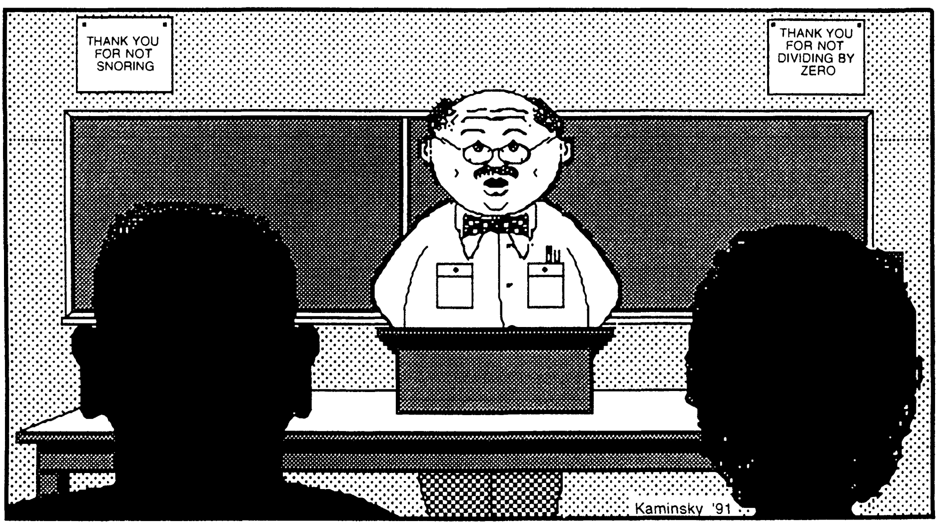
Using the identity $F_k F_n + F_{k-1} F_{n-1} = F_{n+k}$ we easily infer that for $s \geq 9$, inequality (13) is not satisfied anymore, so 9 steps into the desert are not possible. And what about 8 steps? It can be shown that at least 162 pegs are necessary, but this is probably not nearly enough since most horizontal moves (even toward the middle) will decrease the weight-sum. But it *is* possible to move up 8 steps—the reader is challenged to do it. The best solution I know uses 292 pegs, but there should be ample room for improvement. And, of course, there is the d-dimensional variant!

REFERENCE

1. E. Berlekamp, J. Conway, R. Guy: *Winning Ways for Your Mathematical Plays*, Academic Press, London, U.K., 1982.

Professor Fogelfroe

Professor F. Fogelfroe is Professor of Mathematics at ValuPak^{tM} University, in Margo's Forehead, Minnesota.



—Kenneth Kaminsky
Augsburg College
Minneapolis, MN 55454

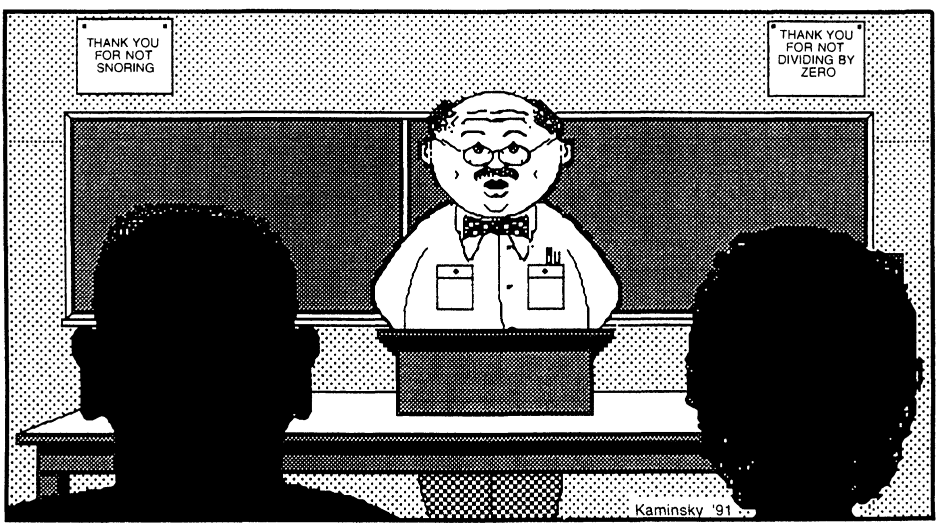
Using the identity $F_k F_n + F_{k-1} F_{n-1} = F_{n+k}$ we easily infer that for $s \geq 9$, inequality (13) is not satisfied anymore, so 9 steps into the desert are not possible. And what about 8 steps? It can be shown that at least 162 pegs are necessary, but this is probably not nearly enough since most horizontal moves (even toward the middle) will decrease the weight-sum. But it *is* possible to move up 8 steps—the reader is challenged to do it. The best solution I know uses 292 pegs, but there should be ample room for improvement. And, of course, there is the d-dimensional variant!

REFERENCE

1. E. Berlekamp, J. Conway, R. Guy: *Winning Ways for Your Mathematical Plays*, Academic Press, London, U.K., 1982.

Professor Fogelfroe

Professor F. Fogelfroe is Professor of Mathematics at ValuPak^{tM} University, in Margo's Forehead, Minnesota.



—KENNETH KAMINSKY
AUGSBURG COLLEGE
MINNEAPOLIS, MN 55454

Golden Fields: A Case for the Heptagon

PETER STEINBACH
Albuquerque Technical-Vocational Institute
Albuquerque, NM 87106

One of the best-kept secrets in plane geometry is the family of ratios of diagonal to side in the regular polygons. So much attention has been given to one member of this family, the *golden ratio* ϕ in the pentagonal case, that the others live in undeserved obscurity. But the wealth of material that pours from the pentagon—proportional sections, recursive sequences, and quasiperiodic systems—can be matched wonder-for-wonder by any other polygon. We will look at some general properties of regular polygons and, in particular, at the case of the heptagon.

I. The Diagonal Product Formula

What are the lengths of the diagonals of a regular polygon with unit side length? As FIGURE 1 shows, the triangle has no diagonals, the square and pentagon each have one kind of diagonal, and the hexagon and heptagon have two. Some diagonal lengths can be discovered by using the Pythagorean theorem or a cosine or sine of a special angle, but to obtain closed forms for ϕ in the pentagon and for ρ and σ in the heptagon we need more. The starred pentagons in FIGURE 2 highlight similar triangles that imply the proportion and the quadratic:

$$\frac{a}{b} = \frac{a+b}{a} = \frac{\phi}{1} \Rightarrow \frac{a}{b} = \frac{a}{a} + \frac{b}{a} \Rightarrow \phi = 1 + \frac{1}{\phi} \Rightarrow \phi^2 - \phi - 1 = 0,$$

which has the positive root $\phi = (1+\sqrt{5})/2 \approx 1.618$.

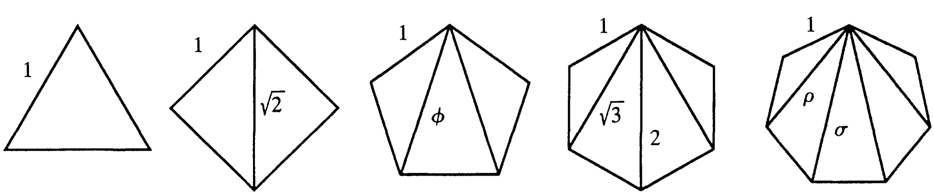


FIGURE 1
Diagonals of regular polygons

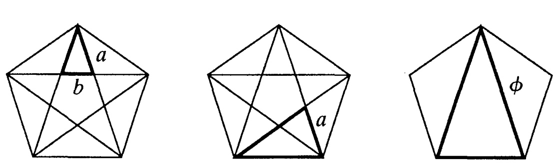


FIGURE 2
Derivation of ϕ

The level of difficulty jumps when we analyze the heptagon. Similar triangles reveal many identities, such as $\rho\sigma = \rho + \sigma$ and $\rho^2 = 1 + \sigma$. These two, solved simultaneously, yield $\rho^3 - \rho^2 - 2\rho + 1 = 0$, and any closed form of the solution involves complex radicals. Rho and sigma are both cubic numbers (roots of irreducible cubic polynomials), and their closed forms are far less useful than the trigonometric forms:

$$\rho = 2\cos(\pi/7) \approx 1.80194, \qquad \sigma = 4\cos^2(\pi/7) - 1 \approx 2.24698.$$

The lengths themselves are less interesting than identities like $\rho\sigma = \rho + \sigma$. Look again at FIGURE 1. The diagonals are arranged in a fan shape to illustrate a pattern: at least up to the hexagon, products of two diagonal lengths are also sums. For the hexagon, we have:

$$\sqrt{3}\cdot 2 = \sqrt{3} + \sqrt{3}, \quad \sqrt{3}\cdot\sqrt{3} = 1 + 2, \quad \text{and} \quad 2\cdot 2 = 1 + 2 + 1.$$

The pattern suggests that the product of two diagonals is a sum of a sequence of diagonals (in the fan, every other one) centered on the longer of the two. This is true also for the heptagon and, indeed, for all regular polygons. We can state it explicitly as follows.

PROPOSITION: DIAGONAL PRODUCT FORMULA. *Consider a regular n-gon* (FIGURE 3) *and let d_0 be the length of a side, d_k the length of a kth internal diagonal, with $k \le \lfloor (n-2)/2 \rfloor$, and $r_k = d_k/d_0$. Then*

$$\begin{aligned}
&r_0 r_k = r_k,\\
&r_1 r_k = r_{k-1} + r_{k+1}, \quad 1 \le k,\\
&r_2 r_k = r_{k-2} + r_k + r_{k+2}, \quad 2 \le k,\\
&r_3 r_k = r_{k-3} + r_{k-1} + r_{k+1} + r_{k+3}, \quad 3 \le k,\\
&\dots\\
&r_h r_k = \sum_{i=0}^{h} r_{k-h+2i} \quad \textit{and} \quad d_h d_k = d_0 \sum_{i=0}^{h} d_{k-h+2i}, \quad h \le k. \qquad (1)
\end{aligned}$$

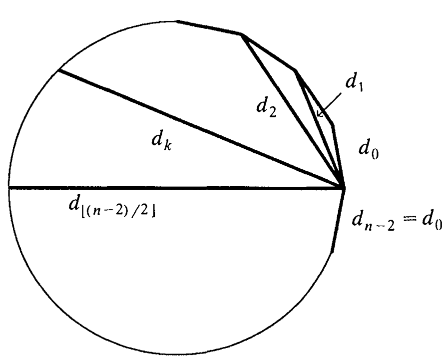


FIGURE 3
Regular n-gon

This can be proved using the n-gon whose vertices are the nth roots of unity in the complex plane. Here is a sketch. Let $\omega = e^{2\pi i/n}$, and express any ratio of diagonal to side as $r_k = |\omega^{k+1} - 1|/|\omega - 1|$. Substitute these ratios in (1) and write out the sum,

$$\frac{|\omega^{h+1} - 1||\omega^{k+1} - 1|}{|\omega - 1|} = |\omega^{k-h+1} - 1| + |\omega^{k-h+3} - 1| + \cdots + |\omega^{k+h+1} - 1|, \quad (2)$$

and manipulate the left side as follows:

$$
\begin{aligned}
&|\omega^h + \omega^{h-1} + \omega^{h-2} + \cdots + 1||\omega^{k+1} - 1| \\
&\quad = |\omega^{k+h+1} - \omega^h + \omega^{k+h} - \omega^{h-1} + \omega^{k+h-1} - \omega^{h-2} + \cdots + \omega^{k+1} - 1| \\
&\quad = |\omega^{k+h+1} - 1 + \omega^{k+h} - \omega + \omega^{k+h-1} - \omega^2 + \cdots + \omega^{k+1} - \omega^h|.
\end{aligned}
$$

The last expression gives a clue for a substitution in the sum on the right side of (2). After substituting appropriately named diagonals, the equation is seen to be an application of the triangle inequality to parallel segments, thus implying equality. Reversing the steps completes the proof.

The diagonal product formula (DPF)(1) allows us to work in the extension field $\mathbb{Q}(r_1)$, wherein we may express products and quotients of diagonals (with $d_0 = 1$) as linear combinations of diagonals. For the pentagon and heptagon the DPF yields the familiar golden ratio identities, $\phi^2 = \phi + 1$ and $1/\phi = \phi - 1$, and the surprising identities:

$$
\begin{array}{lll}
\rho^2 = 1 + \sigma & \sigma/\rho = \sigma - 1 & 1/\sigma = \sigma - \rho \\
\rho\sigma = \rho + \sigma & \rho/\sigma = \rho - 1 & 1/\rho + 1/\sigma = 1 \\
\sigma^2 = 1 + \rho + \sigma & & 1/\rho = 1 + \rho - \sigma
\end{array} \tag{3}
$$

(The quotients are obtained by simple manipulations of the products.)

The polynomials of which diagonal ratios are roots constitute a rich subject for study. The golden polynomial, $x^2 - x - 1 = 0$, has the two roots ϕ and $-1/\phi$. Closer inspection of the cubics for ρ and σ shows that

$$
\begin{array}{ll}
x^3 - x^2 - 2x + 1 = 0 & \text{has roots } \rho, 1/\sigma, -\sigma/\rho; \\
x^3 - 2x^2 - x + 1 = 0 & \text{has roots } \sigma, 1/\rho, -\rho/\sigma.
\end{array} \tag{4}
$$

In absolute value all six heptagonal ratios occur as roots in the two cubics. (I know of no general pattern of roots in such polynomials.)

Using a chain of substitutions in the DPF, one can derive for the regular n-gon a general polynomial ((4) is one case) that has $r_1 = 2\cos\pi/n$ as a root. For odd n this polynomial is

$$
\begin{aligned}
&\binom{k}{0}x^k - \binom{k-1}{1}x^{k-2} + \binom{k-2}{2}x^{k-4} - \cdots \\
&\quad = \binom{k-1}{0}x^{k-1} - \binom{k-2}{1}x^{k-3} + \binom{k-3}{2}x^{k-5} - \cdots, \text{ where } k = \frac{n-1}{2}.
\end{aligned} \tag{5}
$$

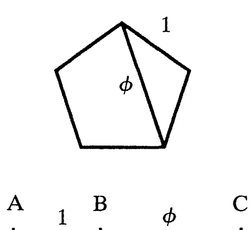
If (5) is written $\boldsymbol{P}_k(x) = 0$, then we have the recurrence $\boldsymbol{P}_{k+1}(x) = x\boldsymbol{P}_k(x) - \boldsymbol{P}_{k-1}(x)$, and $\boldsymbol{P}_k$ can be expressed in terms of derivatives of the Chebychev polynomials. $\boldsymbol{P}_k$ is irreducible over $\mathbb{Q}$ if and only if n is prime. For the 11-gon, for example, r_1 is a quintic number. For the 15-gon, $\boldsymbol{P}_k$ has degree seven, with an irreducible quartic factor that has r_1 as a root. Moreover, when n is prime, the unit side and the set of diagonals form a basis for the field $\mathbb{Q}(r_1)$.

The property of a field that all products and quotients are expressible as linear combinations of basis elements is the feature essential to our study of the behavior of these special numbers. For this reason (and for others to come) I will call the set $\mathbb{Q}(\phi) = \{a\phi + b\colon a, b \in \mathbb{Q}\}$ the *first golden field*. The *second golden field*, $\mathbb{Q}(\rho) = \{a\sigma + b\rho + c\colon a, b, c \in \mathbb{Q}\}$, we will study in more detail.

II. A Family of Golden Proportions

Since ρ and σ are cubic numbers, the heptagon is not classically constructible, which may explain the ancients' silence on the matter. The Greek geometers' method of investigation was construction. This and their limited understanding of irrational numbers would inhibit their analysis of figures like the heptagon. Archimedes **[1]** at least constructed the heptagon with a *marked* straightedge, and may have discovered more. The derivation of ϕ and its properties by similar triangles (FIGURE 2) has been known since ancient times, and one would think that the Greeks would have applied the same reasoning to other figures despite their inconstructibility. (Perhaps they did. Dijksterhuis **[1]** cites evidence of a lost Archimedean manuscript entitled *On the Heptagon in a Circle*.)

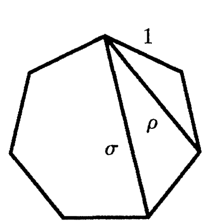
Besides the DPF there is another remarkable pattern that the ancients were able to find but which apparently escaped their notice. The Greeks defined ϕ in terms of a section (cut) of a segment (see FIGURE 4a): *The whole is to the larger part as the larger is to the smaller part*. (Sections such as this began a design tradition, continuing today, in which a harmonious arrangement of elements is defined as one that realizes some special ratio, such as ϕ, and repeats it in proportion. One of the diagonals of the octagon, $\theta = 1 + \sqrt{2}$, is known to architects as the Sacred Cut. See **[4]** and **[7]**.) Numerically, though, the golden *proportion* (realized in the *section*) is the unique solution to the problem of forming a non-trivial proportion, which requires four entries, using only two quantities. When stated as $(a+b)/a = a/b$, the unique solution is $a/b = \phi/1$.



$$\frac{AC}{BC} = \frac{BC}{AB} = \phi \quad \text{or} \quad \frac{\text{whole}}{\text{larger}} = \frac{\text{larger}}{\text{smaller}}$$

FIGURE 4a
Optimal (golden) bisection and proportion

To express ρ and σ in terms of a section (FIGURE 4b), draw a segment of length $1 + \sigma + \rho$ $(= \sigma^2)$. Now the three quantities 1, ρ, σ, in all their combinations as the



$$\begin{aligned} &AB : CD : BC \\ &= CD : AC : BD \\ &= BC : BD : AD \end{aligned} \quad \text{or} \quad \begin{aligned} &1 : \rho \quad : \sigma \\ &= \rho : 1 + \sigma : \rho + \sigma \\ &= \sigma : \rho + \sigma : 1 + \rho + \sigma \end{aligned}$$

FIGURE 4b
Optimal trisection and proportion

six internal segments, fill the 18 entries in these three triple proportions:

$$\frac{\mathrm{AD}}{\mathrm{BC}}=\frac{\mathrm{BD}}{\mathrm{CD}}=\frac{\mathrm{BC}}{\mathrm{AB}}=\sigma \quad \text{or} \quad \frac{\text{whole}}{\text{large}}=\frac{\text{large}+\text{medium}}{\text{medium}}=\frac{\text{large}}{\text{small}};$$

$$\frac{\mathrm{BD}}{\mathrm{BC}}=\frac{\mathrm{AC}}{\mathrm{CD}}=\frac{\mathrm{CD}}{\mathrm{AB}}=\rho \quad \text{or} \quad \frac{\text{large}+\text{medium}}{\text{large}}=\frac{\text{large}+\text{small}}{\text{medium}}=\frac{\text{medium}}{\text{small}};$$

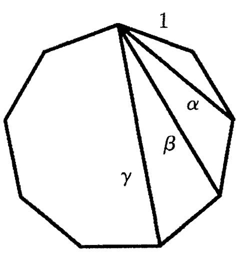
$$\frac{\mathrm{AD}}{\mathrm{BD}}=\frac{\mathrm{BD}}{\mathrm{AC}}=\frac{\mathrm{BC}}{\mathrm{CD}}=\frac{\sigma}{\rho} \quad \text{or} \quad \frac{\text{whole}}{\text{large}+\text{medium}}=\frac{\text{large}+\text{medium}}{\text{large}+\text{small}}=\frac{\text{large}}{\text{medium}}.$$

The enneagon's four diagonals 1, α, β, γ (FIGURE 4c) fill 48 entries in six quadruple proportions.

$$\alpha=\frac{\mathrm{DE}}{\mathrm{AB}}=\frac{\mathrm{AC}}{\mathrm{DE}}=\frac{\mathrm{CE}}{\mathrm{BC}}=\frac{\mathrm{BD}}{\mathrm{CD}} \qquad \frac{\beta}{\alpha}=\frac{\mathrm{BC}}{\mathrm{DE}}=\frac{\mathrm{CE}}{\mathrm{AC}}=\frac{\mathrm{AD}}{\mathrm{CE}}=\frac{\mathrm{BE}}{\mathrm{BD}}$$

$$\beta=\frac{\mathrm{BC}}{\mathrm{AB}}=\frac{\mathrm{CE}}{\mathrm{DE}}=\frac{\mathrm{AD}}{\mathrm{BC}}=\frac{\mathrm{BE}}{\mathrm{CD}} \qquad \frac{\gamma}{\alpha}=\frac{\mathrm{CD}}{\mathrm{DE}}=\frac{\mathrm{BD}}{\mathrm{AC}}=\frac{\mathrm{BE}}{\mathrm{CE}}=\frac{\mathrm{AE}}{\mathrm{BD}}$$

$$\gamma=\frac{\mathrm{CD}}{\mathrm{AB}}=\frac{\mathrm{BD}}{\mathrm{DE}}=\frac{\mathrm{BE}}{\mathrm{BC}}=\frac{\mathrm{AE}}{\mathrm{CD}} \qquad \frac{\gamma}{\beta}=\frac{\mathrm{CD}}{\mathrm{BC}}=\frac{\mathrm{BD}}{\mathrm{CE}}=\frac{\mathrm{BE}}{\mathrm{AD}}=\frac{\mathrm{AE}}{\mathrm{BE}}$$



$$\begin{array}{l} \mathrm{AB}:\mathrm{DE}:\mathrm{BC}:\mathrm{CD} \\ =\mathrm{DE}:\mathrm{AC}:\mathrm{CE}:\mathrm{BD} \\ =\mathrm{BC}:\mathrm{CE}:\mathrm{AD}:\mathrm{BE} \\ =\mathrm{CD}:\mathrm{BD}:\mathrm{BE}:\mathrm{AE} \end{array} \quad \text{or} \quad \begin{array}{l} 1:\alpha:\beta:\gamma \\ =\alpha:1+\beta:\alpha+\gamma:\beta+\gamma \\ =\beta:\alpha+\gamma:1+\beta+\gamma:\alpha+\beta+\gamma \\ =\gamma:\beta+\gamma:\alpha+\beta+\gamma:1+\alpha+\beta+\gamma \end{array}$$

FIGURE 4c
Optimal quadrisection and proportion

Like the golden section, these are also unique solutions, but to precisely what general problem? To answer this question, first rewrite these proportions as in FIGURE 4, as $k \times k$ proportions, consisting of k equal k-part ratios. It is not difficult to show that unity and the diagonals of any regular $(2k+1)$-gon, $1, r_1, r_2, r_3, \ldots, r_{k-1}$, can be arranged end-to-end in such a way that the composite k-part segment illustrates a $k \times k$ proportion using all $\binom{k+1}{2}$ internal segments—a sort of optimal harvest of proportions. The heptagon's 3×3 and the enneagon's 4×4 proportions are shown in FIGURE 4. In fact, for each odd n-gon, only one arrangement (and its reverse) will accomplish this. Only the sequence $1, r_2, r_4, r_6, \ldots, r_{n-5}=r_3, r_{n-3}=r_1$ can accommodate all linear combinations given by the DPF (because of its sums of every-other diagonal), so this is the only sequence that can represent all products as internal segments. But suppose we ask whether any other sets of real numbers can simultaneously fit a $k \times k$ proportion and a sectioned segment. This would define the general problem and the question of uniqueness of solutions.

Let $x_1 = 1 < x_2 < x_3 < \cdots < x_k$, and observe that

$$\begin{aligned} & 1 : x_2 : x_3 : \cdots : x_k \\ = \; & x_2 : x_2^2 : x_2 x_3 : \cdots : x_2 x_k \\ = \; & x_3 : x_2 x_3 : x_3^2 : \cdots : x_3 x_k \\ = \; & \cdots \\ = \; & x_k : x_2 x_k : x_3 x_k : \cdots : x_k^2. \end{aligned}$$

(Each k-part ratio is equal to the first, since it is the first multiplied through by some x_i.) Now, in order to model a k-sectioned segment, we require that each product in this $k \times k$ proportion be equal to a linear combination of the x_i such that every combination occurs *consecutively* in some fixed permutation of the x_i. This is the general problem we are looking for. For $k = 2$,

$$\begin{aligned} & 1 : x \\ & = x : 1 + x \Rightarrow x^2 - x - 1 = 0 \Rightarrow x = \phi = (1 + \sqrt{5})/2, \end{aligned}$$

but for $k = 3$ there are three permutations depending on which quantity is in the middle of the segment, and so there are three cases. For $1 < x < y$, place each in turn in the middle.

order 1, x, $y \Rightarrow$	order x, 1, $y \Rightarrow$	order 1, y, $x \Rightarrow$
$1 : x \quad : y$	$1 : x \quad : y$	$1 : x \quad : y$
$= x : 1 + x : x + y$	$= x : 1 + x : 1 + y$	$= x : 1 + y : x + y$
$= y : x + y : 1 + x + y.$	$= y : 1 + y : 1 + x + y.$	$= y : x + y : 1 + x + y.$

The first two cases above have no real solutions, and the third (realized in FIGURE 4b) uniquely determines that $x = \rho$ and $y = \sigma$, establishing the unique optimal trisection. Similarly, the case $k = 4$ has twelve permutations yielding sixteen cases; fifteen have no real solutions, and the sixteenth uniquely determines the diagonals of the enneagon! The obvious question is whether *all* odd n-gon proportions are unique. I do not know.

(When $k = 4$, twelve permutations yield sixteen cases because of ambiguous inequalities. For instance, if $1 < x < y < z$, then is $1 + z$ less or greater than $x + y$? Both cases must be checked. In general, k induces $k!/2$ segment permutations and k^{k-2} possible proportion cases. So settling the uniqueness of the 11-gon's diagonals for an optimal pentasection involved checking 125 cases, and the pentasection is the largest that has been verified.)

III. Sequences

In number theory, terms of the well-known Fibonacci sequence $0, 1, 1, 2, 3, 5, 8, \ldots$ occur as coefficients of linear combinations expressing powers of ϕ.

$$\begin{aligned} & \phi = 1\phi + 0; \quad \phi^2 = 1\phi + 1; \quad \phi^3 = 2\phi + 1; \quad \phi^4 = 3\phi + 2; \quad \phi^5 = 5\phi + 3; \\ & \phi^6 = 8\phi + 5; \quad \phi^7 = 13\phi + 8 \ldots \end{aligned} \tag{6}$$

The recurrence relation, $\phi^k = a_k\phi + b_k \Rightarrow \phi^{k+1} = (a_k + b_k)\phi + a_k$, yields new coefficients from old. Also, successive ratios of coefficients tend in the limit to ϕ: $a_k/b_k = a_k/a_{k-1} \to \phi$.

Taking powers of σ yields the following:

$$\begin{aligned}
\sigma &= 1\sigma + 0 + 0 \\
\sigma^2 &= 1\sigma + 1\rho + 1 \\
\sigma^3 &= 3\sigma + 2\rho + 1 \\
\sigma^4 &= 6\sigma + 5\rho + 3 \\
\sigma^5 &= 14\sigma + 11\rho + 6 \\
\sigma^6 &= 31\sigma + 25\rho + 14 \\
\sigma^7 &= 70\sigma + 56\rho + 31\ldots
\end{aligned} \tag{7}$$

The recurrence relation is $\sigma^k = a_k\sigma + b_k\rho + c_k \Rightarrow \sigma^{k+1} = (a_k + b_k + c_k)\sigma + (a_k + b_k)\rho + a_k$. In the limit $a_k : b_k : c_k \to \sigma : \rho : 1$, and $x_k/x_{k-1} \to \sigma$ for $x = a$, b, or c. (The sequences $1, 1, 3, 6, 14, \ldots$ and $0, 1, 2, 5, 11, \ldots$ have been included in **[5]**.)

Another interesting sequence and recurrence holds for powers of ρ, and one may also consider linear expansions in terms of the bases $\{1, \rho, \rho^2\}$ or $\{1, \sigma, \sigma^2\}$ instead of $\{1, \rho, \sigma\}$. Note how the recurrence relation in (7) naturally generalizes that in (6). This recurrence pattern can be extended to all odd n-gons for powers of the longest diagonal. (Terauchi **[6]** calls the coefficients in (7) "higher-order Fibonacci sequences." The name has been applied to many things and is now in use by crystal-diffraction physicists. Physicists also employ the terms silver mean, bronze mean, copper mean, etc., but these, with one exception, are unrelated to the polygonal ratios.)

When the transformation $\langle a, b\rangle \to \langle a + b, a\rangle$ is applied to *any* non-negative numbers a and b (initially above $a = 1$, $b = 0$) and iterated, the ratio $a : b$ approaches $\phi : 1$. Similarly, the transformation $\langle a, b, c\rangle \to \langle a + b + c, a + b, a\rangle$ can be applied to any non-negative numbers (above $a = 1$, $b = c = 0$), and the ratio $a : b : c$ must approach $\sigma : \rho : 1$! For its utter simplicity, this may be the most striking manifestation of these algebraic numbers. Extending the recurrence pattern, the transform $\langle a, b, c, d\rangle \to \langle a + b + c + d, a + b + c, a + b, a\rangle$ drives numbers toward the ratios of the enneagon.

Considering the prevalence of the golden proportion and the Fibonacci sequence in the growth of natural forms, one wonders whether these other ratios and sequences occur in nature. Flower symmetries, for example, occur mainly in 4-, 5-, 6-, 7-, and 8-fold types, and the algebraic numbers associated with all but one of these types—the heptagon—have been widely studied in a great variety of natural contexts. Has the heptagon been overlooked? Can ρ and σ be found in plain view?

IV. Quasiperiodics

Numbers of the golden family generate some very deep dissection puzzles, packing problems, and geometric novelties—too many to discuss here. The most interesting and current, though, are quasiperiodic (beware that the term *quasiperiodic* has no universally accepted definition yet).

The first analyzed example of quasiperiodicity was Fibonacci's own model of rabbit populations. In one generation, every baby becomes an adult and every adult reproduces once. Equivalently, we may rule that the quantities 1 and ϕ be multiplied by ϕ. We convey this rule symbolically by $1 \to \phi$, $\phi \to 1\phi$.

When iterated, this rule generates an infinite sequence on two letters. From the initial word 1 we have:

$$\begin{aligned}&1 \to \phi \to 1\phi \to \phi 1\phi \to 1\phi\phi 1\phi \to \phi 1\phi 1\phi\phi 1\phi \to 1\phi\phi 1\phi\phi 1\phi 1\phi\phi 1\phi \\ &\to \phi 1\phi 1\phi\phi 1\phi 1\phi\phi 1\phi\phi 1\phi 1\phi\phi 1\phi \\ &\to 1\phi\phi 1\phi\phi 1\phi 1\phi\phi 1\phi\phi 1\phi 1\phi\phi 1\phi 1\phi\phi 1\phi\phi 1\phi 1\phi\phi 1\phi \to \cdots\end{aligned} \tag{8}$$

In the limiting infinite sequence the characters ϕ and 1 occur in the ratio $\phi : 1$, and their frequencies in any iteration are consecutive Fibonacci numbers. The sequence is not periodic since it cannot be generated by translations of a fundamental domain (the ratio of characters would then be rational), but it is quasiperiodic (QP) in the sense that it exhibits local-isomorphism and local-symmetry properties to be described shortly.

Using the identities (3) given by the DPF we can generate several QP sequences on the characters 1, ρ, σ. Here is one. Begin by applying the following rule, amounting to multiplication by σ:

$$\begin{aligned}&1 \to \sigma, \quad \rho \to \rho\sigma, \quad \sigma \to 1\sigma\rho. \quad \text{Then} \\ &1 \to \sigma \to 1\sigma\rho \to \sigma 1\sigma\rho\rho\sigma \to 1\sigma\rho\sigma 1\sigma\rho\rho\sigma\rho\sigma 1\sigma\rho \\ &\to \sigma 1\sigma\rho\rho\sigma 1\sigma\rho\sigma 1\sigma\rho\rho\sigma\rho\sigma 1\sigma\rho\rho\sigma 1\sigma\rho\sigma 1\sigma\rho\rho\sigma \to \cdots\end{aligned} \tag{9}$$

The characters will occur in the expected ratios, and their frequencies in an iteration we have already seen as coefficients in (7). The order of characters in the last rule ($\sigma \to 1\sigma\rho$) was chosen so that the characters 1 and ρ will never be adjacent, in order to conform to the section order shown in FIGURE 4. This makes it possible to find a linear combination for any power of ρ or σ as a contiguous subsequence infinitely often in the sequence. Another QP sequence based on multiplication by ρ has this same property, known as local isomorphism, and so of course does the ϕ-sequence (8) in powers of ϕ.

Crossing a QP sequence with itself on the Euclidean plane is a very simple way to produce an aperiodic tiling. FIGURE 5 shows the results of applying either the

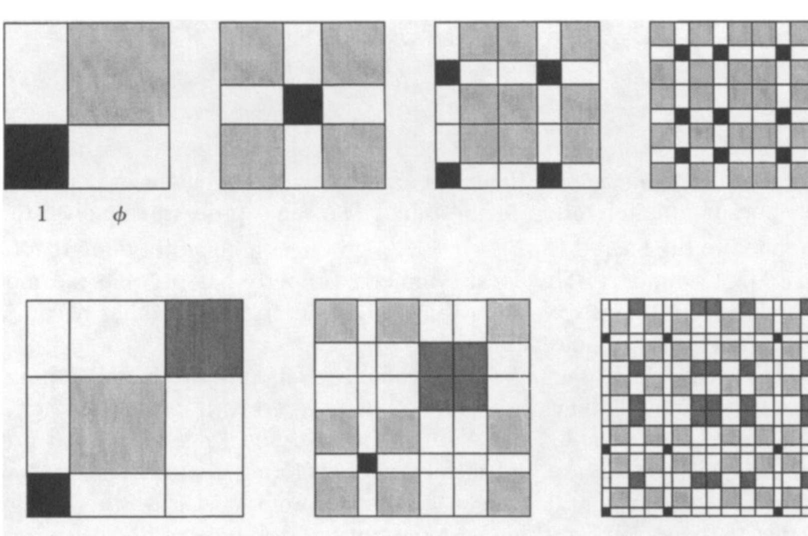
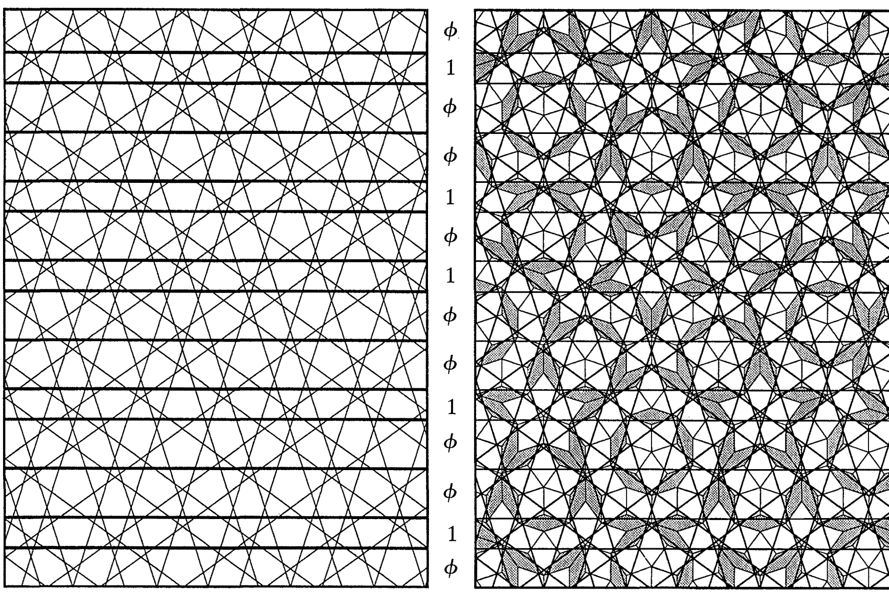


FIGURE 5
ϕ- and σ-aperiodic tilings

ϕ-transform (8) or the σ-transform (9) to the sides of a square. Successive subdivision and magnification covers the plane with tiles aperiodically (though the tilings have one reflective symmetry across the diagonal).

In a very different way the ϕ-sequence can be used to generate Penrose QP tilings. As in FIGURE 6a, form an infinite set of parallel lines, spaced according to the ϕ-sequence. Five sets of these lines, oriented at multiples of $2\pi/5$ and arranged appropriately, form an Ammann 5-grid **[3]**, to which the Penrose rhombus tiling is a sort of dual (FIGURE 6b). If the tiles are decorated with segments from the grid (actually, all varieties of Penrose tiles are mere decorations on the Ammann grid), and tiles are assembled so that segments of the grid form straight lines, then a plane-covering QP tiling is guaranteed.



a b

FIGURE 6
Ammann grid and Penrose tiling

Most interesting is the way that the local isomorphism property of the ϕ-sequence is manifest in the tiling as *local symmetry*: arbitrarily large regions of pentagonal (D_5) symmetry occur infinitely often in the tiling. The same process, sequence to grid to tiling, applied to the Sacred Cut, $\theta = 1 + \sqrt{2}$, produces a beautiful tiling that has local octagonal (D_8) symmetry. This local symmetry property has become the motivation for a search for QP tilings with other dihedral symmetries, and so has become one of the defining elements of quasiperiodicity.

Now if the reader is expecting a heptagonal revelation, there is only a puzzle. The problem of deciding whether the plane can be covered with copies of a given set of tiles is in general unsolvable; it is akin to the so-called "algebraic word problem." What we have just seen is a more fruitful way of exploring tiling: Establish the desired properties by means of a grid, then make tiles to order, that is, dissect the grid into tiles so that the pattern of grid lines is constant on each type of tile. Historically, the grid method has revealed many properties of these tilings. But though there are several ρ, σ-QP sequences analogous to the ϕ-sequence, remarkably, the expected

7-grid has not been found. (The claim was made in **[3]**, but no explanation has appeared.) Analyses to date suggest that there must exist tilings with D_7 local symmetry, but that centers of such symmetry are very far apart, and that if there is a rhombic heptiling analogous to the Penrose sort, then each type of rhombus will appear with more than one type of decoration or matching rule, and the number of types of decorations is unknown. (For an interesting alternative, see Franco's tiling **[2]**. This is a radial tiling; it has one center of global rotational symmetry.)

The difficulty of the grid method in the heptagonal case prompts two observations. First, all of the richly symmetric QP grids found so far—with D_5, D_8, D_{12}, and D_{17} local dihedral symmetry—are based on *constructible* polygons. Do we face the same barrier that stopped the Greek geometers, but for different reasons? Second, the regular pentagon is prohibited from periodic systems of dimension less than four, but can occur in QP systems in $\mathbb{R}^2$ and $\mathbb{R}^3$. The regular heptagon cannot occur in periodic systems of dimension less than six. At what level does the heptagon—using a grid and a finite number of tile shapes—fully participate in the quasiperiodic arena?

Acknowledgment. My study of these patterns began many years ago when I told Abraham Hillman of the University of New Mexico that I wanted to find the Seventh Roots of Unity. "You may find them," he replied, "but they won't be of any use to you." I found them and he was right. (At the time I didn't know that he never speaks hastily.) But the path I took showed me many useful and beautiful things.

REFERENCES

1. E. J. Dijksterhuis, *Archimedes*, Ejnar Munksgaard, Copenhagen, 1956.
2. B. J. O. Franco, Third-order Fibonacci sequence associated to a heptagonal quasiperiodic tiling of the plane. *Physics Letters A* 178 (1993) 119–122.
3. B. Grünbaum and G. C. Shephard, *Tilings and Patterns*, Freeman and Co., New York, NY, 1987.
4. Robert Lawlor, *Sacred Geometry*, Thames & Hudson, London, 1992.
5. Neil J. Sloane and Simon Plouffe, *The Encyclopedia of Integer Sequences*. Academic Press, New York, NY, 1995.
6. H. Terauchi, et al., Synchrotron X-ray diffraction study of third-order Fibonacci lattices. *J. Phys. Soc. Japan* 59 (1990) 405.
7. Kim Williams, The sacred cut revisited: the pavement of the Baptistery of San Giovanni, Florence. *The Mathematical Intelligencer* 16:2 (1994) 18.

Where the Inflection Points of a Polynomial May Lie

BRUCE B. ANDERSON
Kent State University
East Liverpool, OH 43920

1. Introduction

Suppose you have an nth degree polynomial $p(x)$ with n distinct (multiplicity 1) real roots, $r_1 < r_2 < \cdots < r_n$. Rolle's theorem guarantees that for each i, $i = 1 \ldots n-1$, there is a critical point c_i, with $r_i < c_i < r_{i+1}$. Since all $n-1$ critical points are accounted for, the critical points will also be distinct, and in fact will be alternating maximum and minimum points. If $n > 2$, then between each pair of critical points will be a second-order critical point, d_i such that $c_i < d_i < c_{i+1}$. These second-order critical points, also distinct, are the inflection points. Thus every interval defined by a pair of neighboring critical points contains a root of the polynomial r_i as well as an inflection point d_{i-1}. But are there any rules as to how the set of d_i's can fall relative to the original roots r_i? Each inflection point d_{i-1} can be larger than, equal to, or less than the corresponding root r_i.

The situation is depicted in FIGURE 1. The 0's refer to roots of the polynomial, 1's are the critical points, and 2's are the inflection points, all located along the x-axis.

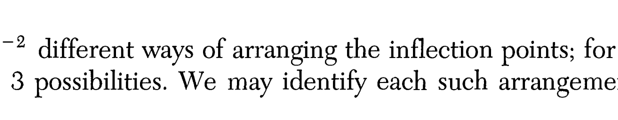


FIGURE 1
A particular arrangement of inflection points

There are 3^{n-2} different ways of arranging the inflection points; for each inflection point there are 3 possibilities. We may identify each such arrangement by defining

$$\sigma_i := \begin{cases} -1 & \text{if} \quad r_i < d_{i-1} \\ 0 & \text{if} \quad r_i = d_{i-1} \\ 1 & \text{if} \quad r_i > d_{i-1} \end{cases}$$

and then specifying a sequence $\sigma = (\sigma_2, \ldots, \sigma_{n-1})$. (Alternatively, one could identify the arrangement by a different sign sequence, given by $\sigma_i := \operatorname{sign}(f''(r_i))$, since the sign of p'' will change across each inflection point.)

Suppose you want to form a particular ordering of roots and inflection points—or equivalently, a particular sign sequence $\sigma = (\sigma_2, \ldots, \sigma_{n-1})$. For example, in creating a calculus exam you might want a function with 5 roots and sign sequence $\sigma = (-1, -1, 1)$. Its graph would look something like FIGURE 2. Is there a polynomial with these characteristics?

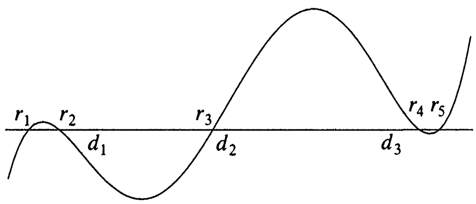


FIGURE 2
A custom-made polynomial

In general, we wish to determine whether any arrangement of inflection points, or, equivalently, any sign sequence $\sigma = (\sigma_2, \dots, \sigma_{n-1})$ is realizable by some nth degree polynomial. We will show that all of these 3^{n-2} arrangements are in fact realizable by polynomials. For 2^{n-2} of the 3^{n-2} arrangements—those where none of the inflection points are equal to the corresponding roots—we shall even provide an exact construction.

While any arrangement of inflection points is realizable by a polynomial, the picture is quite different if we also incorporate third-order critical points. In this case, some arrangements are not realizable by a polynomial **[1]**.

We start with the special case where *every* inflection point equals its corresponding root, as it is used to construct the other arrangements.

2. Inflection Points Coinciding with Roots

We will show first that for any integer $n \geq 3$, there exists an nth degree polynomial $p(x)$ with n distinct real roots, such that the inflection points equal the interior roots of p (all the roots except the first and last). The situation is depicted in FIGURE 3. Indeed, existence of one such polynomial implies that an infinite number exist; translating all the roots an equal amount or multiplying the polynomial by a constant will not affect the position of the roots relative to any critical or inflection points.

$$
\begin{array}{ccccccccccccccccc}
0 & & 0 & & 0 & & 0 & & \cdots & & 0 & & 0 & & 0 & & 0 \\
 & 1 & \Big| & 1 & \Big| & 1 & \Big| & 1 & & 1 & \Big| & 1 & \Big| & 1 & \Big| & 1 & \\
 & & 2 & & 2 & & 2 & & & & 2 & & 2 & & 2 & &
\end{array}
$$

FIGURE 3
Coinciding inflection points

If we restrict attention to monic polynomials with first and last roots at $x = \pm 1$, given by

$$p(x) = (x^2 - 1)\left(x^{n-2} + a_{n-3}x^{n-3} + \cdots + a_1 x + a_0\right),$$

then we will show there is a *unique* such polynomial for which the inflection points coincide with the interior roots. The inflection points are the roots of

$$p''(x) = n(n-1)x^{n-2} + (n-1)(n-2)a_{n-3}x^{n-3} + (n-2)(n-3)(a_{n-4} - 1)x^{n-4} \\ + (n-3)(n-4)(a_{n-5} - a_{n-3})x^{n-5} + \cdots + 2(a_0 - a_2).$$

We want the roots of $q(x) = x^{n-2} + a_{n-3}x^{n-3} + \cdots + a_1 x + a_0$ to be identical to those of $p''(x)$. Two polynomials can have the same set of zeros (real and complex) with the same multiplicity only if they are the same up to a constant multiple. Thus we want $p''(x) = n(n-1)q(x)$. Equating coefficients, we get a series of equations:

$$\begin{aligned} n(n-1)a_{n-3} &= (n-1)(n-2)a_{n-3}, \\ n(n-1)a_{n-4} &= (n-2)(n-3)(a_{n-4}-1), \\ n(n-1)a_{n-6} &= (n-4)(n-5)(a_{n-6}-a_{n-4}), \\ &\cdots \end{aligned}$$

This quickly yields the descending recursive formulae:

$$\begin{aligned} a_{n-3} &= a_{n-5} = a_{n-7} = \cdots = 0 \\ a_{n-2} &= 1 \\ a_{n-2k} &= \frac{a_{n-2k+2}(n-2k+2)(n-2k+1)}{(n-2k+2)(n-2k+1)-n(n-1)}, \end{aligned}$$

where $2 < 2k \leq n$. Note that the denominator will never vanish, since the first product is always less than the second.

Examples of these polynomials up to degree 6 are given in the following table. Their graphs are depicted in FIGURE 4.

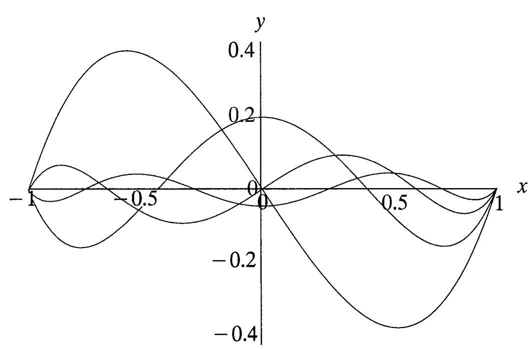


FIGURE 4

degree	polynomial
3	$x(x^2-1) = x^3 - x$
4	$(x^2-1)(x^2-\frac{1}{5}) = x^4 - \frac{6}{5}x^2 + \frac{1}{5}$
5	$x(x^2-1)(x^2-\frac{3}{7}) = x^5 - \frac{10}{7}x^3 + \frac{3}{7}x$
6	$(x^2-1)(x^4-\frac{2}{3}x^2+\frac{1}{21}) = x^6 - \frac{5}{3}x^4 + \frac{5}{7}x^2 - \frac{1}{21}$

In short, for any $n \geq 3$ there is a unique monic polynomial of degree n with two of its roots at $x = \pm 1$ and whose second-order critical points match the remaining roots (with the same multiplicity). Several questions remain in order to meet all our criteria: (1) Are the roots distinct? (2) Are they real? (3) Do they lie inside the interval $[-1, 1]$?

We will answer each of these questions in the affirmative.

(1) *All the roots are distinct.* Suppose not. The roots of $p(x)$ are *not* distinct if and only if there is a root x_0 with multiplicity at least two. Thus we can write

$$p = (x - x_0)^m q$$

where $m \geq 2$ and q is a polynomial with $q(x_0) \neq 0$. Differentiating twice yields

$$p'' = (x - x_0)^{m-2}\left[m(m-1)q + 2m(x - x_0)q' + (x - x_0)^2 q''\right].$$

Since $q(x_0) \neq 0$, p'' has multiplicity $m - 2$ at x_0. If $x_0 \neq \pm 1$ this contradicts the fact that p and p'' were constructed to have zeros of the same multiplicity at x_0. Now suppose that $x_0 = \pm 1$. In this case, by construction, p must have a zero of multiplicity one higher than p'', since p and p'' are essentially the same, except that p has additional roots at ± 1. But by the argument just given, zeros of p can only have multiplicity 2 higher than zeros of p''.

(2) *All the roots are real.* To show this we resort to a result from complex analysis. Lucas' Theorem (see, e.g., **[3]** and **[4]**) tells us that all the critical points of p must lie in the convex hull H of the roots of p. Further, if the zeros of p are not collinear, no critical point of p lies on the boundary of H unless it is a multiple root of p. In our case, p has no multiple roots. If p has complex (but not real) roots, they could not all be collinear, since p has roots at ± 1. Thus Lucas' Theorem guarantees that the critical points of p lie in the interior of H. Applying Lucas' Theorem again, we see that the second-order critical points must lie in the interior of H', the convex hull of the critical points. (See FIGURE 5.) In particular, the second-order critical points must all lie in the interior of H. Thus the roots of p cannot possibly match the roots of p'', as assumed. Therefore all roots of p are real.

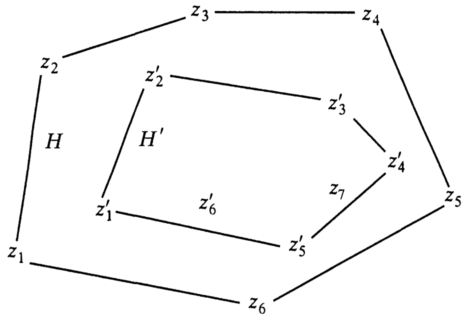


FIGURE 5
Lucas' Theorem, applied twice

(3) *All the roots lie in the interval* $[-1, 1]$. Since p has all real distinct roots, Rolle's theorem implies that every critical point lies in the interior of each pair of neighboring roots. This implies the first and last roots of the polynomial must fall, respectively, before and after the first and last inflection points. Since the only two roots not matched with an inflection point are at $x = \pm 1$, the other roots must lie in $(-1, 1)$.

3. Adjusting the Roots

The next step is to adjust the roots so that a specified subcollection of the roots falls to the right of the inflection points, while the others fall to the left. For this, we employ a slight extension of the Polynomial Root Dragging Theorem **[1]**, **[2]**, **[4]**:

THEOREM. *Let $p(x)$ be a polynomial of degree n with distinct real roots $r_1, \ldots, r_n$. Then as we "drag" some or all of the interior roots a distance at most ε to the right, the critical points will all follow to the right, and each of them will move less than ε units.*

Proof. If c is a critical point of $p(x) = \Pi(x - r_i)$, then $p'(c) = 0$. Since the roots of p are distinct, $p(c) \neq 0$. Thus

$$-\frac{p'(c)}{p(c)} = \frac{1}{r_1 - c} + \frac{1}{r_2 - c} + \cdots + \frac{1}{r_n - c} = 0.$$

Implicit differentiation with respect to r_i gives

$$\frac{\partial c}{\partial r_i} = \frac{\left(\frac{1}{r_i - c}\right)^2}{\left(\frac{1}{r_1 - c}\right)^2 + \left(\frac{1}{r_2 - c}\right)^2 + \cdots + \left(\frac{1}{r_n - c}\right)^2}.$$

It follows that $0 < \Sigma(\partial c/\partial r_i) < 1$, where we sum over any strict subset of $\{1 \dots n\}$. Thinking of c as a function of the roots, consider what happens to c as we shift every root r_i (other than the first and last) by an amount $\varepsilon_i \geq 0$. Consider the function:

$$f(t) := c(r_1, r_2 + t\varepsilon_2, \dots, r_{n-1} + t\varepsilon_{n-1}, r_n),$$

where t varies from 0 to 1. Then

$$f'(t) = \sum_2^{n-1} \frac{\partial c}{\partial r_i} \varepsilon_i \leq \max(\varepsilon_i) \sum \frac{\partial c}{\partial r_i} < \max(\varepsilon_i).$$

By the mean value theorem, $f(1) - f(0) < \max(\varepsilon_i)$, so c will move to the right strictly less than the fastest moving root. Iterating the argument, we find that the inflection points also move slower than the fastest moving root. This completes the proof.

Using this information we can adjust the roots to suit our purposes. Start with the polynomial whose roots coincide with its inflection points. Select those roots which we want to fall to the right of the corresponding inflection points. Drag those roots to the right by a small distance ε, keeping the other roots frozen. As long as ε is less than the minimum separation of the roots of p, there will be no crossover of roots. Since all the inflection points follow behind the moving roots, we arrive at a polynomial whose selected roots lie to the right of the inflection points, while the frozen roots remain to the left of their corresponding inflection points, as illustrated in FIGURE 6.

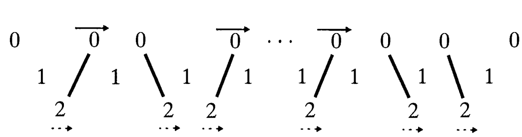


FIGURE 6
Adjusting the roots

4. Construction of a Particular Example

Let us apply the theory to the particular example depicted in FIGURE 2; the graph shows a polynomial with five real roots. We begin with the polynomial $x^5 - \frac{10}{7}x^3 + \frac{3}{7}x$ from the table in Section 2. The polynomial has five real roots; the inner three are aligned with the inflection points. The outer roots are at ± 1, and the inner roots are at $0, \pm \sqrt{21}/7 \approx \pm 0.655$.

We want the first two roots to remain to the left of the corresponding inflection points, so we keep them frozen. The third root needs to shift to the right. Shifting the third root by an arbitrary amount, say $\frac{1}{7}$, produces a new third root at $(\sqrt{21}+1)/7 \approx 0.800$. (As long as we do not shift the root beyond 1, any shift is acceptable.) The new polynomial has the form

$$(x+1)\left(x+\frac{\sqrt{21}}{7}\right)x\left(x-\frac{\sqrt{21}+1}{7}\right)(x-1).$$

The inflection points are located at approximately -0.650, 0.031, and 0.7048—just where we wanted relative to the roots.

5. Keeping Some Roots Lined Up

We are ready to consider the more difficult case: What if we want some of the inflection points to coincide with the roots, while the other inflection points are specified in a certain order relative to the roots, some left and some right? We will show that all such arrangements exist, though we do not provide an exact construction.

Let us begin again with the polynomial $p(x)$ constructed in Section 2, whose inflection points all coincide with the roots. Define "rightward" (resp. "leftward") roots to be to the right (resp. left) of the corresponding inflection points; "neutral" roots are those we want lined up with their inflection points. As in Section 3, begin by "dragging" all the rightward roots an equal distance ε to the right; again, ε must be less than the minimum separation of the roots. The leftward and neutral roots, which do not move, will now be to the left of their inflection points.

We want to drag the neutral roots to the right in order to line them up with their inflection points. As long as none of these roots is dragged more than a total of ε units to the right from its original position, none of the rightward roots will be "passed" by its corresponding inflection points.

We line up the neutral roots (call them $s_1, s_2, \ldots, s_k$) as follows: Start with s_1, which is still in its original position, to the left of its inflection point. By the time we shift s_1 to the right ε units, s_1 will lie to the right of its inflection point. Somewhere in between, therefore, s_1 will coincide with its inflection point. Now begin dragging s_2, which is currently to the left of its inflection point. By the time it is shifted ε units, s_2 will be to the right of its inflection point. Somewhere in between, s_2 coincides with its inflection point. In the process, s_1 is back to the left of its inflection point, since its inflection point was dragged to the right. We continue lining up $s_3, s_4, \ldots, s_k$ and then begin cycling through again, lining up s_1, then lining up s_2, etc. This process is illustrated in FIGURES 7 and 8. Each of these neutral roots and its corresponding inflection point defines an increasing sequence, bounded above by the original value plus ε. Since every increasing bounded sequence has a limit point, each of these neutral roots and inflection points must approach a limit. Thus the difference between

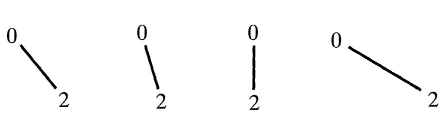


FIGURE 7

The roots shown are only neutral roots. The first two roots were lined up previously by dragging them to the right. The third root was then lined up, so that the first and second are no longer lined up. The fourth has not yet been lined up.

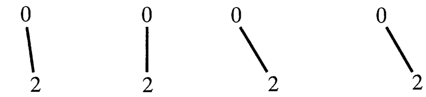


FIGURE 8

After cycling through the four neutral roots, lining each one up consecutively, we go back to the first root and line it up. Now the second root has just been lined up a second time.

the root and the inflection point must also approach a limit, and that limit will be zero, since the difference must equal zero once every cycle. Therefore in the limit the neutral roots will exactly line up with the corresponding inflection points.

Appendix

The graphs of the special polynomials derived in Section 2 merit further study. FIGURES 9, 10, and 11 illustrate the graphs for some even degree polynomials. They

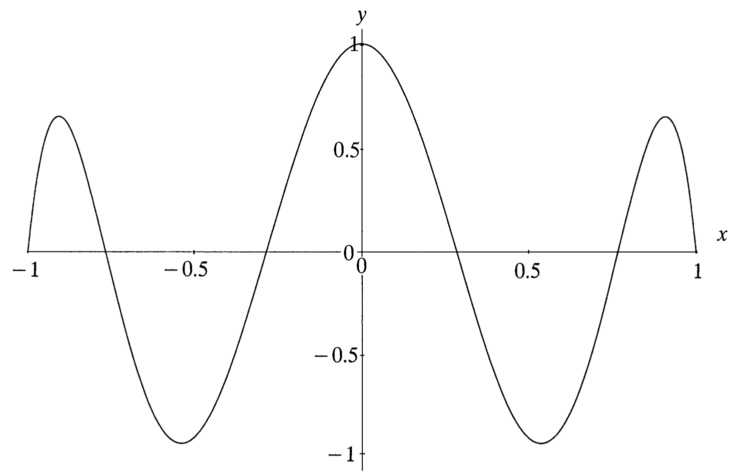


FIGURE 9
$n = 6$

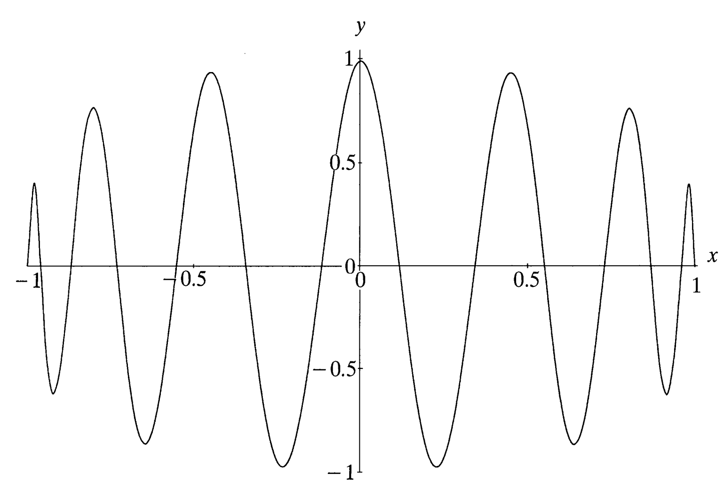


FIGURE 10
$n = 14$

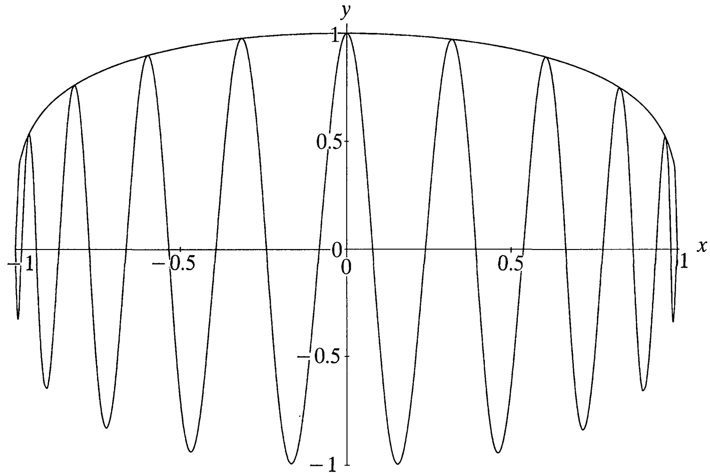


FIGURE 11
$n = 20$ and an asymptotic candidate: $\sqrt[4]{1-x^2}$

have been normalized so that the top (or bottom) peak has magnitude 1. Note (as pointed out by a referee) that as n increases, the "envelope" appears to have a nice asymptotic limit. The author wagers that the asymptotic limit is given by $\sqrt[4]{1-x^2}$, as shown in FIGURE 11. It is left as a challenge to the reader to prove or disprove this, and to find an asymptotic limit for odd degree polynomials as well.

REFERENCES

1. B. Anderson, Polynomial root dragging, *Amer. Math. Monthly*, 100 (1993) 864–866.
2. P. Andrews, Where not to find the critical points of a polynomial—variation on a Putnam theme, *Amer. Math. Monthly*, 102 (1995) 155–158.
3. M. Marden, *Geometry of Polynomials*, American Mathematical Society, Providence, 1966.
4. T. Needham, A visual explanation of Jensen's inequality, *Amer. Math. Monthly*, 100 (1993) 768–771.
5. P. Walker, Bounds for the separation of real zeros of polynomials, *Australian Mathematical Society Journal Series A*, 59 (1995) 330–342.

NOTES

How to Spread Rumors Fast

C. KENNETH FAN
Harvard University
Cambridge, MA 02138

BJORN POONEN
Princeton University
Princeton, NJ 08540

GEORGE POONEN
Renaissance Solutions Inc.
55 Old Bedford Road
Lincoln, MA 01773

The problem Seven gossiping friends are most anxious to share their gossip amongst each other. Each gossiper knows something unique to begin with. Each day, some of the gossipers phone each other and exchange all the gossip they have collected so far, but on any given day, each gossiper can be involved in a phone call with at most one other person. (No conference calls allowed!)

What is the minimum number of days it would take for each gossiper to learn the collective gossip of all the gossipers?

Let us denote by $d(n)$ the minimal number of days required for n gossipers to share all their knowledge. The first few values of $d(n)$ are as follows:

n	1	2	3	4	5	6	7	8
$d(n)$	0	1	3	2	4	3	4	3

Note that $d(n)$ is not an increasing function: for instance, it takes more time for three people to share their gossip than it does for four people! We shall prove

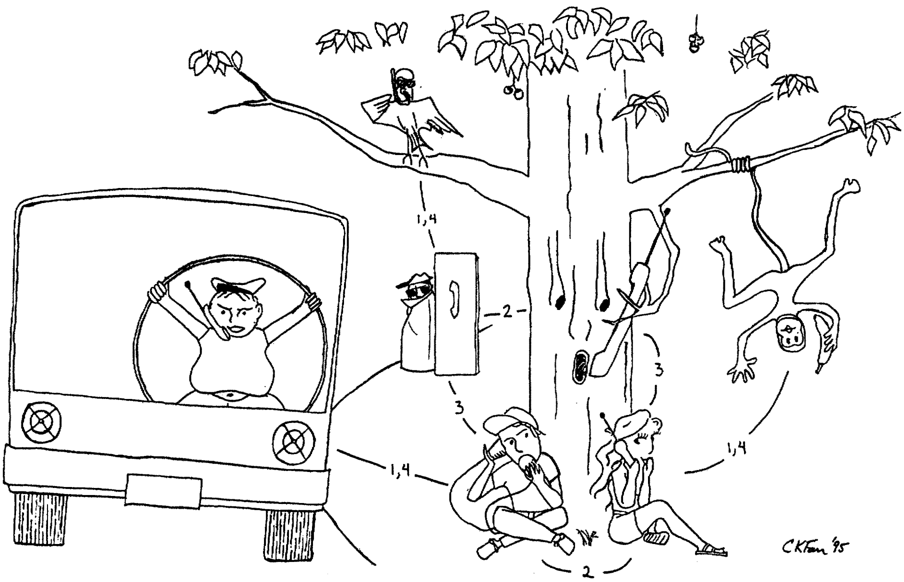
$$d(n) = \begin{cases} \lceil \log_2 n \rceil, & \text{if } n \text{ is even}, \\ \lceil \log_2 n \rceil + 1, & \text{if } n \text{ is odd}, \end{cases}$$

and indicate how the gossipers should behave to achieve this minimum.

A lower bound On any given day, the number of people who can know a given fact, say, fact A, can at most double since each person who knows fact A can tell it to at most one other person. Therefore, we cannot hope to spread all the gossip around in fewer than $\lceil \log_2 n \rceil$ days, where $\lceil x \rceil$ is the least integer greater than or equal to x.

Hypercubes Consider an even number of gossipers $2m$. On the first day, we can divide the gossipers into two groups, group I and group II, each of size m. By definition, in $d(m)$ days, each member of each group can learn all the information within his or her own group. We take $d(m)$ days to accomplish this, working the two groups in parallel. Next, we pair up the gossipers in group I arbitrarily with the gossipers in group II so that after another day, all the gossipers will know all the gossip. In this way, we see that $d(2m) \leq 1 + d(m)$.

When $n = 2^D$, this inductively leads to the inequality $d(2^D) \leq D$. On the other hand, we have already seen that $d(2^D) \geq D$. Consequently, $d(2^D) = D$. In this case, the gossipers can be identified with vertices of a D-dimensional hypercube so that the phone calls made on any given day correspond to the set of edges parallel to some edge.



An odd number of gossipers Now assume n is odd and greater than 1. Here is one method of spreading the rumors. Let h be the positive integer with $2^h < n < 2^{h+1}$. Let H be a subset of 2^h gossipers, and let H^c be the complement of H. Note that $|H| > |H^c|$. Because of this, on the first day we can pair up the members of H^c with distinct members of H. (Some members of H will not talk to anyone on the first day.) After the first day, the members of H will collectively know everything. We can use the results of the previous section and, ignoring temporarily the members of H^c, have the members of H take h days to exchange their knowledge amongst each other. When that is finished, we again pair up members of H^c with the now omniscient members of H, thus taking another day to complete the spread of rumors. See the illustration for the case $n = 7$ (lines with numbers indicate the day on which the connected couple should communicate). This process takes $h + 2 = \lceil \log_2 n \rceil + 1$ days.

For n odd, it turns out that this is the best one can do. To see why, observe that on the first day of conversations, at least one gossiper will be forced to sit out from communicating. This gossiper knows something unknown to the others (by our hypothesis), and the lower bound argument given earlier shows that this bit of information takes at least $\lceil \log_2 n \rceil + 1$ days to spread around because on the second day, still only one person knows the bit of information. We conclude that for n odd, $d(n) = \lceil \log_2 n \rceil + 1$.

An even number of gossipers The method used to spread rumors for odd n applies in general. Thus, we see that

$$\lceil \log_2 n \rceil \leq d(n) \leq \lceil \log_2 n \rceil + 1.$$

When n is odd, we have equality on the right, and when $n = 2^d$, we have equality on the left. We now discuss the remaining case, with n even but not a power of 2. (Actually, the method we describe below also works for $n = 2^d$, and in that case gives a new pattern of phone calls that distributes the information just as fast as the hypercube arrangement.)

For $k \geq 1$, define

$$\Delta_k = \frac{1-(-2)^k}{3}.$$

Note that Δ_k is an odd integer. Label the gossipers 1 through n once and for all. By gossiper p (p any integer), we shall be referring to the gossiper labeled q where $1 \leq q \leq n$ and q and p are congruent modulo n. Since n is even, it makes sense to refer to even and odd gossipers. On day k, make the odd gossiper $2x-1$ call up the even gossiper $2x-1+\Delta_k$.

We claim that at the end of day k, gossipers $2x-1$ and $2x-1+\Delta_k$ each know all the knowledge originally held by gossipers $2x-1+\Delta_k/2-(2^{k-1}-\frac{1}{2})$ through $2x-1+\Delta_k/2+(2^{k-1}-\frac{1}{2})$. Note that when $k \leq \log_2 n$, this implies that each gossiper knows 2^k pieces of information at the end of day k, the maximum possible. And at the end of day $\lceil \log_2 n \rceil$, each gossiper knows everything.

We prove the claim by induction on k. When $k=0$, we have $\Delta_k = 0$ and the claim is that gossiper $2x-1$ has the knowledge of gossiper $2x-1$, which is clear. Assume the claim is true for $k=K$, where K is a nonnegative integer. On day $K+1$, gossiper $2x-1$ and gossiper $2x-1+\Delta_{K+1}$ exchange information. By the inductive hypothesis, gossiper $2x-1$ has the original knowledge of gossipers l_1 through r_1, and gossiper $2x-1+\Delta_{K+1}$ has the original knowledge of gossipers l_2 through r_2, where

$$l_1 = 2x-1+\frac{\Delta_K}{2}-\left(2^{K-1}-\frac{1}{2}\right),$$

$$r_1 = 2x-1+\frac{\Delta_K}{2}+\left(2^{K-1}-\frac{1}{2}\right),$$

$$l_2 = 2x-1+\Delta_{K+1}-\frac{\Delta_K}{2}-\left(2^{K-1}-\frac{1}{2}\right),$$

$$r_2 = 2x-1+\Delta_{K+1}-\frac{\Delta_K}{2}+\left(2^{K-1}-\frac{1}{2}\right).$$

A straightforward computation shows that $r_1+1=l_2$ if K is even, and $r_2+1=l_1$ if K is odd (in other words the two blocks of integers in $[l_1, r_1]$ and $[l_2, r_2]$ are contiguous), and that in either case, after day $K+1$, gossipers $2x-1$ and $2x-1+\Delta_{K+1}$ each know the original knowledge of gossipers $2x-1+\Delta_{K+1}/2-(2^K-\frac{1}{2})$ through $2x-1+\Delta_{K+1}/2+(2^K-\frac{1}{2})$. The claim follows by induction.

Thus, using this strategy, each gossiper will know everything at the end of day $\lceil \log_2 n \rceil$.

Final remarks Admittedly, one day is a bit much for the time it takes for even the most loquacious of rumormongers to communicate with a friend! Perhaps one hour or even one minute would have been more realistic.

Lest the reader find the topic of this note too frivolous, let us remark that the problem can also be interpreted in terms of communications between computers operating in parallel. The result can be useful, for instance, when multiple copies of a database are distributed to several locations and replication is needed to keep the databases synchronized. In fact, this is the situation in which our problem originally arose.

More on the Thinned-Out Harmonic Series

GERALD A. HEUER
Concordia College
Moorhead, MN 56562

Every mathematics student learns early on that the harmonic series $\Sigma_{k=1}^{\infty} 1/k$ is divergent. Many of us were amazed to learn a little later that if we omit all terms $1/k$ where in the decimal representation of k there is at least one occurrence of the digit 9, the resulting series converges. The same is true if "9" is replaced by any other digit $0, 1, \ldots, 7$, or 8, and the corresponding fact holds for any base b in place of 10.

In a recent note **[1]**, G. Hossein Behforooz discusses these series and related ones, and raises an interesting question. Following **[1]**, we write

$$T_r = \sum_{r \notin k} \frac{1}{k}$$

for the series which omits the terms $1/k$ where r occurs as a digit of k. Behforooz quotes from **[2]** the values of $T_0, T_1, \ldots, T_9$ when $b = 10$, where one may observe that $T_1 < T_2 < \cdots < T_9 < T_0$, and asks:

Open question *For any base b, is $T_m < T_n$, whenever $1 \leq m < n$?*

We think that many readers will have seen rather quickly that the question has a positive answer, as we show in Theorem 1 below. An equally interesting, and slightly more difficult question is the following:

Next question *For any base b, is $T_0 > T_r$ whenever $1 \leq r < b$?*

We show in Theorem 2 below that the answer here is again positive. Thus we always have $T_0 > T_{b-1} > \cdots > T_2 > T_1$.

THEOREM 1. *For any base b, if $1 \leq m < n < b$, then $T_m < T_n$.*

Proof. Note that if k contains both the digits m and n, then the term $1/k$ appears in neither sum, and if k contains neither m nor n, then $1/k$ is in both sums. In the remaining terms $1/k$ of T_m, the digit n occurs but m does not. For each such k, let k' be the integer obtained from k by changing each occurrence of the digit n to m. Then $1/k'$ is a term in T_n. Moreover, this mapping $k \mapsto k'$ is a bijection of the set of terms of T_m containing n to the terms of T_n containing m. (The inverse map changes each m to n.) Since $m < n$, we have $k' < k$, and therefore $1/k < 1/k'$. It follows that $T_m < T_n$. ∎

THEOREM 2. *For any base b, $T_0 > T_{b-1}$, and therefore $T_0 > T_r$ for every r, $1 \leq r < b$.*

The idea of the proof is similar, but the mapping is not as simple. We shall write the proof for base 10; it will be clear how the same argument works for base b. Let

$$D_{09} = \{k : 0 \notin k, 9 \notin k\},$$

$$D = \{k : 0 \in k, 9 \notin k\},$$

and

$$D' = \{k' : 0 \notin k', 9 \in k'\}.$$

Then

$$T_0 = \sum_{k \in D_{09}} \frac{1}{k} + \sum_{k' \in D'} \frac{1}{k'},$$

and

$$T_9 = \sum_{k \in D_{09}} \frac{1}{k} + \sum_{k \in D} \frac{1}{k}.$$

These representations reorder the terms of the sums T_0 and T_9, but these are absolutely convergent series, so reordering does not alter their sums. The sets D and D' begin as follows:

$$D = \{10, 20, \ldots, 80, 100, 101, 102, \ldots, 108, 110, \ldots\};$$

$$D' = \{9, 19, \ldots, 79, 89, 91, 92, \ldots, 98, 99, \ldots\}.$$

Let k_n and k'_n be the n-th elements of D and D', respectively, in the natural order. For each term $1/k_n$ with k_n in D occurring in T_9 we replace it by the term $1/k'_n$. This transforms T_9 into T_0. We see that initially each k_n is replaced by a smaller integer k'_n, so that $1/k'_n > 1/k_n$. Thus, $\frac{1}{9} > \frac{1}{10}, \frac{1}{19} > \frac{1}{20}, \ldots, \frac{1}{99} > \frac{1}{110}$. If we can show that this is always the case, i.e., that $k'_n < k_n$ for all n, it will follow that $T_0 > T_9$.

It suffices to show that for each k in D, the number of elements of D' less than k is greater than the number of elements of D less than k. Rather than count these directly, it is easier to include D_{09} with both. For $k \in D$, let $C(k)$ be the number of elements of $D \cup D_{09}$ less than k and $C'(k)$ the number of elements of $D' \cup D_{09}$ less than k. Since D, D' and D_{09} are disjoint, it will suffice to show that $C'(k) > C(k)$ for each $k \in D$. These counts are fairly straightforward, but it will be easier to follow the general case if we illustrate it first in a specific case. In Table 1 below we count $C(83047)$ and $C'(83047)$ in sections, by counting how many such integers there are in the interval $[1, 10)$, how many in $[10, 10^2)$, etc.

TABLE 1. The counts $C'(83047)$ and $C(83047)$

Interval	$C'(83047)$	$C(83047)$
$[1, 10)$	9	8
$[10, 100)$	9^2	$8 \cdot 9$
$[100, 1000)$	9^3	$8 \cdot 9^2$
$[1000, 10000)$	9^4	$8 \cdot 9^3$
$[10000, 80000)$	$7 \cdot 9^4$	$7 \cdot 9^4$
$[80000, 83000)$	$2 \cdot 9^3$	$3 \cdot 9^3$
$[83000, 83040)$	0	$4 \cdot 9$
$[83040, 83047)$	0	7

TABLE 2. The counts $C'(k)$ and $C(k)$ in general

Interval	$C'(k)$	$C(k)$
$[1, 10)$	9	8
$[10, 10^2)$	9^2	$8 \cdot 9$
$\vdots$		
$[10^{r-2}, 10^{r-1})$	9^{r-1}	$8 \cdot 9^{r-2}$
$[10^{r-1}, d_1 10^{r-1})$	$(d_1 - 1)9^{r-1}$	$(d_1 - 1)9^{r-1}$
$[d_1 10^{r-1}, d_1 10^{r-1} + d_2 10^{r-2})$	$(d_2 - 1)9^{r-2}$	$d_2 9^{r-2}$
$\vdots$		
$[d_1 10^{r-1} + \cdots + d_{q-2} 10^{r-q+2},$ $d_1 10^{r-1} + \cdots + d_{q-2} 10^{r-q+2} + d_{q-1} 10^{r-q+1})$	$(d_{q-1} - 1)9^{r-q+1}$	$d_{q-1} 9^{r-q+1}$
$[d_1 10^{r-1} + \cdots + d_{q-1} 10^{r-q+1}, k)$	0	at most 9^{r-q}, (but 0 if $r = q$)

Then we have

$$\begin{aligned} &C'(80347) - C(83047) \\ &\quad = (9-8) + (9-8)9 + (9-8)9^2 + (9-8)9^3 + (2-3)9^3 - 4 \cdot 9 - 7 \\ &\quad = 1 + 9 + 9^2 - 4 \cdot 9 - 7 > 0. \end{aligned}$$

We see that once we come to the 0-digit in k, we have zeros in the rest of the C' column. The most we could have had in the C column from this point on would be 9^2, corresponding to $k = 83088$. In this case $C'(k) - C(k)$ would be $1 + 9 + 9^2 - 9^2$, still positive. Here is how the argument goes in general:

Let $k = d_1 d_2 \ldots d_r$ be an element of D (so $0 \in k$, $9 \notin k$), and let d_q be the leftmost zero digit of k. The counts are then as displayed in Table 2 above:
Then

$$\begin{aligned} C'(k) - C(k) &\geq 1 + 9 + \cdots + 9^{r-2} - 1 \cdot 9^{r-2} - \cdots - 1 \cdot 9^{r-q+1} - 9^{r-q} \\ &= 1 + 9 + \cdots + 9^{r-q-1} > 0, \quad \text{if } q < r, \end{aligned}$$

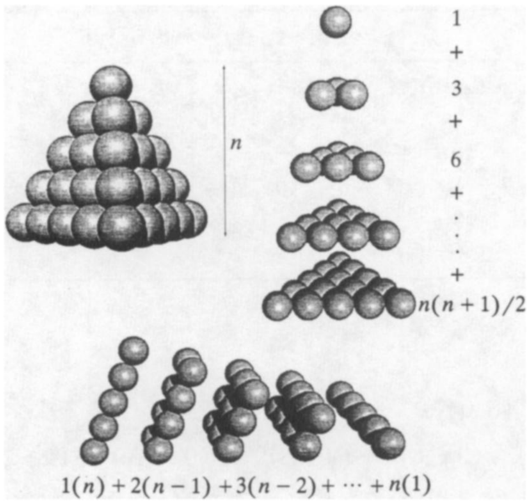
and if $q = r$, then $C'(k) - C(k) = 1$.

REFERENCES

1. G. Hossein Behforooz, Thinning out the harmonic series, this MAGAZINE 68 (1995), 289–293.
2. R. Baillie, Sums of reciprocals of integers missing a given digit, *Amer. Math. Monthly* 86 (1979), 372–374.

Proof Without Words: Counting Cannonballs

$$\sum_{i=1}^{n} i(n-i+1) = \sum_{i=1}^{n} \frac{i(i+1)}{2}$$



—Deanna B. Haunsperger and Stephen F. Kennedy
Carleton College
Northfield, MN 55057

Bernoulli's Identity Without Calculus

KENNETH S. WILLIAMS
Carleton University
Ottawa, Ontario
Canada K1S 5B6

A very simple way of defining the Bernoulli numbers B_m $(m = 0, 1, 2, \dots)$ is by means of the recurrence relation

$$B_m = -\frac{1}{m+1}\sum_{j=0}^{m-1}\binom{m+1}{j}B_j, \qquad m = 1, 2, 3, \dots, \tag{1}$$

where $B_0 = 1$ ([**2**, p. 229]). The more usual (and indeed equivalent) way is by means of the power series expansion of $x/(e^x - 1)$, namely,

$$\frac{x}{e^x - 1} = \sum_{m=0}^{\infty} B_m \frac{x^m}{m!}, \qquad |x| < 2\pi. \tag{2}$$

The definition (1) has the advantage over the definition (2) of not requiring an infinite series and its radius of convergence. The first few Bernoulli numbers are easily calculated from (1). Taking $m = 1$, 2, 3, and 4 in (1), we obtain

$$B_1 = -\frac{1}{2}B_0 = -\frac{1}{2},$$

$$B_2 = -\frac{1}{3}(B_0 + 3B_1) = \frac{1}{6},$$

$$B_3 = -\frac{1}{4}(B_0 + 4B_1 + 6B_2) = 0,$$

$$B_4 = -\frac{1}{5}(B_0 + 5B_1 + 10B_2 + 10B_3) = -\frac{1}{30}.$$

The sequence of Bernoulli numbers is one of the most important number sequences in mathematics. Bernoulli numbers arise naturally in the calculus of finite differences, in combinatorics, in connection with Fermat's last theorem, in the study of class numbers of algebraic number fields, in the Euler–Maclaurin summation formula, and in connection with values of the Riemann zeta function. The main properties of Bernoulli numbers can be found for example in Ireland and Rosen [**2**, Chap. 15]. An extensive bibliography of Bernoulli numbers has been compiled by Dilcher, Skula, and Slavutskii [**1**].

In 1987, Nunemacher and Young [**3**] made use of (2) to give an elegant proof of Bernoulli's identity

$$\sum_{r=1}^{n-1} r^k = \frac{1}{k+1}\sum_{j=0}^{k}\binom{k+1}{j}B_j n^{k+1-j}, \quad n, k = 1, 2, \dots, \tag{3}$$

which expresses the sum of the kth powers of the first $n-1$ natural numbers as a polynomial in n. For example, with $k=2$ we have

$$\begin{aligned} 1^2+2^2+\cdots+(n-1)^2 &= \frac{1}{3}\left(\binom{3}{0}B_0n^3+\binom{3}{1}B_1n^2+\binom{3}{2}B_2n\right)\\ &= \frac{1}{3}\left(n^3-\frac{3}{2}n^2+\frac{1}{2}n\right)\\ &= (n-1)n(2n-1)/6 \end{aligned}$$

and with $k=3$ we have

$$\begin{aligned} 1^3+2^3+\cdots+(n-1)^3 &= \frac{1}{4}\left(\binom{4}{0}B_0n^4+\binom{4}{1}B_1n^3+\binom{4}{2}B_2n^2+\binom{4}{3}B_3n\right)\\ &= \frac{1}{4}(n^4-2n^3+n^2)\\ &= ((n-1)n/2)^2. \end{aligned}$$

The identity (3) for $k=1,2,\dots,10$ was given by Jakob Bernoulli (1654–1705) in his work "Ars Conjectandi" published posthumously in 1713. This was the first occurrence of Bernoulli numbers in the literature. The classical proofs of (3) make use of the Euler–Maclaurin summation formula or the derivatives and integrals of Bernoulli polynomials. They thus involve knowledge of calculus. This paper was motivated by the desire to find a proof of (3) which proceeds directly from the recurrence relation (1) to obtain (3) without any use of calculus or infinite series. We now present such a proof which is therefore accessible to students who have not studied calculus. It only uses the basic properties of binomial coefficients.

We begin by giving a sketch of the proof and then give the missing details. For a nonnegative integer m we define the Kronecker delta symbol δ_m by

$$\delta_m=\begin{cases}1, & \text{if } m=1,\\ 0, & \text{if } m\neq 1.\end{cases} \tag{4}$$

An easy deduction from the recurrence relation (1) is the following expression for δ_m, namely,

$$\delta_m=\sum_{j=0}^{m}\binom{m}{j}B_j-B_m, \qquad m=0,1,2,\dots. \tag{5}$$

Let k and n be positive integers and let r and l be integers with $1\le r\le n-1$ and $0\le l\le k+1$. Appealing to (4) we see that

$$(k+1)r^k=\sum_{l=0}^{k+1}\binom{k+1}{l}\delta_{k+1-l}\,r^l. \tag{6}$$

Using (5) in (6) we obtain

$$(k+1)r^k=\sum_{l=0}^{k+1}\binom{k+1}{l}\left\{\sum_{j=0}^{k+1-l}\binom{k+1-l}{j}B_j-B_{k+1-l}\right\}r^l. \tag{7}$$

Manipulation of the right side of (7) yields

$$(k+1)r^k = \sum_{j=0}^{k+1} \binom{k+1}{j} B_j \{(r+1)^{k+1-j} - r^{k+1-j}\}. \tag{8}$$

Then summing (8) over $r = 1, 2, \ldots, n-1$ the right side telescopes and we obtain (3).

Now we give the details. First we prove (5). For $m = 0$ and $m = 1$ (5) is easily checked. For $m \geq 2$, appealing to (1), we see that

$$\sum_{j=0}^{m} \binom{m}{j} B_j - B_m = \sum_{j=0}^{m-2} \binom{m}{j} B_j + mB_{m-1} = -mB_{m-1} + mB_{m-1} = 0 = \delta_m.$$

Next we deduce (8) from (7). From (7) we see that

$$(k+1)r^k = S_1 - S_2,$$

where

$$S_1 = \sum_{l=0}^{k+1} \sum_{j=0}^{k+1-l} \binom{k+1}{l} \binom{k+1-l}{j} B_j r^l$$

and

$$S_2 = \sum_{l=0}^{k+1} \binom{k+1}{l} B_{k+1-l}\, r^l.$$

Now

$$\begin{aligned} \binom{k+1}{l}\binom{k+1-l}{j} &= \frac{(k+1)!}{l!(k+1-l)!} \frac{(k+1-l)!}{j!(k+1-l-j)!} \\ &= \frac{(k+1)!}{j!(k+1-j)!} \frac{(k+1-j)!}{l!(k+1-l-j)!} \\ &= \binom{k+1}{j}\binom{k+1-j}{l} \end{aligned}$$

so

$$S_1 = \sum_{l=0}^{k+1} \sum_{j=0}^{k+1-l} \binom{k+1}{j} \binom{k+1-j}{l} B_j r^l.$$

Interchanging the order of summation, we obtain

$$S_1 = \sum_{j=0}^{k+1} \binom{k+1}{j} B_j \sum_{l=0}^{k+1-j} \binom{k+1-j}{l} r^l.$$

By the binomial theorem, the inner sum is $(r+1)^{k+1-j}$, so that

$$S_1 = \sum_{j=0}^{k+1} \binom{k+1}{j} B_j (r+1)^{k+1-j}.$$

Changing the summation variable l in the sum S_2 to $j = k+1-l$, we deduce that

$$\begin{aligned} S_2 &= \sum_{j=0}^{k+1} \binom{k+1}{k+1-j} B_j r^{k+1-j} \\ &= \sum_{j=0}^{k+1} \binom{k+1}{j} B_j r^{k+1-j} \end{aligned}$$

Then we have

$$(k+1)r^k = S_1 - S_2 = \sum_{j=0}^{k+1} \binom{k+1}{j} B_j\left((r+1)^{k+1-j} - r^{k+1-j}\right),$$

which is (7).

Finally, we sum (8) over $r = 1, 2, \dots, n-1$ and obtain

$$\begin{aligned} (k+1)\sum_{r=1}^{n-1} r^k &= \sum_{j=0}^{k+1} \binom{k+1}{j} B_j\left(n^{k+1-j} - 1\right) \\ &= \sum_{j=0}^{k+1} \binom{k+1}{j} B_j n^{k+1-j} - B_{k+1} \quad (\text{by } (5)) \\ &= \sum_{j=0}^{k} \binom{k+1}{j} B_j n^{k+1-j}, \end{aligned}$$

completing the proof of Bernoulli's identity (3).

Acknowledgment The author would like to acknowledge suggestions by the referees, the editor, and James G. Huard (Canisius College, Buffalo, NY), which led to improvements in the first version of this article.

REFERENCES

1. K. Dilcher, L. Skula and I. Sh. Slavutskii, *Bernoulli Numbers Bibliography (1713–1990)*, Queen's Papers in Pure and Applied Mathematics No. 87, Queen's University, Kingston, Ontario, Canada, 1991.
2. K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, Springer-Verlag, New York, 1982.
3. J. Nunemacher and R. M. Young, On the sum of consecutive kth powers, this MAGAZINE 60 (1987), 237–238.

A Tree for Generating Bernoulli Numbers

S. C. WOON
Trinity College
Cambridge CB2 1TQ, England

The familiar Bernoulli numbers B_k are most succinctly defined by a generating function [1]

$$\frac{z}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n.$$

The odd-numbered B's vanish, except for $B_1 = -1/2$. This may be established by observing that $z/(e^z - 1) + z/2$ is an even function. This gives rise to an alternate formulation of the generating function

$$\frac{z}{e^z - 1} + \frac{z}{2} = \sum_{n=0}^{\infty} \frac{B_{2n}}{2n!} z^{2n},$$

which defines just the even numbered B's. Bernoulli numbers appear in connection with a variety of topics, including sums of powers of integers and the Riemann zeta function. In this note we display a binary tree generated by simple arithmetic operations that can be used to generate the Bernoulli numbers. We call the tree *Bernoulli's tree*.

Each node of Bernoulli's tree is an expression of the form $\pm 1/a!b!\ldots$. Two operations, $\mathbf{O}_0$ and $\mathbf{O}_1$, are applied to a node to generate the node's children. The operations are defined as follows

$$\frac{\pm}{a!b!\ldots} \overset{\mathbf{O}_0}{\mapsto} \frac{\mp 1}{(a+1)!b!\ldots}$$

and

$$\frac{\pm 1}{a!b!\ldots} \overset{\mathbf{O}_1}{\mapsto} \frac{\pm 1}{2!a!b!\ldots}.$$

The starting node for the tree is $(+1/2!)$. The first several rows in the tree are illustrated in FIGURE 1, with the convention that $\mathbf{O}_0$ is applied when following the left arrow emanating from a node, and $\mathbf{O}_1$ is applied when following the right arrow. For example, from the initial node of $(+1/2!)$ we apply $\mathbf{O}_0$ to obtain $(-1/3!)$ below and to the left, while applying $\mathbf{O}_1$ produces $(+1/2!2!)$ below and to the right.

Let S_n be the sum of the nodes in the nth row of the tree diagram, where the initial node is counted as row 0. The starting node is $(+1/2!)$, so $S_0 = 1/2$. In row 1, there are two nodes: $(-1/3!)$ and $(-1/2!2!)$. Summing these gives $S_1 = 1/12$. In general, it can be shown that $S_{2n} = 0$ for $n \in \mathbb{Z}^+$ (the set of positive integers). The odd numbered S's do not vanish, and are closely related to the even Bernoulli numbers. Indeed, we will show below that $B_{2n} = (2n)!S_{2n-1}$. Thus, we can compute Bernoulli numbers using Bernoulli's tree.

Before proceeding to the proof, we observe that the definition of S_n can be given in purely symbolic terms. It is clear that each term in row n can be obtained by applying a sequence of n operators $\mathbf{O}_0$ or $\mathbf{O}_1$ to the initial node $(+1/2!)$, and that every such

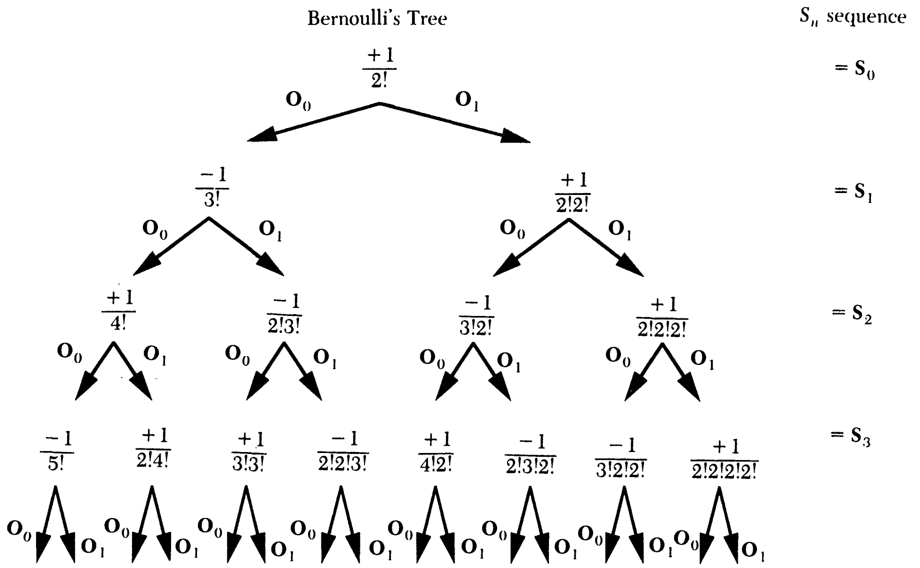


FIGURE 1

sequence of operations corresponds to one node. If $\sigma = (\sigma_1, \ldots, \sigma_n)$, with each σ_k either 0 or 1, define $\mathbf{O}_\sigma$ to be the composition $\mathbf{O}_{\sigma_1}\mathbf{O}_{\sigma_2} \ldots \mathbf{O}_{\sigma_n}$. Then we may write

$$S_n = \sum_\sigma \mathbf{O}_\sigma\left(\frac{+1}{2!}\right).$$

Alternatively, interpret the expression $(\mathbf{O}_0 + \mathbf{O}_1)^n$ as an operator by formally expanding the power, respecting the fact that $\mathbf{O}_0$ and $\mathbf{O}_1$ do not commute. Then the terms in the expression are just $\mathbf{O}_\sigma$. Thus

$$S_n = (\mathbf{O}_0 + \mathbf{O}_1)^n\left(\frac{-1}{2!}\right).$$

For example,

$$\begin{aligned}
S_2 &= (\mathbf{O}_0 + \mathbf{O}_1)^2\left(\frac{-1}{2!}\right) \\
&= (\mathbf{O}_0 + \mathbf{O}_1)(\mathbf{O}_0 + \mathbf{O}_1)\left(\frac{+1}{2!}\right) \\
&= (\mathbf{O}_0\mathbf{O}_0 + \mathbf{O}_0\mathbf{O}_1 + \mathbf{O}_1\mathbf{O}_0 + \mathbf{O}_1\mathbf{O}_1)\left(\frac{+1}{2!}\right) \\
&= \frac{+1}{4!} + \frac{-1}{2!3!} + \frac{-1}{3!2!} + \frac{+1}{2!2!2!} \\
&= 0.
\end{aligned}$$

In this context, though, we bear in mind that it is only valid to reassociate the factors $(\mathbf{O}_0 + \mathbf{O}_1)$ with $(+1/2!)$ so long as we refrain from any arithmetic simplification of the results. Once the formal representation of $(\mathbf{O}_0 + \mathbf{O}_1)(+1/2!)$ is modified, we lose the ability to apply additional operators. After all, the operators are defined in terms of the formal expressions on which they act, rather than in terms of the numerical values of those expressions. Thus, while we may write $S_1 = 1/12 = (1/2!3!)$, it is not valid to conclude that $S_2 = (\mathbf{O}_0 + \mathbf{O}_1)(1/12)$ or even that $S_2 = (\mathbf{O}_0 + \mathbf{O}_1)(+1/2!3!)$.

We conclude the paper by proving, using the Riemann zeta function, that $B_n = n! S_{n-1}$ for $n \in \mathbb{Z}^+$, i.e., S_n vanishes for even $n \geq 2$, and $B_{2n} = (2n)! S_{2n-1}$.

The Riemann zeta function is defined as

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} \qquad \text{for } \operatorname{Re}(s) > 1,\ s \in \mathbf{C}. \tag{1}$$

It can be analytically continued to the left-half of the complex plane by the Euler–Maclaurin summation formula [2][3][4]. For $\operatorname{Re}(s) > -2M - 1$,

$$\zeta(s) = \lim_{N\to\infty} \left[\begin{array}{c} \displaystyle\sum_{n=1}^{N} n^{-s} \frac{1}{-s+1} N^{-s+1} - B_1 N^{-s} \\ \displaystyle - \sum_{k=1}^{M} \frac{B_{2k}}{(2k)!} \frac{\partial^{2k-1}}{\partial N^{2k-1}} (N^{-s}) \\ + O(N^{-s-2M-1}) \end{array} \right]. \tag{2}$$

Now, being adventurous, we choose instead to analytically continue $\zeta(s)$ the long way. Consider the difference between (1) and its analogous integral:

$$\begin{aligned} &\lim_{N\to\infty} \left[\sum_{n=1}^{N} n^{-s} - \int_0^N x^{-s}\,dx \right] \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \left[n^{-s} - \int_{n-1}^{n} x^{-s}\,dx \right] \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \left[n^{-s} - \frac{1}{-s+1} n^{-s+1} + \frac{1}{-s+1} (n-1)^{-s+1} \right]. \end{aligned}$$

Binominal expansion of the last term gives a set of series

$$= \lim_{N\to\infty} \left[\begin{array}{ll} \displaystyle\sum_{n=1}^{N} \left[n^{-s} - \frac{1}{-s+1} n^{-s+1} + \frac{1}{-s+1} n^{-s+1} + (-1) n^{-s} \right] & x(\text{sum to } 0) \\ \displaystyle + \sum_{n=1}^{N} \frac{(-s)}{2!} n^{-s-1} & (3) \\ \displaystyle + \sum_{n=1}^{N} \frac{(-1)(-s)(-s-1)}{3!} n^{-s-2} & (4) \\ \vdots & \\ \displaystyle + \sum_{n=1}^{N} \frac{(-1)^{m+1}(-s)(-s-1)\dots(-s+1-m)}{(m+1)!} n^{-s-m} & (5) \\ \vdots & \end{array} \right]$$

$$= \lim_{N\to\infty} \sum_{m=1}^{\infty} \sum_{n=1}^{N} \left[\left[\prod_{j=1}^{m} (-s+1-j) \right] \left(\frac{(-1)^{m+1}}{(m+1)!} \right) n^{-s-m} \right]$$

$$= \lim_{N\to\infty} \sum_{m=1}^{\infty} \sum_{n=1}^{N} \left[\left[\prod_{j=1}^{m} (-s+1-j) \right] \mathbf{O}_0^{m-1} \left(\frac{+1}{2!} \right) n^{-s-m} \right]$$

where $\mathbf{O}_0^{m-1}$ represents a string of $(m-1)$ $\mathbf{O}_0$ operator(s), and $\mathbf{O}_0^{m-1}(+1/2!)$ corresponds to the *first* term of S_{m-1}, $m \in \mathbb{Z}^+$, in Bernoulli's tree. $\mathbf{O}_0^0$ is taken as the identity operator.

Now, $\lim_{N\to\infty} \int_0^N kx^{-p}\,dx = \lim_{N\to\infty}[kN^{-p-1}/-p+1]$ converges when $\mathrm{Re}(p) > 1$, where constants, $k, p \in \mathbb{C}$. Therefore, $\lim_{N\to\infty} \sum_{n=1}^N kn^{-p}$ converges when $\mathrm{Re}(p) > 1$. Hence,

$$\begin{aligned}\zeta(s) &= \lim_{N\to\infty}\left[\sum_{n=1}^{N} n^{-s} - \int_0^N x^{-s}\,dx\right] \\ &= \lim_{N\to\infty}\left[\sum_{n=1}^{N} n^{-s} - \frac{1}{-s+1}N^{-s+1}\right]\end{aligned}$$

defines an analytic continuation of $\zeta(s)$ for $\mathrm{Re}(s) > 0$, since the series

$$\begin{array}{lcr} (3) & & >0 \\ (4) & \text{converges when } \Re(s) & >-1 \\ (5) & & >-m+1. \end{array}$$

Similarly, find the difference between the series (3) and its analogous integral.

$$\lim_{N\to\infty}\left[\sum_{n=1}^{N}\frac{(-s)}{2!}n^{-s-1} - \int_0^N \frac{(-s)}{2!}x^{-s-1}\,dx\right]$$

$$= \lim_{N\to\infty}\left[\begin{array}{lr} \displaystyle\sum_{n=1}^{N}\frac{(-s)(-s-1)}{2!2!}n^{-s-2} & (6) \\ \vdots & \\ \displaystyle +\sum_{n=1}^{N}\frac{(-1)^{q+1}(-s)(-s-1)\dots(-s-q)}{2!(q+1)!}n^{-s-1-q} & (7) \\ \vdots & \end{array}\right]$$

$$= \lim_{N\to\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{N}\left[\left[\prod_{j=1}^{m+1}(-s+1-j)\right]\left(\frac{(-1)^{m+1}}{2!(m-1)!}\right)n^{-s-m-1}\right]$$

$$= \lim_{N\to\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{N}\left[\left[\prod_{j=1}^{m+1}(-s+1-j)\right]\mathbf{O}_1\mathbf{O}_0^{m-1}\left(\frac{+1}{2!}\right)n^{-s-m-1}\right]$$

where $\mathbf{O}_1\mathbf{O}_0^{m-1}(+1/2!)$ corresponds to the *second* term of S_m, $m \in \mathbb{Z}^+$, in Bernoulli's tree.

Hence,

$$\begin{aligned}\zeta(s) &= \lim_{N\to\infty}\left[\sum_{n=1}^{N} n^{-s} - \int_0^N x^{-s}\,dx - \int_0^N \frac{(-s)}{2!}x^{-s-1}\,dx\right] \\ &= \lim_{N\to\infty}\left[\sum_{n=1}^{N} n^{-s} - \frac{+1}{-s+1}N^{-s-+1} - \frac{1}{2!}N^{-s}\right] \\ &= \lim_{N\to\infty}\left[\sum_{n=1}^{N} n^{-s} - \frac{+1}{-s+1}N^{-s+1} - S_0 N^{-s}\right]\end{aligned}$$

defines an analytic continuation of $\zeta(s)$ for $\operatorname{Re}(s) > -1$ since the series

(6) > -1

(7) converges when $\operatorname{Re}(s)$ $> -q.$

Note that $(+1/2!)$, the coefficient of N^{-s}, corresponds to S_0, the starting node of Bernoulli's tree.

To analytically continue $\zeta(s)$ to $\operatorname{Re}(s) > -2$, we again subtract in a similar way the set of series of n^{-s-2} (e.g., (4) and (5)), which diverge for $\operatorname{Re}(s) \leq -1$ to get sets of series of higher order that converge for $\operatorname{Re}(s) > -2$.

Similarly, the difference between the series (4) and its analogous integral, and the corresponding one for series (5) are

$$\lim_{N\to\infty} \sum_{m=1}^{\infty} \sum_{n=1}^{N} \left[\left[\prod_{j=1}^{m+3} (-s+1-j) \right] \mathbf{O}_0 \mathbf{O}_1 \mathbf{O}_0^{m-1} \left(\frac{-1}{2!} \right) n^{-s-m-3} \right]$$

and

$$\lim_{N\to\infty} \sum_{m=1}^{\infty} \sum_{n=1}^{N} \left[\left[\prod_{j=1}^{m+3} (-s+1-j) \right] \mathbf{O}_1^2 \mathbf{O}_0^{m-1} \left(\frac{+1}{2!} \right) n^{-s-m-3} \right]$$

respectively, where $\mathbf{O}_0\mathbf{O}_1\mathbf{O}_0^{m-1}(+1/2!)$ and $\mathbf{O}_1^2\mathbf{O}_0^{m-1}(+1/2!)$ correspond to the *third* and *fourth* terms respectively of S_{m+1}, $m \in \mathbb{Z}^+$, in Bernoulli's tree.

Hence,

$$\zeta(s) = \lim_{N\to\infty} \begin{bmatrix} \displaystyle\sum_{n=1}^{N} n^{-s} - \int_N^0 x^{-s}\, dx - \int_0^N \frac{(-s)}{2!} x^{-s-1}\, dx \\ \displaystyle -\int_0^N (-s)(-s-1)\mathbf{O}_0\left(\frac{+1}{2!}\right) x^{-s-2}\, dx \\ \displaystyle -\int_0^N (-s)(-s-1)\mathbf{O}_1\left(\frac{+1}{2!}\right) x^{-s-2}\, dx \end{bmatrix}$$

$$= \lim_{N\to\infty} \begin{bmatrix} \displaystyle\sum_{n=1}^{N} n^{-s} - \frac{1}{-s+1} N^{-s+1} - S_0 N^{-s} \\ -S_1(-s) N^{-s-1} \end{bmatrix}$$

defines an analytic continuation of $\zeta(s)$ for $\operatorname{Re}(s) > -2$, where $S_0 = (+1/2!)$ and $S_1 = (\mathbf{O}_0 + \mathbf{O}_1)(+1/2!)$.

Comparing the last line with (2) while letting $M = 1$

$$\zeta(s) = \lim_{N\to\infty} \left[\sum_{n=1}^{N} n^{-s} - \frac{1}{-s+1} N^{-s+1} - B_1 N^{-s} - \frac{B_2}{2!}(-s) N^{-s-1} \right]$$

we find that $B_2 = 2! S_1$.

At this point, we observe that every time we find the difference between a divergent series and its analogous integral, the following binomial expansion has the effect of inserting additional operators $\mathbf{O}_1\mathbf{O}_0^{m-1}$ immediately before $(+1/2!)$. When we write down these sequences, Bernoulli's tree appears.

To obtain further analytic continuation of $\zeta(s)$, we make use of the tree and write

$$\zeta(s) = \lim_{N\to\infty} \left[\begin{array}{l} \displaystyle\sum_{n=1}^{N} n^{-s} - \frac{1}{-s+1} N^{-s+1} - S_0 N^{-s} \\ \displaystyle - \sum_{m=1}^{M} \left[\left[\prod_{j=1}^{m} (-s+1-j) \right] S_m N^{-s-m} \right] \end{array} \right],$$

where $$S_m = (\mathbf{O}_0 + \mathbf{O}_1)^m \left(\frac{+1}{2!} \right).$$

Comparing the above with (2), and noting that B_n vanishes for odd $n \geq 3$, and so the coefficients of N^{-s-n+1} also vanish for odd $n \geq 3$, we get

$$B_n = n! S_{n-1} \quad \text{for } n \in \mathbb{Z}^+.$$

Conclusion

No adventure, no tree!

Acknowledgements. Many thanks to the referees for helpful comments. Special thanks go to Peter Magyar, a penfriend and then a graduate student at Harvard while I was finishing high school. Initially, I wanted to evaluate $\zeta(s)$ in the left-half of the complex plane. After I tried extending it in the straightforward but laborious way, he pointed out to my chagrin that I had actually done "Euler–Maclaurin summation" term by term! Ironically, it's amusing that had I come across the Euler–Maclaurin summation formula earlier, I might have missed bumping into Bernoulli's tree in the forest.

REFERENCES

1. Michael Spivak, *Calculus*, Benjamin, New York, NY, 1967, p. 478, Problem 15, and p. 482, Problem 17.
2. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, Cambridge University Press, Cambridge, U.K., 1927, p. 128.
3. E. Jahnke, F. Emde, and F. Lösch, *Tables of Higher Functions*, McGraw-Hill, New York, NY, 1960, p. 38.

Mapping the Cantor Set onto [0, 1]

MELISSA RICHEY
Millsaps College
Jackson, MS 39210

Introduction The example of a continuous surjection of the Cantor set C onto [0, 1] that is given in most topology texts has been around for some time ([**1**], [**5**]). It involves writing each $x \in C$ in ternary form $x = .x_1x_2\ldots$ where each $x_i \in \{0, 2\}$, then defining $f\colon C \to [0,1]$ by $f(x) = \sum_{i=1}^{\infty} x_i/2^{i+1}$. Our construction is quite different, although the functions are the same. We use the classic middle-thirds description of C and employ a technique for defining the continuous functions previously used in [**2**] and [**3**].

The Construction. We begin with the following two facts [**4**] and some notation.

(a) The nested intersection of non-empty compact subsets of $\mathbb{R}^2$ is non-empty and compact.
(b) Let $X \subset \mathbb{R}$ be compact. The function $f\colon X \to \mathbb{R}$ is continuous if and only if the graph of f is compact.

Let $\Pi_x, \Pi_y\colon \mathbb{R}^2 \to \mathbb{R}$ denote the x- and y-coordinate projections, respectively (i.e., $\Pi_x((a,b)) = a$, $\Pi_y((a,b)) = b$) and, for $A \subset \mathbb{R}^2$, let $A[x] = \{y\colon (x,y) \in A\}$. Define compact sets $C_0 \supset C_1 \supset C_2 \supset \ldots$ as in FIGURE 1. We obtain C_1 from C_0 as follows: Remove the open middle-third rectangle from C_0 to obtain FIGURE 2.

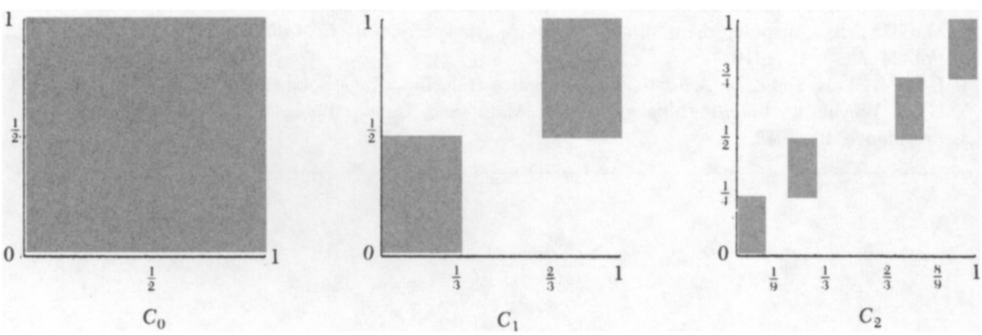


FIGURE 1

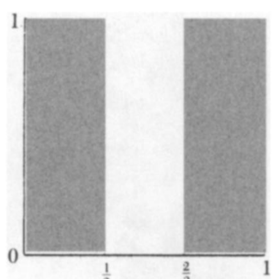


FIGURE 2

Next, keep the lower half of the left rectangle and the upper half of the right rectangle. This yields C_1. Notice that $\text{lub}\{|y_2 - y_1|: y_1, y_2 \in C_1[x]\} = \text{diam}(C_1[x]) \leq \frac{1}{2}$ for each $x \in [0,1]$. To get C_2 from C_1, repeat the above procedure on each rectangle in C_1. Again, $\text{diam}(C_2[x]) \leq 1/2^2$ for each $x \in [0,1]$. *Continue*. We obtain a collection of compact sets $C_0 \supset C_1 \supset C_2 \supset \dots$ with the following properties.

(i) $\text{diam}(C_n[x]) \leq 1/2^n$ for each $x \in [0,1]$;
(ii) $\Pi_x(\bigcap C_n) = \bigcap \Pi_x(C_n) = C$, the Cantor set;
(iii) $\Pi_y(\bigcap C_n) = \bigcap \Pi_y(C_n) = [0,1]$.

Note that the inclusions $\supseteq$ in (ii) and (iii) do not hold in full generality but hold in this case since the C_n's are compact. Let $K = \bigcap C_n$. By (i), $\text{diam}(K[x]) = 0$ for each $x \in [0,1]$ so that K is the graph of a function f (this is the vertical line test). By (ii), the domain of f is the Cantor set and, by (iii), the range of f is $[0,1]$. Hence, $f\colon C \to [0,1]$ is onto. Since K is compact (by (a)), f is continuous (by (b)).

Remark. The function f defined above is not one-to-one. Of course, it cannot be: A one-to-one continuous map between compact Hausdorff spaces is a homeomorphism, but C and $[0,1]$ are not homeomorphic since one is connected and the other is totally disconnected. This can also be seen from the geometric construction of f. When C_1 is formed from C_0, the upper right-hand corner of the lower rectangle, namely $(\frac{1}{3},\frac{1}{2})$, has the same y-coordinate as the lower left-hand corner of the upper rectangle, namely $(\frac{2}{3},\frac{1}{2})$. These corner points and their analogues survive in the intersection $\bigcap C_n$, so $f(\frac{1}{3}) = f(\frac{2}{3}) = \frac{1}{2}$.

REFERENCES

1. R. Engelking, *General Topology*, Polish Scientific Publ., Warszawa, 1977.
2. Mark Lynch, A continuous, nowhere differentiable function. *Amer. Math. Monthly*, 99 (1992), 8–9.
3. Mark Lynch, A mapping theorem for compact metric spaces, *Int. J. Math. Educ. Sci. Technol.*, 1993, Vol. 24, no. 3, 413–416.
4. J. R. Munkres, *Topology: A First Course*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1975.
5. G. T. Whyburn, *Analytic Topology*, Amer. Math. Soc. Colloq. Publ., Vol. 28, Amer. Math. Soc., Providence, RI, 1942.

Some Remarks on the Evaluation of $\int \frac{dt}{t^m + 1}$

JUDITH A. PALAGALLO
THOMAS E. PRICE
University of Akron
Akron, OH 44325-4002

Ramanujan evaluated the integral $I_m = \int_0^x \frac{1}{t^m+1}\,\mathrm{d}t$ for $m = 1, 2, 3, 4, 5, 6, 8$, and 10 ([**2**]). Recently, Gopalan and Ravichandran ([**1**]) discussed a method for computing I_m when $m = 2^n$ $(n \geq 2)$. In this note we describe a technique for calculating

$$F_m(t) := \int \frac{\mathrm{d}t}{t^m+1} \tag{1}$$

for any rational number m. Since $\frac{1}{t^{-m}+1} = 1 - \frac{1}{t^m+1}$ we need only consider the case for positive values of m. (The case when $m = 0$ is trivial.) We begin by finding F_m for m a positive integer.

Let $\theta_j := [(2j-1)\pi/m]$ and $t_j := \cos\theta_j + i\sin\theta_j$ for $j = 1, \dots, m$. By DeMoivre's theorem $t_j^m = -1$ for each j so that $\{t_j\}_{j=1}^m$ are the *distinct* zeros of $P(t) := t^m + 1$. Thus, by partial fractions decomposition, there are constants a_j, $j = 1, \dots, m$, such that

$$\frac{1}{t^m+1} = \sum_{j=1}^{m} \frac{a_j}{t-t_j}, \qquad t \neq t_j. \tag{2}$$

For a fixed k, $1 \leq k \leq m$, we have

$$1 = \lim_{t\to t_k} P(t) \sum_{j=1}^{m} \frac{a_j}{t-t_j} = a_k \lim_{t\to t_k} mt^{m-1} = -a_k m/t_k,$$

which implies that $a_j = -t_j/m$, $j = 1, \dots, m$.

First suppose that $m = 2n$. A little manipulation with trigonometric identities verifies that $t_j = \bar{t}_{m-j+1}$ and $a_j = \bar{a}_{m-j+1}$, $j = 1, \dots, n$. (In view of these relationships the decomposition (2) above could have been obtained using real arithmetic with quadratic denominators. Complex arithmetic is often convenient to use with partial fractions decomposition since the unknown coefficients are more easily determined. In the spirit of [**1**] we will solve the remainder of this problem using real integration. Although more abstract, complex integration has certain advantages and will be used below to derive a formula for F_m when m is rational.) A few simple calculations starting with (2) establish

$$\begin{aligned}\frac{1}{t^m+1} &= 2\sum_{j=1}^{n} \operatorname{Re}\left(\frac{a_j}{t-t_j}\right)\\ &= -\frac{2}{m}\sum_{j=1}^{n} \frac{\cos\theta_j(t-\cos\theta_j) - \sin^2\theta_j}{(t-\cos\theta_j)^2 + \sin^2\theta_j}.\end{aligned} \tag{3}$$

Since

$$\int \frac{\cos\theta_j(t-\cos\theta_j)-\sin^2\theta_j}{(t-\cos\theta_j)^2+\sin^2\theta_j}\,dt$$

$$=\cos\theta_j\int \frac{(t-\cos\theta_j)}{(t-\cos\theta_j)^2+\sin^2\theta_j}\,dt-\int \frac{\sin^2\theta_j}{(t-\cos\theta_j)^2+\sin^2\theta_j}\,dt$$

$$=\frac{1}{2}\cos\theta_j\log\left(t^2-2t\cos\theta_j+1\right)-\sin\theta_j\arctan\left(\frac{t-\cos\theta_j}{\sin\theta_j}\right)+C,$$

we have from (3)

$$F_m(t)=-\frac{1}{m}\sum_{j=1}^{n}\left[\cos\theta_j\log\left(t^2-2t\cos\theta_j+1\right)-2\sin\theta_j\arctan\left(\frac{t-\cos\theta_j}{\sin\theta_j}\right)\right]+C. \tag{4}$$

To see that this reduces to the same formula obtained in [**1**] for m a power of 2, let $m=2^k$ $(k\geq 2)$. Using the identity for the tangent of the sum of two angles and the fact that $\theta_j=\pi-\theta_{\frac{m}{2}-j+1}$, $j=1,\dots,\frac{m}{4}=2^{k-2}$, equation (4) becomes

$$F_m(t)=\frac{1}{m}\sum_{j=1}^{m/4}\left[\cos\theta_j\log\left(\frac{t^2+2t\cos\theta_j+1}{t^2-2t\cos\theta_j+1}\right)+2\sin\theta_j\arctan\left(\frac{2t\sin\theta_j}{1-t^2}\right)\right]+C.$$

Since $\sin\theta_{\frac{m}{4}-j+1}=\sin(\pi/2-\theta_j)=\cos\theta_j$, $j=1,\dots,m/4$, we obtain

$$F_m(t)=\frac{1}{m}\sum_{j=1}^{m/4}\cos\theta_j\left[\log\left(\frac{t^2+2t\cos\theta_j+1}{t^2-2t\cos\theta_j+1}\right)+2\arctan\left(\frac{2t\cos\theta_j}{1-t^2}\right)\right]+C. \tag{5}$$

By repeatedly using the half-angle formula for the cosine function one easily verifies that, for each j, $\cos\theta_j$ agrees with the value prescribed by equations (1) and (6) in [**1**].

Example 1. Consider $I:=\int_0^1\frac{dt}{t^{12}+1}$. Since the exponent is not a power of two, the formula given in (1) does not apply; and evidently F_{12} was not evaluated by Ramanujan. Note that equation (5) remains valid as long as m is divisible by four. Hence,

$$I=\lim_{t\to 1^-}\left[F_{12}(t)-F_{12}(0)\right]=\frac{1}{12}\sum_{j=1}^{3}\cos\theta_j\left[\log\left(\frac{1+\cos\theta_j}{1-\cos\theta_j}\right)+\pi\right]\approx 0.94747.$$

If $m=2n+1$, then the decomposition of (2) takes the form

$$\frac{1}{t^m+1}=-\frac{1}{m}\left[2\sum_{j=1}^{n}\mathrm{Re}\left(\frac{t_j}{t-t_j}\right)-\frac{1}{t+1}\right].$$

Hence,

$$F_m(t) = -\frac{1}{m}\sum_{j=1}^{n}\left[\cos\theta_j \log\left(t^2 - 2t\cos\theta_j + 1\right) - 2\sin\theta_j \arctan\left(\frac{t-\cos\theta_j}{\sin\theta_j}\right)\right] + \frac{\log(t+1)}{m} + C.$$

Finally, let $m =: p/q$ be any positive rational number where p and q are relatively prime. We assume that $p = 2n$; the case for p odd is similar. For $j = 1, \dots, p$, let $\theta_j := (2j-1)\pi/p$ and $t_j := e^{i\theta_j}$, and set $u = t^{1/q}$. A simple modification of the argument used to determine the coefficients for equation (2) yields

$$\frac{1}{t^{p/q}+1} = \frac{1}{u^p+1} = -\frac{1}{p}\sum_{j=1}^{p}\frac{t_j}{u-t_j}.$$

Noting that $dt = qu^{q-1}\,du$ we have

$$F_{p/q}(t) = -\frac{q}{p}\sum_{j=1}^{p} t_j \int \frac{u^{q-1}}{u-t_j}\,\mathrm{d}u.$$

Substituting

$$\frac{u^{q-1}}{u-t_j} = \frac{t_j^{q-1}}{u-t_j} + \sum_{s=0}^{q-2} u^s t_j^{q-2-s}$$

into this last equation yields

$$\begin{aligned} F_{p/q}(t) &= -\frac{q}{p}\sum_{j=1}^{p} t_j\left[t_j^{q-1}\int\frac{\mathrm{d}u}{u-t_j} + \sum_{s=0}^{q-2} t_j^{q-2-s}\int u^s\,\mathrm{d}u\right] \\ &= -\frac{q}{p}\left[\sum_{j=1}^{p} t_j^q \log(u-t_j) + \sum_{s=0}^{q-2}\frac{u^{s+1}}{s+1}\sum_{j=1}^{p} t_j^{q-1-s}\right] + C. \end{aligned} \tag{6}$$

At this point we will exploit the property that the last sum in (6) is zero for "many" values of s. To see this, fix s and consider

$$\begin{aligned} \sum_{j=1}^{p} t_j^{q-1-s} &= \sum_{j=1}^{p}\exp[i(2j-1)(q-1-s)\pi/p] \\ &= \exp[-i(q-1-s)\pi/p]\sum_{j=1}^{p}\{\exp[i2(q-1-s)\pi/p]\}^j \\ &= \exp[i(q-1-s)\pi/p]\sum_{j=0}^{p-1}\{\exp[i2(q-1-s)\pi/p]\}^j. \end{aligned} \tag{7}$$

This last sum is a finite geometric series with ratio $\exp[i2(q-1-s)\pi/p]$ that sums to zero unless $\exp[i2(q-1-s)\pi/p] = 1$; i.e., unless p divides $(q-1-s)$. In this case the series sums to p. We write $q = bp + r$ where $0 < r < p$, excluding $r = 0$ because p and q are relatively prime. Then p divides $(q-1-s) = (bp + r - 1 - s)$ if and only if $s = r - 1 + \alpha p$ where $\alpha = 0, \dots, b-1$. For such values of s, $q - 1 - s = (b-\alpha)p$. Consequently, (7) reduces to $p(-1)^{b-\alpha}$ so that indexing the second sum in (6) over α yields

$$F_{p/q}(t) = -\frac{q}{p}\left[\sum_{j=1}^{p} t_j^q \log(u-t_j) + (-1)^b p \sum_{\alpha=0}^{b-1}{}' (-1)^\alpha \frac{u^{r+\alpha p}}{r+\alpha p}\right] + C, \tag{8}$$

where the prime indicates that the sum is zero if $b=0$. Since $p=2n$ and

$$t_j^q \log(u-t_j)+\overline{t_j^q}\log(u-\overline{t_j}) = \cos q\theta_j \log(u^2-2u\cos\theta_j+1) + 2\sin q\theta_j \arctan\left(\frac{\sin\theta_j}{u-\cos\theta_j}\right),$$

a little manipulation similar to that used to establish (4) verifies that (8) can be expressed as

$$F_{p/q}(t) = -\frac{q}{p}\sum_{j=1}^{n}\left[\cos q\theta_j \log(t^{2/q}-2t^{1/q}\cos\theta_j+1) - 2\sin q\theta_j \arctan\left(\frac{t^{1/q}-\cos\theta_j}{\sin\theta_j}\right)\right] + (-1)^b q \sum_{\alpha=0}^{b-1}{}' \frac{t^{(r+\alpha p)/q}}{r+\alpha p} + C'. \tag{9}$$

(The changes in the argument of the arctangent function and the constant of integration follow from the identity $\arctan(1/x)=\pi/2-\arctan(x)$.) As expected, if $q=1$ then $b=0$ and (9) reduces to equation (4).

Example 2. Using equation (9) to compute $\int \frac{dt}{t^{2/3}+1}$ we get

$$F_{2/3}(t) = -\frac{3}{2}\left[\cos\frac{3\pi}{2}\log\left(t^{2/3}-2t^{1/3}\cos\frac{\pi}{2}+1\right) + 2\sin\frac{3\pi}{2}\tan^{-1}\left(\frac{t^{1/3}-\cos\frac{\pi}{2}}{\sin\frac{\pi}{2}}\right)\right] + 3t^{1/3} + C$$
$$= 3[t^{1/3}+\tan^{-1}t^{1/3}] + C.$$

In a similar manner, using the m roots of unity, one can calculate $G_m(t) := \int (t^m - 1)^{-1}\,dt$ for any integer m. An alternate approach would be to use the change of variable $t = e^{i\pi/m}s$ to obtain $G_m(t) = -e^{i\pi/m}F_m(-e^{i\pi/m}t)$. A slight modification of this change of variable works for rational values of m.

Decomposing an integrand using the m roots of a given number has many interesting applications. For example, in **[3]** we used a partial fractions decomposition involving the roots of unity to evaluate

$$\int \cot t/\sin\left(\frac{\pi-2t}{2m}\right)dt$$

for m a positive integer. With appropriate hints, this is an interesting exercise for the complex variables student.

Acknowledgement The authors wish to express appreciation to the reviewers for their helpful comments.

Note: This work was done while the authors were on professional leave from The University of Akron.

REFERENCES

1. Gopalan, M. A. and Ravichandran, V., Note on the evaluation of $\int_0^x \frac{1}{1+t^{2^n}}\,dt$, this MAGAZINE (1994), **67**, 53–54.
2. *Notebooks of Srinivasa Ramanujan*, Vol. 2, Tata Institute of Fundamental Research, Bombay, 1957.
3. Palagallo, J. A. and Price, T. E., Near-best approximation by averaging polynomial interpolants, *IMA J. of Numerical Analysis* (1987) **7**, 107–122.

Editor's Note. Professor Yongzhi (Peter) Yang of the University of St. Thomas, St. Paul, Minnesota, independently supplied the following alternative approach to evaluating integrals of the same form as in the preceding article:

The integral

$$I=\int_0^x \frac{1}{1+t^{2m+\delta_{j,1}}}\,dt \qquad (j=0,1)$$

can be evaluated for any integer $m \geq 1$ by the following formula:

$$\int_o^x \frac{1}{1+t^{2m+\delta_{j,1}}}\,dt=\sum_{l=1}^{m}\left(\frac{b_l}{2}\ln\left|x^2+a_lx+1\right|+\frac{(2c_l-a_lb_l)}{\sqrt{4-a_l^2}}\arctan\left(\frac{x\sqrt{4-a_l^2}}{2+a_lx}\right)\right)$$
$$+\frac{\delta_{j,1}}{(2m+1)}\ln|x+1|,$$

where

$$a_l=-2\cos\left(\frac{\pi+(l-1)2\pi}{2m+\delta_{j,1}}\right),$$

$$b_l=\frac{1}{2\,\mathrm{Im}(\alpha_l)i}$$
$$\times\left[\frac{1}{(\delta_{j,1}\alpha_l+1)\prod\limits_{\substack{s=1\\ s\neq l}}^{m}(\alpha_l^2+a_s\alpha_l+1)}-\frac{1}{(\delta_{j,1}\overline{\alpha}_l+1)\prod\limits_{\substack{s=1\\ s\neq l}}^{m}(\overline{\alpha}_l^2+a_s\overline{\alpha}_l+1)}\right],$$

$$c_l=\frac{1}{2\,\mathrm{Im}(\alpha_l)i}$$
$$\times\left[\frac{\overline{\alpha}_l}{(\delta_{j,1}\alpha_l+1)\prod\limits_{\substack{s=1\\ s\neq l}}^{m}(\alpha_l^2+a_s\alpha_l+1)}-\frac{\alpha_1}{(\delta_{j,1}\overline{\alpha}_l+1)\prod\limits_{\substack{s=1\\ s\neq l}}^{m}(\overline{\alpha}_l^2+a_s\alpha\;ay_l+1)}\right],$$

$$\alpha_l=\cos\left(\frac{\pi+(l-1)2\pi}{2m+\delta_{j,1}}\right)+i\sin\left(\frac{\pi+(l-1)2\pi}{2m+\delta_{j,1}}\right),$$

$\overline{\alpha}_i$ is the complex conjugate of α_l, $i=\sqrt{-1}$, $\mathrm{Im}(\cdot)$ denotes the imaginary part, and $\delta_{j,1}$ is the Kronecker delta function.

PROBLEMS

GEORGE T. GILBERT, *Editor*
Texas Christian University

ZE-LI DOU, KEN RICHARDSON, and SUSAN G. STAPLES, *Assistant Editors*
Texas Christian University

Proposals

To be considered for publication, solutions should be received by July 1, 1997.

1514. *Proposed by Hoe Teck Wee, Lengkok Bahru, Singapore.*

Let p be an odd prime and k be a natural number. Find the sum of the elements of the subsets of $\{1, 2, \dots, kp\}$, the sum of whose elements is divisible by p. (For instance, when $p = 3$ and $k = 1$, the relevant subsets are $\{1, 2\}$, $\{3\}$, and $\{1, 2, 3\}$, and the required sum is 12.)

(This generalizes problem 6 of the 36th International Mathematical Olympiad, held in July 1995.)

1515. *Proposed by Isaac Sofair, Fredericksburg, Virginia.*

The edges of a parallelepiped emanating from one vertex are given by the vectors **a**, **b**, and **c**, of lengths a, b, and c, respectively. If α, β, and γ are the angles between **b** and **c**, **c** and **a**, and **a** and **b**, respectively, and $\sigma = (\alpha + \beta + \gamma)/2$, show that the volume of the parallelepiped is

$$2abc\sqrt{\sin \sigma \sin(\sigma - \alpha)\sin(\sigma - \beta)\sin(\sigma - \gamma)}\ .$$

1516. *Proposed by David Doster, Choate Rosemary Hall, Wallingford, Connecticut.*

Let $S_n = \sum_{k=1}^{n} \sqrt{4n^2 - k^2}$. Find the unique value of c for which $\lim_{n \to \infty}(cn - S_n/n)$ exists, and evaluate the limit for this value of c.

We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals must, in general, be accompanied by solutions and by any bibliographical information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution.

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

Solutions and new proposals should be mailed to George T. Gilbert, Problems Editor, Department of Mathematics, Box 298900, Texas Christian University, Fort Worth TX 76129, or mailed electronically (ideally as a LATEX file) to g.gilbert@tcu.edu. Readers who use e-mail should also provide an e-mail address.

1517. *Proposed by Charles Vanden Eynden, Illinois State University, Normal, Illinois.*

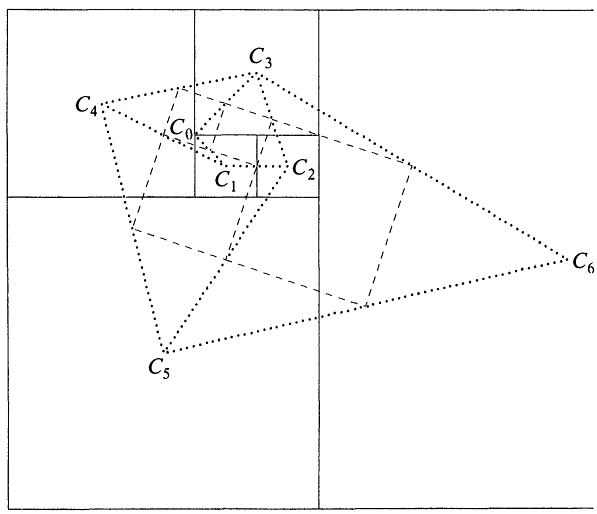
Let M_n be the $n \times n$ matrix with entries the integers from 1 to n^2 spiraling clockwise inwardly, starting in the first row and column. For example

$$M_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 12 & 13 & 14 & 5 \\ 11 & 16 & 15 & 6 \\ 10 & 9 & 8 & 7 \end{pmatrix}.$$

Evaluate the determinant of M_n.

1518. *Proposed by Edward Kitchen, Santa Monica, California.*

Let C_n be the center of a square whose side-length is F_n, $n \geq 0$, where (F_n) is the Fibonacci sequence $0, 1, 1, 2, 3, \ldots$. Place the squares side-by-side in a spiral as in the diagram below. For $n > 0$ join the midpoints of adjacent sides of each quadrangle $C_{n-2}C_{n-1}C_{n+2}C_{n+1}$ (where $C_{-1} = C_1$ by convention). Prove that the resulting pattern is another sequence of squares whose side-lengths are a constant multiple of the Fibonacci sequence.



Quickies

Answers to the Quickies are on page 73.

Q859. *Proposed by Sammy Yu and Jimmy Yu, students, University of South Dakota, Vermillion, South Dakota.*

In ΔABC, prove that

$$\frac{bcA + caB + abC}{bc + ca + ab} \le 60 \le \frac{aA + bB + cC}{a + b + c},$$

where A, B, and C denote the angle measures in degrees, and a, b, and c are the lengths of their respective opposite sides.

Q860. *Proposed by David Bradley, Simon Fraser University, Burnaby, British Columbia, Canada.*

Let z be a complex number with nonzero real part. Evaluate the bilateral sum

$$\sum_{n=-\infty}^{\infty} \frac{1}{\cosh((2n+1)z) + \cosh(z)}.$$

Q861. *Proposed by David Callan, University of Wisconsin, Madison, Wisconsin.*

Suppose the function $f\colon \mathbb{N} \to \mathbb{Z}$ satisfies the property that $m+n$ divides $f(m)+f(n)$ for all $m, n \in \mathbb{N}$ (for example, $f(x) = x^3$). Show that $m-n$ divides $f(m)-f(n)$ for all $m, n \in \mathbb{N}$.

Solutions

Divisibility for Factorials, Least Common Multiples **February 1996**

1489. *Proposed by David Callan, University of Wisconsin, Madison, Wisconsin.*

For what integers k is

$$\frac{m!n!\operatorname{lcm}\{1, 2, \dots, m+n+k\}}{(m+n+k)!}$$

an integer for all nonnegative integers m and n such that $m+n+k > 0$ (lcm denoting the least common multiple)?

Solution by Thomas Jager, Calvin College, Grand Rapids, Michigan.

The condition holds for $k \le 1$. If N is a positive integer and p is a prime number, let N_p denote the exponent of p in the prime factorization of N and let $L(N) = \operatorname{lcm}\{1, 2, \dots, N\}$. First, suppose $k \ge 2$. Then the condition fails for $m = n = 1$ since $(2+k)!$ does not divide $L(2+k)$.

Now, suppose $k = 0$, $m, n \ge 0$, and $m+n > 0$. Then, for p prime and s defined by $p^s \le m+n < p^{s+1}$,

$$\left(\frac{(m+n)!}{m!n!}\right)_p = \sum_{t=1}^{s}\left(\left\lfloor\frac{m+n}{p^t}\right\rfloor - \left\lfloor\frac{m}{p^t}\right\rfloor - \left\lfloor\frac{n}{p^t}\right\rfloor\right) \le s = L(m+n)_p.$$

Next, suppose $k=1$ and $m, n \geq 0$. If the prime p does not divide $n+1$, then

$$(m+n+1)!_p \leq (m!(n+1)!L(m+n+1))_p = (m!n!L(m+n+1))_p$$

by the $k=0$ case. Hence, by symmetry, we may assume $p|(n+1)$ and $p|(m+1)$. Then, p cannot divide $m+n+1$, so that

$$(m+n+1)!_p = (m+n)!_p \leq (m!n!L(m+n))_p \leq (m!n!L(m+n+1))_p.$$

Finally, suppose $k<0$, $m, n \geq 0$, and $m+n+k>0$. Then, we can find $k_1, k_2 \leq 0$ such that $k=k_1+k_2$, $m+k_1 \geq 0$, and $n+k_2 \geq 0$, so that we may apply the $k=0$ case to $m+k_1$ and $n+k_2$. Hence, $(m+n+k)! = (m+k_1+n+k_2)!$ divides $(m+k_1)!(n+k_2)!L(m+n+k)$. Since the latter divides $m!n!L(m+n+k)$, we are done.

Also solved by Merrill Barnebey, Qais Haider Darwish (Oman), Lorraine L. Foster, J. S. Frame, Thelma W. Hedgepeth, Colonel Johnson, Jr., Helen M. Marston, Peter D. Yom, Monte J. Zerger, and the proposer.

A Cat and Mouse Pursuit **February 1996**

1490. *Proposed by Evgenii S. Freidkin, Rutgers-The State University of New Jersey, Piscataway, New Jersey.*

A mouse with maximum speed v_m sits at the center of a regular pentagon. At each vertex of the pentagon is a cat with maximum speed v_c. If the cats must remain on the boundary of the pentagon, find a necessary and sufficient condition which guarantees the mouse can escape from the pentagon. (Assume that the animals are points and that changes in velocity may be instantaneous.)

Solution by Irl C. Bivens and Benjamin G. Klein, Davidson College, Davidson, North Carolina.

The mouse can escape if and only if $v_m/v_c > 1/(2\sin 36°) \approx 0.8507$.

Assume that $v_m/v_c > 1(2\sin 36°)$. In this case the mouse can escape by running at top speed towards a vertex A, "dodging" at the last moment to one of the edges adjacent to A. To prove that such an escape is indeed possible, we will first show that the mouse can avoid capture by the cat originally at A. Let P denote a point between the center O of the pentagon and vertex A such that $OP/OA > \cos 72°$, and assume Q is a point on an edge adjacent to A such that PQ is perpendicular to PA. Then ΔQAP is a $36° - 54° - 90°$ triangle and

$$\frac{PQ}{AQ} = \sin 54° < \frac{\sin 54°}{\sin 72°} = \frac{1}{2\sin 36°} < \frac{v_m}{v_c}.$$

Consequently, the mouse can run from P to Q in less time than it takes a cat to run from A to Q. The mouse can avoid capture by the cat originally at vertex A by making an appropriate 90 degree turn to an edge adjacent to A. (If the cat originally at A is at A when the mouse makes its turn, then the mouse runs to either of the two edges adjacent to A. Otherwise, the mouse runs to the edge this cat is not on at the moment the mouse turns.) Furthermore, since $v_m/v_c > 1/(2\sin 36°)$, the mouse can run from O to A in less time than it takes a cat to run the length of an edge. Thus, it is clear that by choosing its turning point P sufficiently close to A, the mouse can run a path from O to an edge adjacent to A before a cat at any other vertex of the pentagon can reach the mouse. Therefore, the mouse can escape.

Next, assume that $v_m/v_c \leq 1/(2\sin 36°)$, and note that then $v_m < v_c$. The cats can now keep the mouse trapped inside the pentagon. It suffices to show that a cat at vertex A can prevent the mouse from exiting through edge AB. The cat's strategy is to immediately begin running to vertex B with speed v_c until it reaches either the foot S of the perpendicular through the mouse's position to the line through A and B or the vertex B. Should the cat's position and S happen to coincide before the cat reaches B, then from that moment on the cat moves in such a way that its position is given by the varying S if possible, and the closer of A and B to S if not. (This is possible since changes in velocity are assumed to be instantaneous and because the speed of S on the line through A and B is at most $v_m < v_c$.) The same strategy holds from the moment the cat reaches B should this occur. In either case we conclude that once the cat's position is S or B, it will be impossible for the mouse to exit through edge AB. For should the mouse contact this edge, the mouse's position and S, and thus also the position of the cat, will coincide. Consequently, the only chance the mouse has of exiting through AB is for the mouse to exit through a point Q in front of the advancing cat. But this is easily seen to be impossible. Suppose Q is any point on AB. Then in ΔAQO, $\angle A = 54°$ and $\angle O \leq 72°$. By the law of sines,

$$\frac{OQ}{AQ} = \frac{\sin 54°}{\sin \angle O} \geq \frac{\sin 54°}{\sin 72°} = \frac{1}{2\sin 36°} \geq \frac{v_m}{v_c}.$$

Since the mouse must run at least a distance of OQ to exit AB at Q, it cannot reach Q before the advancing cat. Consequently, the mouse is unable to exit the pentagon through edge AB. Since a similar argument applies to the other four edges as well, the mouse is trapped inside the pentagon.

Note that the initial placement of cats is a very important factor. Were the cats initially at the midpoints of the edges of the pentagon, then the mouse cannot escape if $v_m/v_c \leq 1$.

Also solved by Western Maryland College Problems Group and the proposer. There were four incorrect solutions.

Inversions Satisfying a Functional Equation **February 1996**

1491. *Proposed by Wu Wei Chao, He Nan Normal University, Xin Xiang City, He Nan Province, China.*

Find all functions $f\colon \mathbb{R} \to \mathbb{R}$ such that
(i) $f(x+f(y)+yf(x)) = y+f(x)+xf(y)$ for all x, y in $\mathbb{R}$;
(ii) $\{f(x)/x\colon x \in \mathbb{R},\ x \neq 0\}$ is a finite set.

Solution by Robert L. Doucette, McNeese State University, Lake Charles, Louisiana.

The function f must be the identity function on $\mathbb{R}$.

By (ii) there is some real number λ such that $S_\lambda := \{x\colon f(x) = \lambda x\}$ contains infinitely many elements. By (i),

$$f(x+\lambda x+\lambda x^2) = f(x+f(x)+xf(x)) = x+\lambda x+\lambda x^2$$

for all $x \in S_\lambda$. It follows that S_1 has infinitely many elements.

Suppose that there is some $y \neq 0$ such that $\mu := f(y)/y \neq 1$. Condition (i) implies that for all $x \in S_1$,

$$f(x+\mu y+yx) = f(x+f(y)+yf(x)) = y+f(x)+xf(y) = y+x+x\mu y.$$

By (ii) the function

$$\varphi\colon S_1 - \left\{ -\frac{\mu y}{y+1} \right\} \to \mathbb{R},$$

$$x \mapsto \frac{f(x+\mu y + yx)}{x+\mu y + yx} = \frac{(1+\mu y)x + y}{(1+y)x+\mu y},$$

takes on only a finite number of values. This is impossible since φ is one-to-one and has infinitely many elements in its domain. It follows that $f(x) = x$ for $x \neq 0$.

By choosing $x \neq 0$, $y \neq 0$ such that $x + y + xy = 0$, we see by (i) that $f(0) = 0$.

Comments. The above solution applies to any infinite field F. Thomas Jager points out that, whenever F is uncountable, (ii) may be relaxed to the condition that $\{f(x)/x\colon x \in F,\ x \neq 0\}$ is a countable set. What functions on $\mathbb{R}$ satisfy (i)? One might note that complex conjugation satisfies (i) over the complex numbers. The editors will consider publishing an example of a function other than the identity that satisfies (i) over $\mathbb{R}$, or a proof that none exists.

Also solved by Bilkent Mathematical Society Problem Solving Group (Turkey), Con Amore Problem Group (Denmark), Tim Flood, Jennifer Hyndman (Canada), Thomas Jager, Kee-Wai Lau (Hong Kong), Can A. Minh (student), Kandasamy Muthuvel, Joel Schlosberg (student), Achilleas Sinefakopoulos (student, Greece), Western Maryland College Problems Group, Peter D. Yom, Paul J. Zwier, and the proposer.

A Laplace Transform **February 1996**

1492. *Proposed by Michael Golomb, Purdue University, West Lafayette, Indiana.*

Derive the Laplace transform of the function $(2\sin x - \sin 2x)/x^2$; i.e., for $\mathrm{Re}\, s > 0$ evaluate the integral

$$\int_0^\infty e^{-sx} \frac{2\sin x - \sin 2x}{x^2}\, dx.$$

I. *Solution by J. S. Frame, Michigan State University, East Lansing, Michigan.*

For s a positive real number, the second derivative $F''(s)$ of this transform integral $F(s)$ is

$$F''(s) = \int_0^\infty e^{-sx}\, 2\sin x\, dx - \frac{1}{2}\int_0^\infty e^{-(s/2)t} \sin t\, dt = \frac{2}{s^2+1} - \frac{2}{s^2+4},$$

where $t = 2x$ in the second integral. Since $\lim_{s\to\infty} F'(s) = \lim_{s\to\infty} F(s) = 0$, we have, upon integrating from s to ∞,

$$\begin{aligned}
F'(s) &= 2\arctan s - \arctan\left(\frac{s}{2}\right) - \frac{\pi}{2},\\
F(s) &= \left(2s\arctan s - \ln(s^2+1)\right) - \left(s\arctan\left(\frac{s}{2}\right) - \ln(s^2+4)\right) - \frac{\pi}{2}s\\
&= s\arctan\left(\frac{2}{s}\right) - 2s\arctan\left(\frac{1}{s}\right) + \ln\left(\frac{s^2+4}{s^2+1}\right)\\
&= \ln\left(\frac{s^2+4}{s^2+1}\right) - s\arctan\left(\frac{2}{s^3+3s}\right).
\end{aligned}$$

Analytic continuation yields the result for $\mathrm{Re}\, s > 0$.

II. *Solution by Benjamin G. Klein, Davidson College, Davidson, North Carolina.*

In the following, we assume $s > 2$ to ensure convergence of all of the series that appear. We have

$$\int_0^\infty e^{-sx}\frac{2\sin(x)-\sin(2x)}{x^2}\,dx = \int_0^\infty e^{-sx}\sum_{k=1}^\infty \frac{(-1)^k(2-2^{2k+1})x^{2k-1}}{(2k+1)!}\,dx$$
$$= \sum_{k=1}^\infty \left(\int_0^\infty e^{-sx}x^{2k-1}\,dx\right)\frac{(-1)^k(2-2^{2k+1})}{(2k+1)!}$$
$$= \sum_{k=1}^\infty \left(\frac{(2k-1)!}{s^{2k}}\right)\frac{(-1)^k(2-2^{2k+1})}{(2k+1)!}$$
$$= \sum_{k=1}^\infty \frac{(-1)^k(2-2^{2k+1})}{(2k)(2k+1)s^{2k}}$$
$$= \sum_{k=1}^\infty \frac{(-1)^k(2-2^{2k+1})}{s^{2k}}\left(\frac{1}{2k}-\frac{1}{2k+1}\right).$$

This last series can be written as

$$\sum_{k=1}^\infty \frac{\left(\frac{-1}{s^2}\right)^k}{k} - \sum_{k=1}^\infty \frac{\left(\frac{-4}{s^2}\right)^k}{k} - 2s\sum_{k=0}^\infty \frac{(-1)^k\left(\frac{1}{s}\right)^{2k+1}}{2k+1} + s\sum_{k=0}^\infty \frac{(-1)^k\left(\frac{2}{s}\right)^{2k+1}}{2k+1}$$
$$= -\ln\left(1+\frac{1}{s^2}\right) + \ln\left(1+\frac{4}{s^2}\right) - 2s\arctan\left(\frac{1}{s}\right) + s\arctan\left(\frac{2}{s}\right)$$
$$= \ln\left(\frac{4+s^2}{1+s^2}\right) - 2s\arctan\left(\frac{1}{s}\right) + s\arctan\left(\frac{2}{s}\right).$$

Analytic continuation yields the result for $\operatorname{Re} s > 0$.

Also solved by Wayne Dennis and Edward D. Onstott, Jayanthi Ganapathy, M. L. Glasser, Paul R. Johnson, Peter A. Lindstrom, O. P. Lossers (The Netherlands), Felix Magnotta, Kim McInturff, Can A. Minh (student), Darryl K. Nester, William A. Newcomb, Jeremy Ottenstein (Israel), Fary Sami and Reza Akhlaghi, Nora S. Thornber, Robert J. Wagner, Harry Weingarten, Joseph Wiener, and the proposer.

A Generalization of Napoleon's Theorem **February 1996**

1493. *Proposed by Jiro Fukuta, Gifu-ken, Japan.*

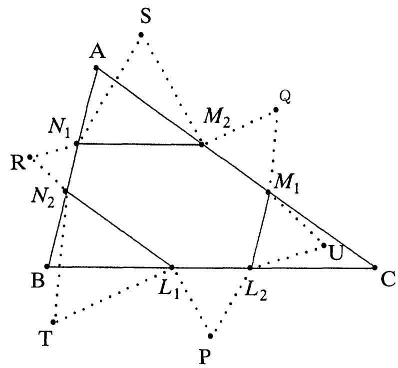
In ΔABC, let L_1 and L_2, M_1 and M_2, N_1 and N_2 be distinct points on the sides BC, CA, AB, respectively, such that

$$\frac{BL_1}{L_1C} = \frac{CL_2}{L_2B} = \frac{CM_1}{M_1A} = \frac{AM_2}{M_2C} = \frac{AN_1}{N_1B} = \frac{BN_2}{N_2A} < 1.$$

Let PL_1L_2, QM_1M_2, RN_1N_2, SM_2N_1, TN_2L_1, and UL_2M_1 be the equilateral triangles built outwards on the sides of the hexagon $L_1L_2M_1M_2N_1N_2$.

(i) Prove that the segments PS, QT, and RU have equal lengths, and the lines PS, QT, and RU intersect at 120° angles and are concurrent.

(ii) If G_1, G_2, G_3, G_4, G_5, and G_6 are the centroids of triangles QSR, SRT, RTP, TPU, PUQ, and UQS, respectively, prove that $G_1G_2G_3G_4G_5G_6$ is a regular hexagon whose centroid coincides with that of ΔABC.



Solution by O. P. Lossers, Technical University Eindhoven, Eindhoven, The Netherlands.

We represent the points by complex numbers and assume $A+B+C=0$. Let $\rho = L_1C/BC$ and $\sigma = BL_1/BC$, and let $\varepsilon = \exp(2\pi i/3)$.

The points L_1, L_2, M_1, M_2, N_1, and N_2 satisfy

$$L_1 = \rho B + \sigma C,\ M_1 = \rho C + \sigma A,\ N_1 = \rho A + \sigma B,$$
$$L_2 = \sigma B + \rho C,\ M_2 = \sigma C + \rho A,\ N_2 = \sigma A + \rho B.$$

For a positively-oriented equilateral triangle VWX, we have $V - X = \varepsilon(X - W)$, or $V = -\varepsilon W - \varepsilon^2 X$, hence

$$P = -\varepsilon L_2 - \varepsilon^2 L_1 = -(\varepsilon\sigma + \varepsilon^2\rho)B - (\varepsilon\rho + \varepsilon^2\sigma)C,$$
$$Q = -\varepsilon M_2 - \varepsilon^2 M_1 = -(\varepsilon\sigma + \varepsilon^2\rho)C - (\varepsilon\rho + \varepsilon^2\sigma)A,$$
$$R = -\varepsilon N_2 - \varepsilon^2 N_1 = -(\varepsilon\sigma + \varepsilon^2\rho)A - (\varepsilon\rho + \varepsilon^2\sigma)B,$$
$$S = -\varepsilon N_1 - \varepsilon^2 M_2 = \rho A - \varepsilon\sigma B - \varepsilon^2\sigma C,$$
$$T = -\varepsilon L_1 - \varepsilon^2 N_2 = \rho B - \varepsilon\sigma C - \varepsilon^2\sigma A,$$
$$U = -\varepsilon M_1 - \varepsilon^2 L_2 = \rho C - \varepsilon\sigma A - \varepsilon^2\sigma B.$$

We find

$$S - P = \rho(A + \varepsilon^2 B + \varepsilon C),$$
$$T - Q = \rho(B + \varepsilon^2 C + \varepsilon A) = \varepsilon(S - P),$$
$$U - R = \rho(C + \varepsilon^2 A + \varepsilon B) = \varepsilon^2(S - P).$$

This proves $|S-P| = |T-Q| = |U-R|$ and

$$\frac{T-Q}{S-P} = \frac{U-R}{T-Q} = \frac{S-P}{U-R} = \varepsilon,$$

i.e., the first two assertions of (i).

Write the equation of the line passing through two points V and W in the form

$$(\overline{V} - \overline{W})Z - (V - W)\overline{Z} + V\overline{W} - W\overline{V} = 0.$$

Since the three lines passing through P and S, Q and T, and U and R are clearly not parallel, the concurrency assertion of (i) is equivalent to the vanishing of the determinant

$$\begin{vmatrix} \overline{S}-\overline{P} & P-S & P\overline{S}-S\overline{P} \\ \overline{T}-\overline{Q} & Q-T & Q\overline{T}-T\overline{Q} \\ \overline{U}-\overline{R} & R-U & R\overline{U}-U\overline{R} \end{vmatrix}.$$

Setting $Y = S - P$, this determinant may be written as

$$-Y\overline{Y}\begin{vmatrix} 1 & 1 & P\overline{S}-S\overline{P} \\ \varepsilon^2 & \varepsilon & Q\overline{T}-T\overline{Q} \\ \varepsilon & \varepsilon^2 & R\overline{U}-U\overline{R} \end{vmatrix} = Y\overline{Y}\begin{vmatrix} 1 & 1 & Y\overline{S}-\overline{Y}S \\ \varepsilon^2 & \varepsilon & \varepsilon Y\overline{T}-\varepsilon^2\overline{Y}T \\ \varepsilon & \varepsilon^2 & \varepsilon^2 Y\overline{U}-\varepsilon\overline{Y}U \end{vmatrix}.$$

Writing S, T, and U in terms of A, B, and C, and using basic properties of the determinant, this last determinant simplifies to

$$Y\overline{Y}\begin{vmatrix} 1 & 1 & Y\overline{A}-\overline{Y}A \\ \varepsilon^2 & \varepsilon & Y\overline{B}-\overline{Y}B \\ \varepsilon & \varepsilon^2 & Y\overline{C}-\overline{Y}C \end{vmatrix}.$$

Since $A + B + C = 0$, the product of $(1,1,1)$ and this last matrix is 0, hence the determinant is 0 and the three lines are concurrent.

We now compute

$$G_1 = \frac{Q+S+R}{3} = \frac{M_2+N_1-\varepsilon^2 M_1-\varepsilon N_2}{3},$$

$$G_2 = \frac{S+R+T}{3} = \frac{N_1+N_2-\varepsilon^2 M_2-\varepsilon L_1}{3},$$

$$G_3 = \frac{R+T+P}{3} = \frac{N_2+L_1-\varepsilon^2 N_1-\varepsilon L_2}{3}.$$

It follows immediately from $A + B + C = 0$ that

$$L_1 + M_1 + N_1 = 0 \quad \text{and} \quad L_2 + M_2 + N_2 = 0.$$

Then

$$\varepsilon^2 G_1 + G_2 = \frac{-\varepsilon(L_1+M_1+N_1)}{3} = 0,$$

$$\varepsilon^2 G_2 + G_3 = \frac{-\varepsilon(L_2+M_2+N_2)}{3} = 0.$$

These and similar computations imply that

$$|G_1| = |G_2| = |G_3| = |G_4| = |G_5| = |G_6|$$

and that

$$G_{i+1}/G_i = -\varepsilon^2 \text{ for } i = 1,\dots,5 \; (\text{and } G_1/G_6 = -\varepsilon^2).$$

Therefore, $G_1G_2G_3G_4G_5G_6$ is a regular hexagon centered at the origin, which is also the centroid of ΔABC. This proves (ii).

The proofs are valid as well if $\varepsilon = \exp(-2\pi i)/3$, i.e., if the equilateral triangles are drawn inwards.

Also solved by Con Amore Problem Group (Denmark), Victor Kutsenok, Wilfred Reyes (Chile), Hoe-Teck Wee (Singapore), Robert L. Young, and the proposer.

Answers

Solutions to the Quickies on page 65.

A859. Write

$$aA = \frac{2}{3}aA + \frac{1}{3}a(180 - B - C) = 60a + \frac{2}{3}aA - \frac{1}{3}aB - \frac{1}{3}aC,$$

and so forth. Because larger angles in a triangle occur opposite larger sides, the above yields

$$\begin{aligned} &aA + bB + cC \\ &\quad = 60(a+b+c) + \frac{1}{3}[(a-b)(A-B) + (b-c)(B-C) + (c-a)(C-A)] \\ &\quad \geq 60(a+b+c), \end{aligned}$$

with equality if and only if ΔABC is equilateral. Dividing by $a+b+c$ proves the second inequality. Similarly,

$$bcA = 60bc + \frac{2}{3}bcA - \frac{1}{3}bcB - \frac{1}{3}bcC,$$

and so forth, from which it follows that

$$\begin{aligned} &bcA + caB + abC \\ &\quad = 60(bc + ca + ab) - \frac{1}{3}[c(a-b)(A-B) + a(b-c)(B-C) + b(c-a)(C-A)] \\ &\quad \geq 60(bc + ca + ab) \end{aligned}$$

(again with equality if and only if ΔABC is equilateral). Dividing by $bc + ca + ab$ demonstrates the first inequality and completes the proof.

A860. Let

$$S(z) = \sum_{n=-\infty}^{\infty} \frac{1}{\cosh((2n+1)z) + \cosh(z)}.$$

If we put $q = e^z$ and factor, we obtain

$$S(z) = \sum_{n=-\infty}^{\infty} \frac{2}{(q^{n+1} + q^{-n-1})(q^n + q^{-n})} = \sum_{n=-\infty}^{\infty} \frac{2q^{2n+1}}{(q^{2n+2}+1)(q^{2n}+1)}.$$

By the method of partial fractions, we rewrite the previous expression in the form

$$S(z)=\frac{2q}{1-q^2}\sum_{n=-\infty}^{\infty}\left(\frac{1}{q^{2n+2}+1}-\frac{1}{q^{2n}+1}\right),$$

which telescopes, leaving us with

$$\begin{aligned}S(z)&=\frac{2q}{1-q^2}\left(\lim_{n\to\infty}\frac{1}{q^{2n+2}+1}-\lim_{m\to\infty}\frac{1}{q^{-2m}+1}\right)\\&=\begin{cases}\dfrac{2q}{1-q^2}, & |q|<1\\[2ex] \dfrac{2q}{q^2-1}, & |q|>1\end{cases}\\&=\frac{\operatorname{sgn}(\operatorname{Re}(z))}{\sinh(z)}.\end{aligned}$$

A861. It is clearly sufficient to show that k divides $f(l+k)-f(l)$ for all $k,l\in\mathbb{N}$. Let r be the remainder when l is divided by k and apply the hypothesis with $m=l$ and $n=k-r$ to obtain that k divides $f(l)+f(k-r)$. Apply it a second time with $m=l+k$ and $n=k-r$ to obtain that k divides $f(l+k)+f(k-r)$. Hence k divides the difference $f(l+k)-f(l)$, as desired.

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 mathematics literature. Readers are invited to suggest items for review to the editors.

Mathematical Digest: Short Summaries of Articles about Mathematics in the Popular Press. `http://www.ams.org/new-in-math/mathdigest/` .

This useful Web page from AMS's e-Math contains links to one-paragraph summaries of two or three articles per month from the popular press, including *Discover*, *Science*, *Wall Street Journal*, *New York Times*, *American Scientist*, *Scientific American*, *The Scientist*, *The Sciences*, and *Physics Today*. The paragraphs are unattributed to authors. Entries seem to be almost two months behind current—so maybe there is still room for this column of reviews, which tends to be two to four months behind current when an issue reaches you.

Stewart, Ian, Mathematical recreations, Monopoly revisited, *Scientific American* 273 (4) (October 1996) 116–119.

Stewart's readers refine his earlier modeling of the game of Monopoly as a Markov chain, taking fuller account of the rules of the game. Stewart does not mention the analysis by Robert B. Ash and Richard L. Bishop ("Monopoly as a Markov process," THIS MAGAZINE 45 (1972) 26–29). Their frequencies differ from Stewart's in part because of different modeling assumptions but also because they take a throw of the dice as the basic transition (and a turn may consist of as many as three throws).

Borwein, J., P. Borwein, R. Girgensohn, and S. Parnes, Making sense of experimental mathematics, *Mathematical Intelligencer* 18 (4) (Fall 1996) 12–18.

In 1993 Enrico Au-Yeung, an undergraduate at the University of Waterloo, observed that

$$\sum_{k=1}^{\infty}\left(1+\frac{1}{2}+\cdots+\frac{1}{k}\right)^2 k^{-2} = 4.599987\ldots \approx \frac{17}{4}\zeta(4) = \frac{17\pi^4}{360}.$$

The authors used this discovery as a springboard to search for and prove similar identities; in this article, they discuss the nature and place of such experimental mathematics

Hargittai, István, Lifelong symmetry: A conversation with H.S.M. Coxeter, *Mathematical Intelligencer* 18 (4) (Fall 1996) 35–41. Coxeter, H.S.M., The trigonometry of Escher's woodcut "Circle Limit III," 42–46.

This interview is a celebration of H.S.M. Coxeter, who will turn 90 in 1997. In it, he mentions his recent investigation of what appear to be troubling inconsistencies in Escher's *Circle Limit III*, which is based on hyperbolic geometry. What Coxeter observed is that arcs of the print meet the circumference not at angles of 90° (which would make them lines of the geometry) but at angles very close to 80°, so that they are branches of equidistant curves. In fact, Escher did the drawings with extraordinary accuracy, agreeing to 0.1 cm with calculated values; the calculated angle is $\cos^{-1}\frac{1}{2}\left(2^{1/4}-2^{-1/4}\right) \approx 79^\circ 58'$.

Konhauser, Joseph D.E., Dan Velleman, and Stan Wagon, *Which Way Did the Bicycle Go ... and Other Intriguing Mathematical Mysteries*, MAA, 1996; xvi + 237 pp, $24.95. ISBN 0-88385-325-6.

For almost 30 years, Joe Konhauser and successor Stan Wagon have posted a problem of the week at Macalester College. This book collects 190 of those, with solutions. "[T]he problems had to involve almost no prerequisites and be succinctly stated and inherently attractive." The problems are arranged by topic: plane geometry, algebra, combinatorics and graph theory, three-dimensional geometry, and miscellaneous. Calculus appears in the solutions of three problems. The shortest problem statement is: "Find the largest regular hexagon that fits into a square of side-length 2."

Reid, Constance, *Julia: A Life in Mathematics*, MAA, 1996; xii + 124 pp, $27. ISBN 0-88385-520-8.

Constance Reid, well-known for her biographies of mathematicians Hilbert, Courant, and Neyman, here features her sister, Julia Bowman Robinson, a distinguished mathematician who later became President of the American Mathematical Society. The book reprints Reid's "Autobiography of Julia Robinson" (two-thirds of the book), an article by Lisl Gaal on Robinson's dissertation, and articles by Martin Davis and Yuri Matijasevich on their collaborations with Robinson. Included are some previously unpublished photographs and memorabilia. This attractive and engaging book is not only for mathematicians but for mathematics teachers, students, and even nonmathematicians.

Aczel, Amir D., *Fermat's Last Theorem: Unlocking the Secret of an Ancient Mathematical Problem*, Four Walls Eight Windows. 1996; xi + 147 pp, $18. ISBN 1-56858-077-0.

This is a very readable account of the setting, development, and history of the proof of Fermat's Last Theorem, drawing the reader into much other history of mathematics. "This book tells the entire story, ... a story of deception, intrigue, and betrayal." Aczel is not referring to the Pythagorean who let out the secret of irrationality of $\sqrt{2}$, nor to Cardano's betraying Tartaglia's secret about how to solve a cubic equation (both of which somehow find their way into the book), but to the history of the Shimura-Taniyama conjecture (which Wiles proved, and which implies FLT). Aczel's main sources are E.T. Bell's dubious *Men of Mathematics*, Kenneth A. Ribet and Brian Hayes's outstanding "Fermat's Last Theorem and modern arithmetic" (*American Scientist* 82 (1994) 144–156), and Serge Lang's polemic "Some history of the Shimura-Taniyama conjecture" (*Notices of the American Mathematical Society* 42 (11) (November 1995) 1301–1307). A surprising assertion: that the tomb of Archimedes was rediscovered in 1963 (p. 31). As popularization rather than absolute history, this is an inspiring book.

Green, David (ed.), *Teaching Statistics at Its Best*, The Teaching Statistics Trust, Dept. of Probability and Statistics, Univ. of Sheffield, S3 7RH, England; iv + 166 pp, (P). ISBN 0-946554-08-0.

As a subscriber to the journal *Teaching Statistics*, I had seen all these article before, as they are selected from Vols. 6–14. But leafing through the volume, I rediscovered more than half a dozen articles that may be useful to me in teaching statistics this spring.

ERRATUM: In Vol. 69, No. 4 (October 1996), p. 313, this column passed on the assertion that there are no known Perrin pseudoprimes—composites n such that n divides $A(n)$, where $A(0) = 3$, $A(1) = 0$, $A(2) = 2$, and $A(n+1) = A(n-1) + A(n-2)$ for $n \geq 3$. In fact, however, many Perrin pseudoprimes are known (e.g., 271,441, and a further 54 Perrin pseudoprimes less than one billion). (Thanks to John P. Robertson, Berwyn, PA.)

Brozan, Nadine, Jeff Bridges and the math professor, *New York Times* (City Edition) (13 November 1996) B5.

The recently-released movie "The Mirror Has Two Faces," which stars Barbra Streisand and Jeff Bridges, features Bridges as a Columbia University mathematics professor. In fact, Prof. Henry Pinkham of Columbia coached Bridges and made sure that all the math that he did was real. But while Bridges' students are seen yawning in class, Pinkham says that at Columbia, "We know how to make it interesting and how to tell a story."

Kolata, Gina, With major math proof, brute computers show flash of reasoning power, *New York Times* (Ntl. Ed.) (10 December 1996) B5, B10; (City Ed.) C1, C8. Robbins Algebras Are Boolean, `http://www.mcs.anl.gov/home/mccune/ar/robbins/` .

At last, a computer has proved a mathematical theorem that had frustrated mathematicians for decades. The theorem, in the field of equational algebras, was conjectured by Herbert Robbins (Rutgers) in the 1930s; it asserts that a particular set of three equations is equivalent to a boolean algebra. In particular, an algebra satisfying commutativity, associativity, and the *Robbins condition* $n(n(x+y)+n(x+n(y)))=x$ is called a *Robbins algebra*; the new result is that every Robbins algebra is a boolean algebra. The proof proceeded by deriving the *Huntington equation* $n(n(x)+y)+n(n(x)+n(y))=x$; it had been known since 1933 that commutativity, associativity, and the Huntington equation are equivalent to standard axioms for a boolean algebra. The computer proof, obtained by William McCune, took place at the Argonne National Laboratory in Illinois, where Larry Wos supervises a computer reasoning project. The proof will be published in the *Journal of Automated Reasoning* and be available on the Internet.

Gadbois, Steve, Poker with wild cards—A paradox, THIS MAGAZINE 69 (4) (October 1996) 283–285. Emert, John, and Dale Umbach, Inconsistencies in "wild-card" poker, *Chance* 9 (3) (Summer 1996) 17–22. Wild cards in poker make the game less challenging by far, Richard Harris, All Things Considered, National Public Radio, 29 November 1996, `http://www.prognet.com/rafiles/npr/password/nc6n2901-8.ram` (4:45 audio; requires RealAudio, available free at) (tape and transcript available from (888)–NPR–NEWS).

Steve Gadbois's article in THIS MAGAZINE about wild cards in poker games attracted the attention of National Public Radio, which ran a short segment interviewing Gadbois and independent discoverer Emert. The gist of their findings: When wild cards are included but the usual order of hands is maintained, the hand with smallest probability does not always win the game.

Barwise, Jon, and Lawrence Moss, *Vicious Circles*, Center for the Study of Language and Information (distributed by Cambridge University Press), 1996; x + 390 pp, \$49.95, \$24.95 (P). ISBN 1–57586–009–0, 1–57586–008–2.

Bertrand Russell banned self-referential paradoxes from logic and mathematics because of their circularity. Banned they have stayed; the Axiom of Foundation in Zermelo-Fraenkel set theory forces the relation of set membership to be well-founded. But what kind of set theory results if we abandon this axiom? The authors do just that, replacing it by their Anti-Foundation Axiom and proving that the resulting system of "hypersets" is consistent if ZFC is. They use it to model streams of symbols and some kinds of graphs, streamline modal logic, resolve paradoxes about metagames, and illuminate self-referential paradoxes. Despite their efforts to minimize abstraction, this is a book for mathematicians and logicians; others will have to wait for a popularization.

NEWS AND LETTERS

J. Arthur Seebach

J. Arthur Seebach, Jr., editor of *Mathematics Magazine* from 1976 through 1980, died on Dec. 3, 1996 after a long struggle with complications of diabetes. He was 58 years old and had been a member of the Department of Mathematics at St. Olaf College for 30 years.

Arthur was known to many mathematicians through his extensive contributions to the publication program of the Association, a culmination of his life-long fascination with language and mathematics that has roots in his twin undergraduate majors: mathematics and Greek. In addition to serving as editor of this *Magazine*, he served for fifteen years as Associate Editor of the *American Mathematical Monthly*, beginning in 1971 as Reviews editor, and concluding in 1986 as editor of the Mathematical Notes section. He served as a member and chair of prize committees for the Allendoerfer Award and the MAA Book Prize, as a member of the Board of Governors, the Committee on Publications, and as co-Chair of the Advisory Committee for FOCUS during its initial years (1981–85). For several years Arthur was a member of the writing committee for the Graduate Record Examination in mathematics, serving as chair in 1986 and 1987.

In addition to his role in the MAA publications program, Arthur was an active early advocate for hands-on computing, beginning by building his own computers from Heathkits. Even before the IBM PC and Apple Macintosh came on the market, Arthur helped prod St. Olaf to make plans for a campus-wide network of microcomputers, and served as newsletter editor for CIMSE, the special interest group on Computers in Mathematics and Science Education. From 1983 through 1987 he served on the MAA Committee on Computers in Mathematics Education.

In addition to mathematics, editing, and computing, Arthur maintained a very active side business in antique Studebakers—stocking and reselling parts, editing a national newsletter, and maintaining several cars of his own. He also fulfilled a long-standing interest in music by singing regularly with the Bach Society of Minnesota. All these interests, and more, he connected to mathematics through his passion for seeing patterns in the most unlikely places.

Beginning in 1965 when we both joined the St. Olaf faculty (at that time a department of five, now over twenty), Arthur and I worked together as co- authors, co-editors, and co-conspirators on many projects. The first was a series of NSF-sponsored summer undergraduate research projects that led to the publication of *Counterexamples in Topology*. Arthur always viewed this monograph as itself a counterexample to the view, widespread at the time, that undergraduates could neither do nor even contribute to research in mathematics. Now, of course, MAA and AMS have joint committees and program activities specifically designed to recognize and promote research activities by undergraduates.

By happy coincidence, Arthur and I published a joint article in the *Monthly* at about the same time as Kenneth O. May, who invented the Telegraphic Reviews in 1965, decided to find a successor. Knowing the effort it took to write, single handedly (and without word processors) 2000 reviews in four years, May and *Monthly* Editor Harley Flanders asked Arthur and me to set up a reviewing service among the many mathematicians at the two colleges in Northfield, Minnesota. This system is still in place, now overseen by Arnold Ostebee, and includes nearly fifty mathematicians at St. Olaf, Carleton and Macalester Colleges.

In 1975, the MAA asked Arthur and me to serve as co-editors of *Mathematics Magazine*, which the Association had acquired in 1961 from its long-time editor, UCLA mathematician Glenn James. To increase the appeal of this little-known publication, we redesigned the format and, with much trepidation, replaced the table of contents on the front cover with eye- catching (but not very professional) artwork–often student-drawn cartoons. Arthur's sense of whimsy, his love of puns, and his proclivity for obscure connections totally transformed the visual image of *Mathematics Magazine*. Cover art, viewed as radical at the time, has since been emulated by other MAA journals, now even by the *Notices* of the AMS.

Another innovation in *Mathematics Magazine*—a few pages of news and announcements (which in September 1976 included the first published exposition of the new computer-based proof of the four color theorem)—caught the eye of Ed Beckenbach, then chair of the MAA Publications Committee. With Ed's urging, the Association began a newsletter, FOCUS, and asked Arthur and me to serve as co-chairs of the Advisory Committee during its initial years. Arthur's experience as editor and publisher of Studebaker newsletters provided much-needed grounding in reality as new editor Marcia Sward began the daunting task of creating a publication *ex nihilo*. Even then, fifteen years before it became feasible, Arthur dreamed of online electronic distribution. It is indeed a pity that his eyesight failed before he could enjoy MAA Online.

Lynn Arthur Steen
December, 1996

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

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