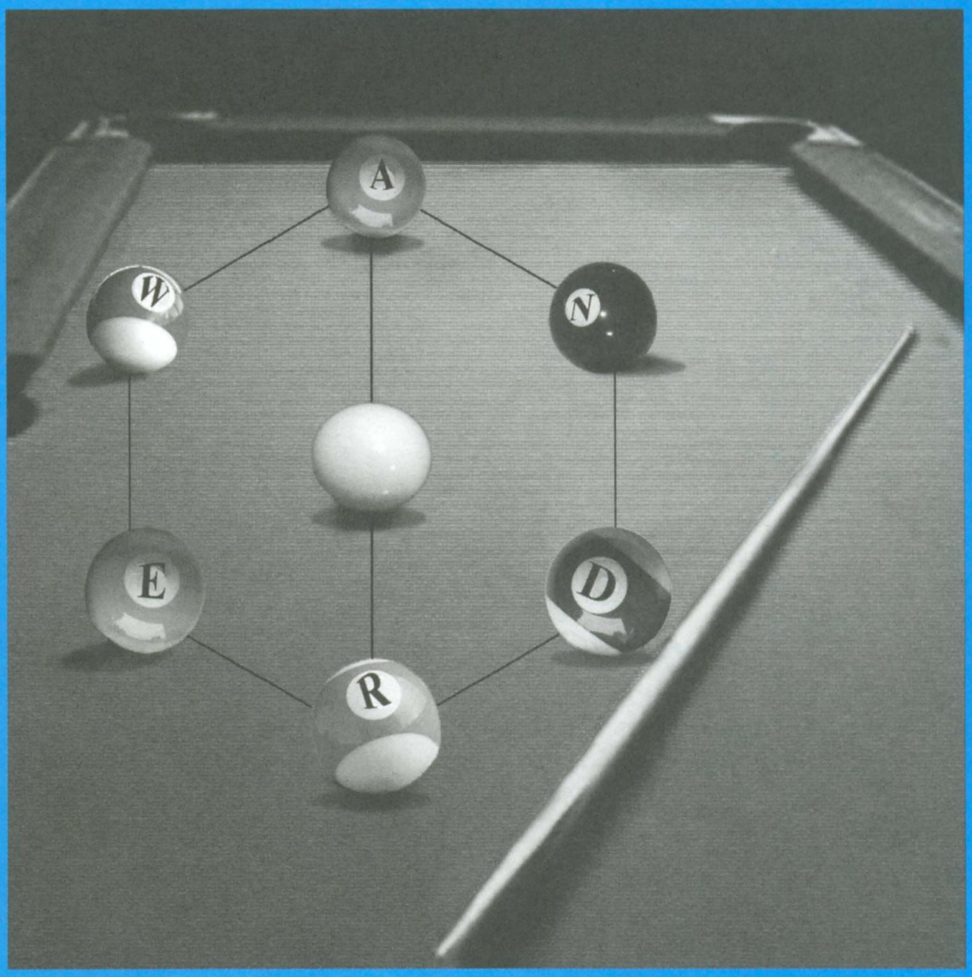
Vol. 82, No. 2, April 2009

MATHEMATICS MAGAZINE



Rick's Tricky Six Puzzle

- Rick's Tricky Six Puzzle: S_5 Sits Specially in S_6
- The Geometry behind Paradoxes of Voting Power
- Counting on Chebyshev Polynomials

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

EDITORIAL POLICY

Mathematics Magazine aims to provide lively and appealing mathematical exposition. The *Magazine* is not a research journal, so the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate 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 among various branches of mathematics and between mathematics and other disciplines.

A more detailed statement of author guidelines appears in this *Magazine*, Vol. 74, pp. 75–76, and is available from the Editor or at www.maa.org/pubs/mathmag.html. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Please submit new manuscripts by email to Editor-Elect Walter Stromquist at mathmag@maa.org. A brief message with an attached PDF file is preferred. Word-processor and DVI files can also be considered. Alternatively, manuscripts may be mailed to Mathematics Magazine, 132 Bodine Rd., Berwyn, PA 19312-1027. If possible, please include an email address for further correspondence.

Cover image: *Rick's Tricky Six Puzzle*, by Hunter Cowdery, art student at West Valley College, who is animating his way to San Jose State University, and Jason Challas, who lectures on computer graphics and fine art at West Valley College.

A legal move in the puzzle exchanges the blank cue ball with any ball to which it is connected by a line. When read clockwise from the top, the balls currently spell ANDREW. According to authors Alex Fink and Richard Guy, the right moves (revealed at the end of their article) lead to a configuration that spells out the answer to all our problems.

AUTHORS

Alex Fink is a third-year mathematics graduate student at the University of California, Berkeley, working under Bernd Sturmfels and Federico Ardila on problems in combinatorics and its interface with algebra and geometry. His undergraduate degrees were in mathematics and computer science at the University of Calgary, in his home town, where he held several undergraduate research grants from NSERC, two supervised by Richard Guy. He has represented Canada and the University of Calgary in several competitions including the Math Olympics, the Putnam Competition, and the ACM International Collegiate Programming Championship. Alex's interests also include linguistics, and he enjoys constructing languages in his spare time.

Richard Guy has enjoyed practicing and teaching mathematics across three continents and at all levels from kindergarten to post-graduate research. He has been privileged to work (or is it play?) with Berlekamp, Conway, Erdős, the Lehmers, Oppenheim, Selfridge, and many other brilliant mathematicians; also many bright students, including his present co-author, a mere seventy years his junior. *Unsolved Problems in Number Theory* and *Winning Ways* are among the better known of his dozen books, and he expects this article to become one of the better known of his three hundred papers. Richard and his 90-year-young wife Louise continue to hike and ski in the Rocky Mountains.

Michael A. Jones earned his doctorate in mathematics at Northwestern University in 1994 under the direction of Donald G. Saari. This article originated as a presentation for the June 2005 DIMACS Reconnect Conference on the Mathematics of Decisions and Elections for which Saari and Jones were the primary and secondary lecturers, respectively. The Reconnect was hosted by Montclair State University where Jones was on the faculty for 10 years until June 2008.

Arthur T. Benjamin teaches at Harvey Mudd College and has served the MAA as co-editor of *Math Horizons* (2004–2008) and Pólya Lecturer (2006–2008). His book, *Proofs That Really Count*, received the MAA's Beckenbach Prize, and his most recent book, *Biscuits of Number Theory* (co-edited with Ezra Brown), was published by the MAA this year.

Daniel Walton earned his B.S. in mathematics at Harvey Mudd College and is currently pursuing his Ph.D. in mathematics at University of California, Los Angeles, supported by an NSF Graduate Fellowship. In his free time, he enjoys playing ultimate frisbee.

Vol. 82, No. 2, April 2009

MATHEMATICS MAGAZINE

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 *MATHEMATICS MAGAZINE* to an individual member of the Association is $131. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount; and new members receive a 20% dues discount for the first two years of membership.)

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

MAA Advertising
1529 Eighteenth St. NW
Washington DC 20036

Phone: (866) 821-1221
Fax: (202) 387-1208
E-mail: advertising@maa.org

Further advertising information can be found online at www.maa.org

Change of address, missing issue inquiries, and other subscription correspondence:

MAA Service Center, *maahq@maa.org*

All at the address:

The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036

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

Postmaster: Send address changes to 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

Rick's Tricky Six Puzzle: S_5 Sits Specially in S_6

ALEX FINK
University of California, Berkeley
Berkeley, CA 94720

RICHARD GUY
The University of Calgary
Calgary, Alberta, Canada T2N 1N4

Many of you will be familiar with the Fifteen Puzzle (FIGURE 1, left). Singmaster [**16**, §5A, pp. 77–84] gives nearly a hundred references to it. It is often associated with the name of Sam Loyd, but Sam continues to be a controversial figure [**9**, Chapter 2, pp. 18–30; **17**]. In the unlikely event that you've never seen the Fifteen Puzzle, you can read about it in the review quoted in the next section.

Sliding block puzzles may be represented by graphs in which the vertices represent possible positions of the blocks and the edges represent the permissible moves of a block from one position to another. For example, the Fifteen Puzzle may be thought of as being played on the sixteen vertices of the graph in FIGURE 1. In this graph, don't think of the numbers as labels for the vertices, but as labeled blocks that can be slid from a vertex to an empty vertex. For example, in the figure, either block 12 or block 15 may be slid onto the vertex where □ indicates that there is no block.

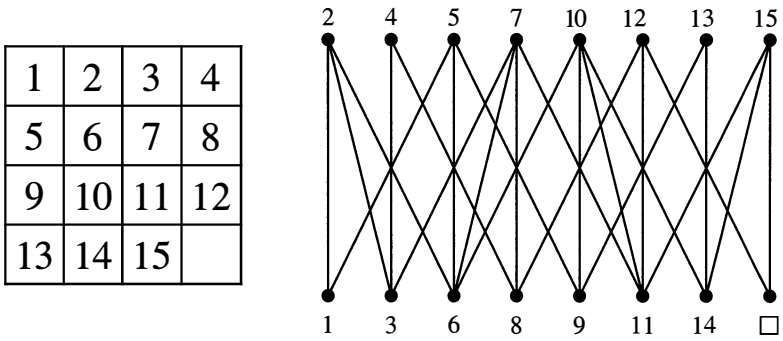


Figure 1 The Fifteen Puzzle and its bipartite graph

The notoriety of the puzzle derives from the impossibility of being able to swap the positions of 14 and 15 in the bottom row, while keeping all the other numbers fixed. This parity property was noted as early as 1879 [**18**, Chapter 1; **19**].

How many people know Rick Wilson's general theorem on sliding block puzzles? We retain Rick's first name to avoid confusion with the well known theorem of Sir John Wilson, first proved by Lagrange, that if p is a prime then $(p-1)!+1$ is divisible by p.

The set of attainable positions in a sliding block puzzle of n pieces sliding on the edges of a graph with $n + 1$ vertices form a group. Rick Wilson's theorem [**25**] states that, apart from simple polygons, and the graph that is the subject of this article, the group of permutations of attainable positions is either S_n, the full symmetric group, if the graph contains an odd circuit, or A_n, the alternating group of even permutations, if the graph contains only even circuits. In the latter case the graph is bipartite, the vertices separate into two sets and there are no edges between members of the same set—the Fifteen Puzzle is the classical example.

We mention that Rick Wilson's theorem applies only to nonseparable graphs, that is, graphs that are 2-connected, or without cut-points, so that there are always at least two paths between any pair of vertices that have no intermediate vertex in common.

What is the exception?

Math Reviews **48** #10882 offers a review by Derek Smith of Wilson's paper [**25**], quoted here with permission from the AMS.

> The 15-puzzle consists of fifteen small movable square tiles numbered 1, 2, ..., 15 and one empty square, arranged in a 4×4 array. One is permitted to interchange the empty square with a tile next to it as often as desired. The challenge is to move by a sequence of such interchanges from one position of the tiles to another specified position. The author generalizes this problem to an arbitrary simple graph and proves that for a finite simple nonseparable graph, with one exception, any position can be reached from any other position unless the graph is bipartite. In the bipartite case, the set of positions splits into two sets, with no position in one set reachable from a position of the other set.

This might be misconstrued to read as though the exception is the set of bipartite graphs. In fact the exception is shown in FIGURE 2. It is a graph on 7 points with 8 edges. It contains two 5-circuits and a 6-circuit, so that we might expect to be able to obtain all $6! = 720$ permutations of the six counters, labeled with the symbols $\mathbf{0}, \mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}, \infty$. Why do we use ∞ instead of **5**? Our labels represent the field $\mathbb{F}_5$ with ∞ adjoined; this will make the connection with the automorphism group of the puzzle clearer.

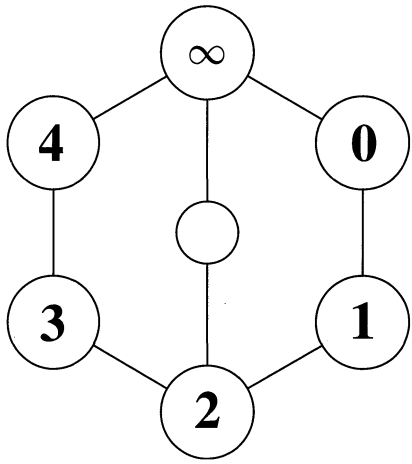


Figure 2 Rick's Tricky Six Puzzle

A little experimentation reveals that there are many arrangements that cannot be attained. The 6! possible arrangements separate into six equivalence classes, with 5! positions in each class. We shall see that

$$\infty\mathbf{01234}, \quad \infty\mathbf{01243}, \quad \infty\mathbf{01324}, \quad \infty\mathbf{01342}, \quad \infty\mathbf{01423}, \quad \infty\mathbf{01432}$$

are representatives, one from each equivalence class. Note that we always read a position clockwise, starting from twelve o'clock. It is not possible to get from any one of these six positions to any other by sliding the disks along the eight edges of the graph.

Not much of a puzzle?

John Conway tells us that he once made a copy of the Tricky Six Puzzle, and we made one that Art Benjamin helped us demonstrate at the 2006 MathFest, but we doubt if it will ever catch on commercially. However, it does have considerable mathematical interest. We shall see that it is related to the projective plane of order 4, to the Hoffman-Singleton graph, to the Steiner system $S(5, 6, 12)$, to a binary (12, 132, 4) code, to the ternary Golay code $\mathcal{C}_{12}$, and to shuffling a deck of cards [**15, 6**]. It is also related to the invariant theory of six points, to "mystic pentagons" and the two-colorings of the three-subsets of a six-element set [**10**], and to the tetracode, the Minimog, and the Rubicon [**5**, pp. 320–330], and to many other things that we don't have room for here.

Many mathematicians are interested in word play, so we asked our favorite anagrammatist, Andrew Bremner, to supply a set of six letters that had many anagrams. He suggested A, C, E, N, R, T. Among the 720 possibilities we found the following twenty words, names and acronyms.

TABLE 1: Six equivalence classes of anagrams

RECANT	ARCNET	CARTEN	CENTRA	CARNET	TANCER
	CANTER	CRANET	CRETAN	CANTRE	TRANCE
	CERANT		NECTAR	CREANT	
	ENCART		TARNEC	NETCAR	
	TERCAN		TRACEN	TANREC	

If you encode these anagrams with R $= \infty$, E $= 0$, C $= 1$, A $= 2$, N $= 3$, and T $= 4$, you will find that it's possible to get from one word to any other in the same column of TABLE 1, but not to any word in a different column. For example, from RECANT, you can't CANTER to any of the other words. We list below four things you CAN do (have we always found the shortest sequence of moves?). If you want to follow along, and to avoid what Conway calls the "alias-alibi problem" (is it the counter? or the position it's in?), then you should label six counters or slips of paper with the symbols ∞, 0, 1, 2, 3, 4 and the letters R, E, C, A, N, T and slide them about on an improvised board. When we write a permutation (ABC...Z) this means that A ends up where B started, B ends up where C started, and so on, cyclically, with Z arriving where A started. By the usual convention, when we string together several such permutations it is the one on the right that acts first: they don't act in the order in which you would normally read them. Compare the out-shuffle with the in-shuffle in the second example below.

1. Cut the deck: swap the first three symbols ∞, 0, 1, with the last three, 2, 3, 4 respectively. The moves 210∞4310∞4310∞432 take RECANT into ANTREC.

This is the permutation $(\infty 2)(03)(14)$. [In anticipation of the next section we will also write this as $x \to (x+2)/(3x+4)$ mod 5. Such a mapping is called a *Möbius transformation*.]

2. Perform an *out-shuffle*, or an *in-shuffle*: cut the deck RECANT into REC and ANT and interleave letters alternately from each half. In an out-shuffle the top card remains on top: RAENCT = (0132) $[x \to 2x+1]$. This can be achieved by the moves $234\infty23102\infty413$. An in-shuffle results in ARNETC = $(\infty 02)(431)$ $[x \to 2/(2x+1)]$ and results from the moves $\infty 012\infty 012\infty 3412\infty 30$. Note that shuffling one way then unshuffling the other performs a cut: $(\infty 20)(134)(0132) = (\infty 2)(03)(14)$. On the other hand, unshuffling then shuffling swaps alternate cards: $(0132)(\infty 20)(134) = (\infty 0)(12)(34)$ $[x \to 2/x]$.

These manipulations of cards don't generate the whole group of the puzzle; they only yield 4! of the 5! possible states, those in which the pairs of cards $\infty 4$, 03, 12, that are equidistant from the centre of the deck, remain so. It doesn't take much experimentation to discover sequences of moves that break up these pairs and generate the whole group:

3. Cycle the first four symbols. The moves $210\infty 2$ followed by $10\infty 21$ and $0\infty 210$ and $\infty 210\infty$ take RECANT $\to$ ARECNT $\to$ CARENT $\to$ ECARNT and back into RECANT. These are the transformations $(\infty 012)$ $[x \to 1/(2x+1)]$, $(\infty 012)^2 = (\infty 1)(02)$ $[x \to (2x+1)/(2x+3)]$, $(\infty 012)^3 = (\infty 210)$ $[x \to (2x+3)/x]$, and $(\infty 012)^4 =$ the identity $[x \to x]$.
4. Fix the first symbol and cycle the other five. The moves $\infty 432104\infty$ send RECANT to RTECAN, $\infty 01234$ to $\infty 40123$, the permutation (01234) $[x \to x+1]$. In fact, combined with the out-shuffle (0132) $[x \to 2x+1]$, this cycle allows us to apply any invertible linear polynomial mod 5 to the finite symbols 0, 1, 2, 3, 4, yielding positions such as (0412) $[x \to 3x+4]$, and its inverse (0214) $[x \to 2x+2]$. These are illustrated in the first of the six diagrams of FIGURE 4 below as all ways to travel round the pentagon or the pentagram.

What is the group of the Tricky Six Puzzle?

As you may have guessed from the brackets in the last section, it is the group $PGL(2, \mathbb{F}_5)$ of Möbius transformations over the field $\mathbb{F}_5$.

$$x \to \frac{px+q}{rx+s} \qquad ps - qr \neq 0$$

This $\mathbb{F}_5$ is the first of several *finite fields* we will encounter. In fact for each prime power q there is a unique field with q elements, which we will denote by $\mathbb{F}_q$. So working in $\mathbb{F}_5$ means working modulo 5—but only because 5 is prime.

There are $5^2 - 1 = 24$ possible nonzero vectors (p, q) for the top row of the matrix $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$, and then $5^2 - 5 = 20$ vectors (r, s) that are independent of the first row, as possibilities for the second row; a total of $24 \times 20 = 480$ nonsingular matrices. But the matrices M, $2M$, $3M$, $4M$, for example

$$\begin{pmatrix} 1 & 0 \\ 4 & 4 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 3 & 3 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 \\ 2 & 2 \end{pmatrix}, \quad \begin{pmatrix} 4 & 0 \\ 1 & 1 \end{pmatrix},$$

all give the same transformation, $(0)(3)(\infty 4)(1\,2)$, taking $\infty 01234$ into 40213∞, or RECANT into TEACNR, so that the number of different transformations is only $480/4 = 120$.

To the surprise of at least one of the authors, this group is isomorphic to S_5, the group of permutations of five objects. We will show that the isomorphism establishing this extends naturally to an automorphism of S_6, under which the group of the puzzle maps to an S_5 subgroup of S_6 given by fixing a point. It's in this context that the isomorphism is most illuminatingly presented.

Two different group actions

An *inner automorphism* of a group is one given by conjugation, that is, each element $x \mapsto a^{-1}xa$ for some fixed element a. The automorphisms of a group themselves form a group, of which the inner automorphisms form a normal subgroup [**2**, pp. 140–141]. The *outer automorphisms* are those automorphisms this doesn't account for: by one definition any non-inner automorphism is outer; by another the outer automorphism group is the quotient of the automorphism group by the inner automorphism group. The symmetric group S_6 is the only finite symmetric group that supports a (nontrivial) outer automorphism [**11**; **14**, Theorem 7.3].

Suppose an abstract group acts on a finite set T (that is, each element of the group permutes T, and permuting by two group elements in succession is the same as permuting by their product). If we were to relabel the elements of T by a permutation a, then an element that acts via the permutation x after the relabelling would have acted by $a^{-1}xa$ before it. Now suppose our abstract group was the symmetric group S_T all along. Then a is in S_T, so $x \mapsto a^{-1}xa$ is an inner automorphism of S_T.

So the existence of an outer automorphism of S_6 means that it can act on sets of size 6 in a fundamentally different way than the obvious one. We'll realize the outer automorphism by constructing such an action, following Sylvester [**20, 21, 22, 23, 24**].

Consider the complete graph on the six points A, B, C, D, E, F. Sylvester calls the six points *monads*, and its $\binom{6}{2} = 15$ edges *duads*. These duads form $15 = 5 \times 3$ matchings, or triads of independent edges, that Sylvester called *synthemes*, and graph theorists know as one-factors. Note that there are 5 choices for A's partner and 3 ways to pair the remaining four.

The graph supports six partitions, or *synthematic totals*, into five synthemes, shown in TABLE 2 and labeled with their associated Tricky Six blocks, ∞, **0, 1, 2, 3, 4**.

TABLE 2: The six totals: the edge-colorings of K_6 with five colors

color	∞	**0**	**1**	**2**	**3**	**4**
r	AB CF DE	AB DE CF	AB FD CE	AB DC FE	AB FE DC	AB EC DF
o	AC DB EF	AC FD EB	AC EF DB	AC BE DF	AC ED BF	AC BF ED
y	AD EC FB	AD CB FE	AD BE FC	AD FB EC	AD CF EB	AD FE CB
i	AE FD BC	AE BF DC	AE DC BF	AE CF BD	AE BC FD	AE DB FC
v	AF BE CD	AF EC BD	AF CB ED	AF ED CB	AF DB CE	AF CD BE

The complete graph K_6 underlying this construction shouldn't be confused with FIGURE 2, the graph of the puzzle itself. As an example, the coloring associated with the label **2**, with AB DC FE colored red, AC BE DF colored orange, etc., is illustrated in FIGURE 3.

If we fix the monad A and operate on the six totals with the $5! = 120$ permutations of the other five monads, we generate the set of possible arrangements of the Tricky Six symbols.

Consider the action of our inner automorphism on conjugacy classes. Within a symmetric group such as S_6 conjugacy classes are just *cycle shapes*, which we write as

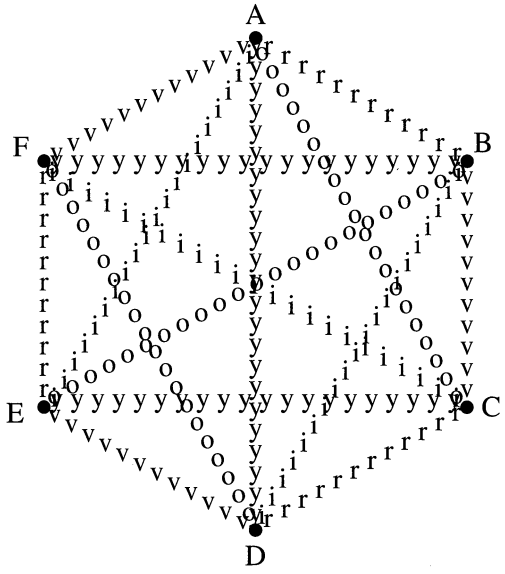


Figure 3 The edge-coloring **2** of K_6, the complete graph on six points

partitions of 6. The cycle shapes on the totals attainable in the puzzle are those that arise from permutations of the monads which fix A, and these have a fixed point in their cycle shape.

For example, if we fix A and three other vertices, we obtain $\binom{5}{2} = 10$ odd permutations of order 2. These are involutions; each is its own inverse. They appear as the first ten entries in TABLE 3:

TABLE 3: Swapping two vertices of K_6

(DE)	(EF)	(FB)	(BC)	(CD)
$(\infty 0)(12)(34)$	$(\infty 1)(23)(40)$	$(\infty 2)(34)(01)$	$(\infty 3)(40)(12)$	$(\infty 4)(01)(23)$
0210	1114	4121	1124	4111
(CF)	(DB)	(EC)	(FD)	(BE)
$(\infty 0)(13)(24)$	$(\infty 1)(24)(30)$	$(\infty 2)(30)(41)$	$(\infty 3)(41)(02)$	$(\infty 4)(02)(13)$
0310	1214	4321	1324	4211
(AB)	(AC)	(AD)	(AE)	(AF)
$(\infty 0)(14)(23)$	$(\infty 1)(20)(34)$	$(\infty 2)(31)(40)$	$(\infty 3)(42)(01)$	$(\infty 4)(03)(12)$

together with the permutations of ∞ 0 1 2 3 4 that they realize, and the entries $pqrs$ of the corresponding Möbius transformation.

For later reference we include as well the five transpositions that move the monad A; these don't realize Möbius transformations.

We thus find that permutations of ABCDEF of shape

$$1^6 \quad 2\cdot 1^4 \quad 2^2 1^2 \quad 2^3 \quad 3\cdot 1^3 \quad 321 \quad 3^2 \quad 4\cdot 1^2 \quad 42 \quad 51 \quad 6$$

map respectively to permutations of $\infty 01234$ of shape

$$1^6 \quad 2^3 \quad 2^2 1^2 \quad 2\cdot 1^4 \quad 3^2 \quad 6 \quad 3\cdot 1^3 \quad 4\cdot 1^2 \quad 42 \quad 51 \quad 321.$$

When A is fixed, respectively

$$1 \quad 10 \quad 15 \quad 0 \quad 20 \quad 20 \quad 0 \quad 30 \quad 0 \quad 24 \quad 0$$

of these are attainable. For example, at the entry $4 \cdot 1^2$ we fix A, and one other letter (5 ways) and cycle the remaining four ($4!/4 = 6$ ways), contributing $5 \times 6 = 30$ to the total of 120. As another example, if we fix A and two other vertices and cycle the rest, $3 \cdot 1^3$, we obtain $\binom{5}{3} \times 2 = 20$ even permutations of order 3. They are displayed in TABLE 4.

TABLE 4: Cycling three of five vertices of K_6

(FBC)	(BCD)	(CDE)	(DEF)	(EFB)
$(\infty 41)(032)$	$(\infty 02)(143)$	$(\infty 13)(204)$	$(\infty 24)(310)$	$(\infty 30)(421)$
1341	0113	1312	2311	3110
(FCB)	(BDC)	(CED)	(DFE)	(EBF)
$(\infty 14)(023)$	$(\infty 20)(134)$	$(\infty 31)(240)$	$(\infty 42)(301)$	$(\infty 03)(412)$
1211	1341	1123	3121	0112
(DEB)	(EFC)	(FBD)	(BCE)	(CDF)
$(\infty 32)(014)$	$(\infty 43)(120)$	$(\infty 04)(231)$	$(\infty 10)(342)$	$(\infty 21)(403)$
1121	2131	0411	1410	1132
(DBE)	(ECF)	(FDB)	(BEC)	(CFD)
$(\infty 23)(041)$	$(\infty 34)(102)$	$(\infty 40)(213)$	$(\infty 01)(324)$	$(\infty 12)(430)$
1431	3211	1140	0141	1213

It will be found that any of the $6 \times 5 \times 4 = 120$ possible arrangements of the first three, or indeed of any three, symbols in a Tricky Six position is attainable, the order of the remaining three then being determined.

All 120 positions are conveniently displayed as the set of six diagrams of FIGURE 4. The first symbol is in the middle of the appropriate diagram. The next two symbols determine a directed edge of a pentagon or pentagram. The final three symbols are then found by continuing to cycle round the pentagon or pentagram in the sense defined by the edge. For example, the position $241xyz$ is found in the diagram having 2 in the middle, where the edge 41 defines the counterclockwise pentagram $41\infty30$, so that $xyz = \infty 30$.

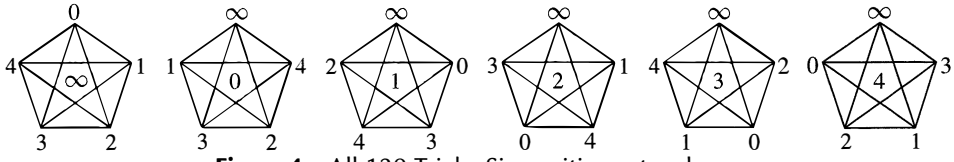


Figure 4 All 120 Tricky Six positions at a glance

The six diagrams of FIGURE 4 are also conveniently viewed as the six pentagonal pyramids that may be sliced from the icosahedron of FIGURE 5, whose opposite vertices are identified. Each pyramid comprises four cycles. For example,

$$\infty(01234)^1 = \infty(01234), \quad \infty(01234)^2 = \infty(02413),$$
$$\infty(01234)^3 = \infty(03142), \quad \infty(01234)^4 = \infty(04321),$$

where the superscripts denote powers, that is the lengths of the steps round the pentagon.

As you can see from TABLE 2, there is a unique synthematic total that is invariant under any five-cycle (JKLMN) on the monads A, B, C, D, E, F. Conway, who introduced

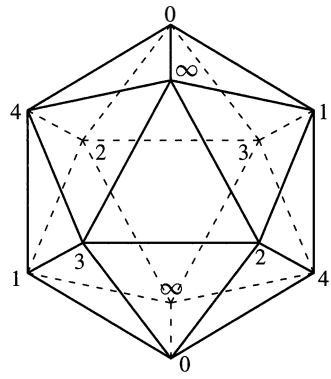


Figure 5 Another good way to see them all

us to Sylvester's notation, denotes it by I(JKLMN). The total I(JKLMN) contains the syntheme IJ KN LM and its images under powers of (JKLMN). FIGURE 6 shows an example.

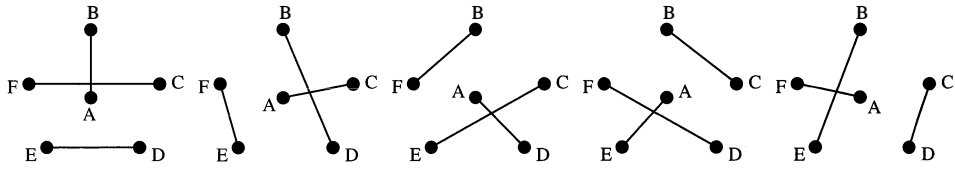


Figure 6 A(BCDEF), the unique synthematic total, also known as ∞, invariant under (BCDEF)

The identities

$$\mathrm{I(JKLMN)} = \mathrm{I(JKLMN)}^{\mathrm{power}} = \overleftrightarrow{\mathrm{J}}(\mathrm{I}\,\overleftrightarrow{\mathrm{LK}}\,\overleftrightarrow{\mathrm{NM}})$$

let us bring any of the 6 monads into the initial position, and write the remainder as any of 5 presentations of any of 4 powers of the five-cycle left over, giving $6 \times 5 \times 4 = 120$ names for each total.

For instance $\mathrm{A(BCDEF)} = \mathrm{A(BCDEF)}^{?}$ for any exponent ? not divisible by 5, and its other names are $\mathrm{B(ADCFE)}^{?} = \mathrm{C(AEDBF)}^{?} = \mathrm{D(AFECB)}^{?} = \mathrm{E(ABFDC)}^{?} = \mathrm{F(ACBED)}^{?}$. Each group of names can be thought of as associated with a pentagram labeled with letters, with the first letter in the centre, like those in FIGURE 4. Such pentagrams are fixed by one of the six subgroups of S_6 of order 20 that fixes $\infty = \mathrm{A(BCDEF)}$.

Indeed, observe that there is a duality of our construction exchanging monads with totals and duads with synthemes, realizable as $\infty 01234 \leftrightarrow \mathrm{ABCDEF}$. Under this exchange the names of the total ∞ become just the attainable Tricky Six permutations. Our situation can be schematized as in FIGURE 7, the symmetry of which makes the duality obvious.

We saw that the first three symbols determine the whole position, and how to read it from FIGURE 4. In fact *any* three symbols determine the position. For example, to find which of ∞, 0, 2 should be assigned to x in $x31y4z$, look for the edge 31 in the ∞, 0 and 2 diagrams of FIGURE 4. It respectively defines the pentagram $(\infty)31420$, the pentagram $(0)31\infty42$, and the pentagon $(2)310\infty4$, of which the second has 4 in

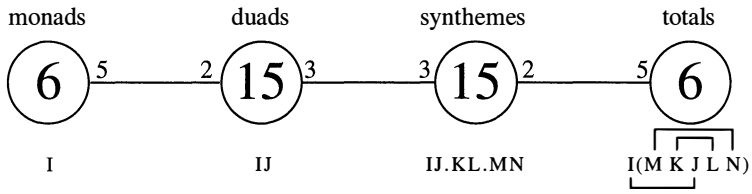


Figure 7 Schematic view of Sylvester's construction

the required position, 031∞42. For another example we may complete $x3y1z4$, by looking in the same three diagrams for the edge 43 (why 43? Think of x as fixed, and notice that 4 and 3 are adjacent in the remaining cycle $3y1z43$). This determines the pentagons (∞)43210, (0)43∞21, (2)4310∞, of which the first has 1 in the required position, ∞32104.

It is through this automorphism that Rick's Tricky Six puzzle is related to the other objects named at the start of the "Not much of a puzzle" section.

Here's a first brief example. Implicit in the way we've written TABLE 2 is another set of six objects paired with the totals, the *mystic pentagons* which begin the interesting paper [**10**]. The ten duads that don't contain A form two sets of five: the second and third columns of each total. Each monad appears twice in each column. If we forget the synthemes and remember only the column divisions, we get a mystic pentagon, that is, a partition of the edges of complete graph on vertices BCDEF into two five-cycles. There are in fact only six mystic pentagons, and we get each of them once (FIGURE 8). Therefore the permutations of the mystic pentagons which can be attained by permuting BCDEF exactly form the Tricky Six group.

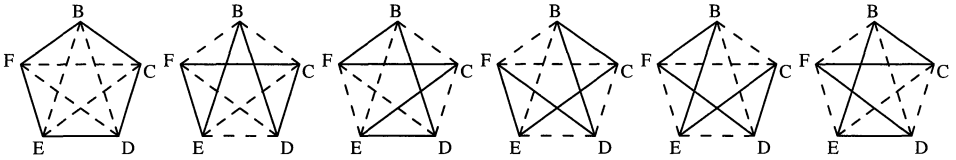


Figure 8 The six mystic pentagons

The remainder of this paper is devoted to a more leisurely examination of several other examples.

The projective plane of order 4

The projective plane of order four, PG(2, $\mathbb{F}_4$), is often defined by means of a *cyclic difference set*, for example {3, 6, 12, 7, 14} modulo 21, whose five members generate the $\binom{5}{2}$ differences $\pm1, \pm2, \ldots, \pm10$. Note that the first three elements generate the multiples of 3, and the last two generate the multiples of 7. Think of the difference set as a complete pentagon which cycles round a complete regular 21-gon as in FIGURE 9. Among its 10 edges there is exactly one of every possible length, so that every pair of the 21 points belongs to just one pentagon. Dually, any two pentagons have just one vertex in common.

Call the pentagon {3, 6, 12, 7, 14} the *line* **0**. Subtract 3, 4, 9, 11 modulo 21 to give the respective lines

3: {0, ♠, 9, 4, 11}, **4**: {20, 2, 8, ♡, 10},
9: {15, 18, ◇, 19, 5}, **11**: {13, 16, 1, 17, ♣}.

These four lines each pass through the point 3 which is denoted differently in each of them, by ♠, ♡, ◇, and ♣ in turn, and is circled in FIGURE 9. The other four points on each of these lines are represented in the figure by the corresponding suit symbols. They exactly cover the 16 points which are not on line **0**.

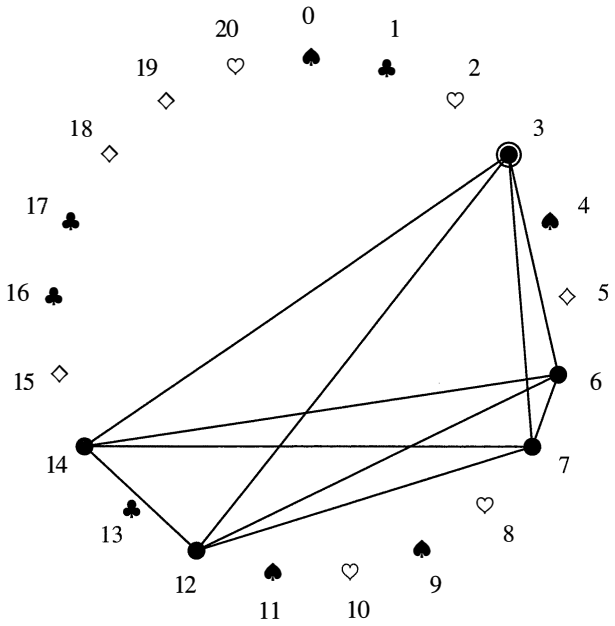


Figure 9 A difference set generates the projective plane of order 4

In general, we give the *line* $\{3-n, 6-n, 12-n, 7-n, 14-n\}$ modulo 21 the name **n**, $\mathbf{0} \leq \mathbf{n} \leq \mathbf{20}$, as in TABLE 5, which displays a configuration of 21 points and 21 lines with 5 points on each line, 5 lines through each point, every pair of lines intersecting in a point and every pair of points determining a line. Bold numbers refer to lines, ordinary numbers to points (or vice versa, since the configuration is self-dual). The line **i** passes through the point j if and only if the point i lies on the line **j**.

TABLE 5: Incidences in the projective plane of order 4

lines	**0**	**20**	**19**	**18**	**17**	**16**	**15**	**14**	**13**	**12**	**11**	**10**	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
points	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	0	1	2
points	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	0	1	2	3	4	5
points	12	13	14	15	16	17	18	19	20	0	1	2	3	4	5	6	7	8	9	10	11
points	7	8	9	10	11	12	13	14	15	16	17	18	19	20	0	1	2	3	4	5	6
points	14	15	16	17	18	19	20	0	1	2	3	4	5	6	7	8	9	10	11	12	13

Twenty-one is not a prime power, so the numbers $0, 1, \ldots, 20$ do not form a field. However, they do form an additive cyclic group, and the twelve numbers which are not multiples of 3 or 7 form a multiplicative group, of which the powers of 2 are a subgroup.

Let's find two different actions of S_6 in this projective plane. As the two sets of size six let us take the points 1 2 4 8 16 11 (the powers of two, $2^0, 2^1, 2^2, 2^3, 2^4, 2^5$,

mod 21) and the lines **0 18 15 9 14 7** (zero and the negatives of the original difference set).

We can begin to rewrite TABLE 2 for the projective plane by replacing the labels A B C D E F of the vertices of K_6 with the respective point numbers 1 2 4 8 16 11. We also relabel the totals, ∞ **0 1 2 3 4** with the respective line numbers **18 15 14 0 7 9**. Then, with TABLE 5 as our guide, we label the edges AB, CF, DE which join the points 1 & 2, 4 & 11, 8 & 16, with the line-numbers **5 3 19** and similarly for all the fifteen synthemes. The lines **5 3 19** concur in the point 9 and each syntheme corresponds to a point. The labels of these fifteen points are just those numbers that are not powers of two, and TABLE 2 turns into TABLE 6. You can check that this is the same configuration, with the same labelling, as before.

TABLE 6: An assignment of numbers to TABLE 2

18				**15**				**14**				**0**				**7**				**9**			
5	**3**	**19**	9	**5**	**19**	**3**	9	**5**	**16**	**8**	19	**5**	**20**	**17**	7	**5**	**17**	**20**	7	**5**	**8**	**16**	19
2	**4**	**17**	10	**2**	**16**	**12**	12	**2**	**17**	**4**	10	**2**	**12**	**16**	12	**2**	**19**	**1**	5	**2**	**1**	**19**	5
6	**8**	**1**	6	**6**	**10**	**17**	18	**6**	**12**	**3**	0	**6**	**1**	**8**	6	**6**	**3**	**12**	0	**6**	**17**	**10**	18
11	**16**	**10**	17	**11**	**1**	**20**	13	**11**	**20**	**1**	13	**11**	**3**	**4**	3	**11**	**10**	**16**	17	**11**	**4**	**3**	3
13	**12**	**20**	15	**13**	**8**	**4**	20	**13**	**10**	**19**	14	**13**	**19**	**10**	14	**13**	**4**	**8**	20	**13**	**20**	**12**	15

The points 1, 2, 4, 8, 16, of which no three are collinear, form a *conic*, that is, the solution set of a homogeneous quadratic over the field of order four. The *tangents* to the conic are the lines that meet the conic in just one point (indicated by a hat):

$$\mathbf{13}\,\{11\ 14\ 20\ 15\ \hat{1}\},\quad \mathbf{16}\,\{\hat{8}\ 11\ 17\ 12\ 19\},\quad \mathbf{1}\,\{\hat{2}\ 5\ 11\ 6\ 13\},$$
$$\mathbf{17}\,\{7\ 10\ \widehat{16}\ 11\ 18\},\quad \mathbf{3}\,\{0\ 3\ 9\ \hat{4}\ 11\}.$$

These are the five lines through the point 11. This point combines with the conic to form a *hyperconic*, six points no three of which are collinear. These six points are the monads, and determine $\binom{6}{2} = 15$ lines, the duads, which meet in threes at the other fifteen points; these correspond to the synthemes. The remaining six lines (**0**, **7**, **14**, **9**, **18**, **15**) that don't meet the hyperconic correspond to the totals; no three of them concur and they form a set of lines dual to the set of six points.

We repeat FIGURE 7 as FIGURE 10, annotating the nodes further to make clear the interpretation of the figure as the projective plane.

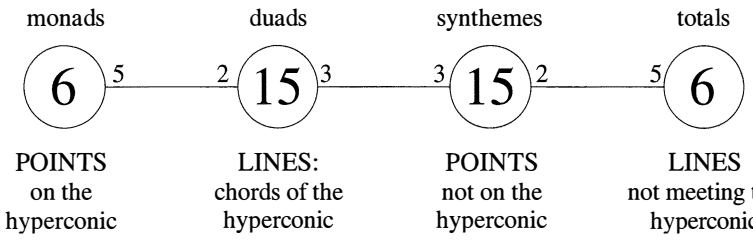


Figure 10 Schematic view of the projective plane of order 4

Our two nonisomorphic S_6-actions show up here as the action that permutes the points of any six-point hyperconic, like 1 2 4 8 16 11, and the action induced on the lines not meeting it, in this case **0 7 14 15 18 9**. Our numbering makes it easy

to check that doubling all the vertex labels modulo 21 is an automorphism that fixes the hyperconic under which line labels are also doubled, so the cycle (1 2 4 8 16 11) induces the permutation (**0**)(**7 14**)(**9 18 15**) of the six lines. If we swap 1 and 2 and fix the other four points, (1 2)(4)(8)(16)(11), this induces (**0 7**)(**15 18**)(**9 14**) on the lines and these two automorphisms are enough to generate the whole group.

We can't draw the plane with straight lines, so, in FIGURE 11, although the twenty-one points 0, 1, 2, ... , 20 are clear, the lines are less so. The line **9** is the incircle of the pentagon and the lines **0, 14, 15, 18, 7** look like petals. The lines **3, 16, 17, 13, 1** are the diameters through the point 11. The lines **4, 8, 6, 12, 2** are pentagram edges, that need to be bent round to pass through the respective points 3, 19, 18, 15, 5; and the lines **11, 5, 10, 20, 19** are pentagon edges, both ends of which should be bent round to pass through the respective pairs of points 17&13, 7&9, 14&17, 13&7, 9&14.

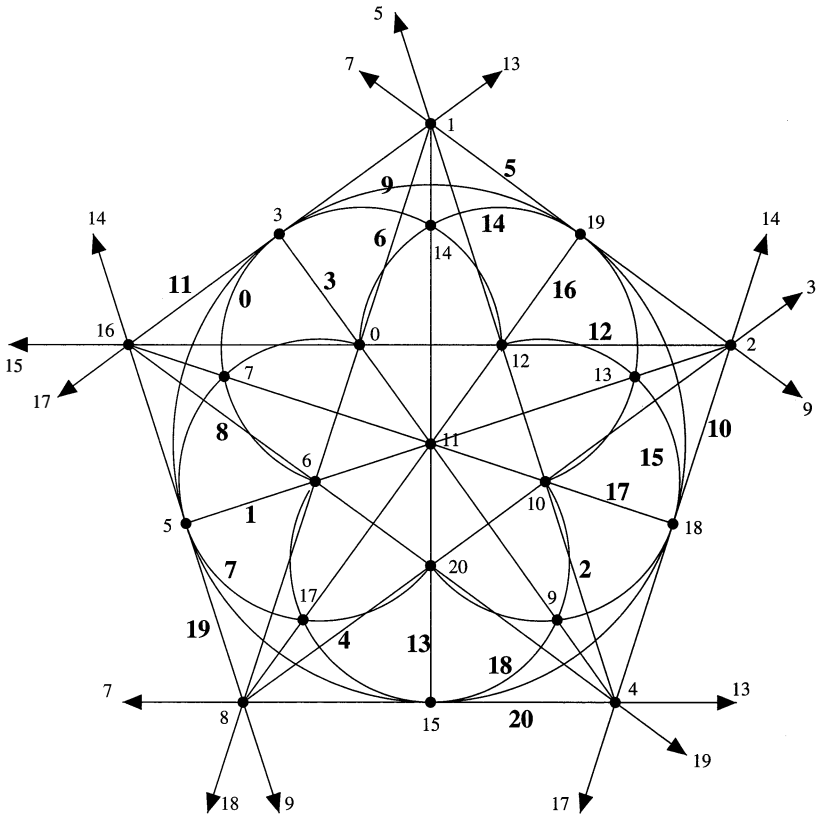


Figure 11 The projective plane of order 4

The points 3, 6, 12, 7, 14 of line **0** thus lie on the respective lines **3, 6, 12, 7, 14** and, of course, lie just one on each of the remaining fifteen lines. The other four points on such a line comprise two pairs that form triples with the line number, each member of a triple being the number of the line containing the other two points. For example, line **18** contains the point 6 and the four points 9, 15, 10, 17 whose joins to the point 18 are the respective lines **15, 9, 17, 10** which form the triples {**18**, **15**, **9**}, and {**18**, **17**, **10**}. There are ten such triples and they exhibit the ten differences $1 \leq d \leq 10$ exactly three times each. For example, the difference 5 occurs in the triples {**8**, **13**, **4**}, {**11**, **16**, **7**}, and {**15**, **20**, **13**}. These ten triples correspond to the sets of edges of pairs of opposite faces of an icosahedron, half of which is shown in FIGURE 12.

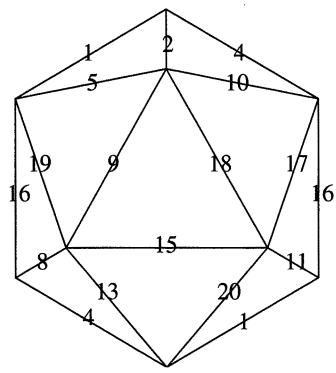


Figure 12 Ten triples form half an icosahedron

Buy one; get several free!

We noticed that the difference set {3, 6, 12, 7, 14} comprised two difference sets: {7, 14} generates the multiples of 7 and {3, 6, 12} generates the multiples of 3. So the projective plane of order four contains the not very exciting projective plane of order one: the triangle {0, 7, 14} and 1119 other copies of it, and the much more interesting projective plane of order two, the so-called Fano configuration (although it was known more than 40 years earlier to the Rev. T. P. Kirkman [**12**]). Besides the obvious example, whose point-numbers are congruent to 0 modulo 3, which is self-dual in the sense that it has the same line-numbers, and is shown in FIGURE 13, there are 359 others: including the dual pair whose point- and line-numbers are respectively congruent to 1 and 2 (or to 2 and 1) modulo 3. The figure also shows a dual pair whose point-numbers differ by 3 from the line-numbers.

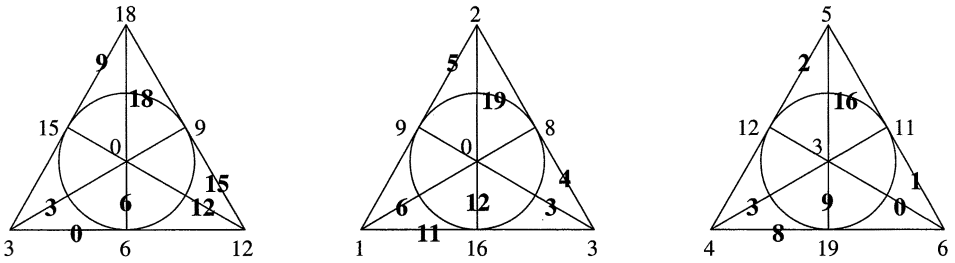


Figure 13 Kirkman-Fano configurations

More surprising is the fact [**1**] that if we throw away a hyperconic we are left with fifteen points which form a projective geometry of order two in three dimensions! For example, throw away 1, 2, 4, 8, 16, 11. The remaining points are those of the line {0, 5, 7, 17, 20}, and its double {0, 10, 14, 13, 19}, together with the multiples of 3. FIGURE 14 shows this geometry as a tetrahedron, as Polster would draw it [**13**]. Its fifteen points are the vertices, 5 7 17 20, the midpoints of the edges (multiples of 3), the centroids of the faces, 10 14 13 19, and the centroid, 0. Fifteen of the thirty-five lines, those which meet the hyperconic, are inherited from the plane: they are the twelve medians of the faces and the three joins of midpoints of opposite edges. The other twenty are the vertex sets of triangles formed by three of the six lines **0**, **7**, **14**, **9**, **18**, **15** which avoid the hyperconic. They appear as the six edges of the tetrahedron, the four joins of the vertices to the centroids of the opposite faces, and ten lines which cannot

be drawn in Euclidean space: the four incircles of the faces and six similar curves circumscribing the "medial triangles":

{3, 14, 19} {6, 10, 14} {9, 10, 13} {12, 13, 14} {15, 10, 19} {18, 13, 19}

formed by the triples of lines

14, 9, 0 **14, 0, 18** **14, 15, 18** **14, 0, 15** **14, 9, 18** **14, 9, 15.**

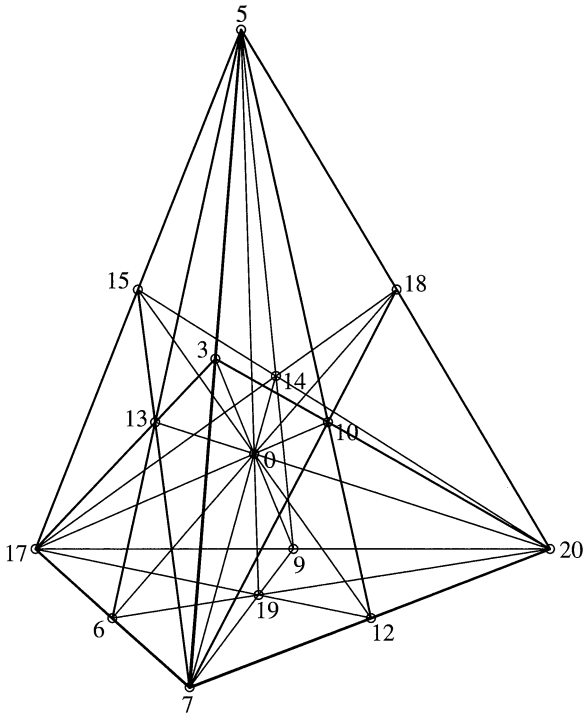


Figure 14 The projective geometry PG$(3, \mathbb{F}_2)$

A different and quite revealing labelling of the $15 = 2^4 - 1 = 4 + 6 + 4 + 1$ points is to assign 1, 2, 4, 8 to the vertices, sums of pairs of these to the midpoints of the edges, sums of three to the centroids of the faces, and the sum of all four, 15, to the centroid.

old numbers	5	7	3	17	15	6	13	20	18	12	10	9	14	19	0
new numbers	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

The thirty-five lines are then those triples whose nim-sums (XOR, binary addition without carry) are zero: the ten "noneuclidean" lines correspond to those nim-sums which are not ordinary sums, for example, $3 \oplus 5 = 6$ and $5 \oplus 11 = 14$. The $15 = 2^4 - 1 = 4 + 6 + 4 + 1$ planes of the geometry are Kirkman-Fano configurations: the four faces of the tetrahedron, the six "medial planes" joining the midpoint of an edge to the opposite edge, the four "cones" joining a vertex to the incircle of the opposite face, and the "sphere" of midpoints of edges together with its centre, 15.

Remarkably, the thirty-five lines can be partitioned, in 240 different ways, into seven sets of five lines, with no two of the five intersecting, each set exactly covering the fifteen points. That is, the thirty-five lines can be arranged as rows in a *Kirkman*

(15, 3, 1)-*design*; they provide solutions to the famous Kirkman schoolgirls problem, with which readers of the previous issue of this MAGAZINE will already be familar [**4**]. An example is shown in TABLE 7.

TABLE 7: The thirty-five lines of PG(3,$\mathbb{F}_2$) form a Kirkman (15,3,1)-design

Sun	Mon	Tue	Wed	Thu	Fri	Sat
1 2 3	1 4 5	1 6 7	1 8 9	1 10 11	1 12 13	1 14 15
5 8 14	3 9 10	3 8 11	2 4 6	2 5 7	3 4 7	3 5 6
4 11 15	2 12 14	2 13 15	3 12 15	3 13 14	2 9 11	2 8 10
7 9 14	7 8 15	5 9 12	5 11 14	4 8 12	5 10 15	4 9 13
6 10 12	6 11 13	4 10 14	7 10 13	6 9 15	6 8 14	7 11 12

The fifteen Kirkman-Fano planes each appear as seven triples, one from each day of the week. For example, the "cone" 1 6 7 10 11 12 13 is represented by 6 10 12, 6 11 13, 1 6 7, 7 10 13, 1 10 11, 1 12 13, and 7 11 12.

A somewhat surprising connection between PG(3, $\mathbb{F}_2$) and the Lehmers' method of factoring integers by means of quadratic forms is made in [**7**, §§26 & 27].

The Hoffman-Singleton graph

A *Moore graph* of type v, k is a regular graph of valence v and diameter k with the maximum possible number of vertices, namely $\big(v(v-1)^k - 2\big)/(v-2)$. This formula doesn't make sense if $v = 2$, but it tends to the limit $2k + 1$ as v approaches 2, and this is the number of vertices in the valence 2 case. Hoffman & Singleton [**8**] showed that for diameter 2 there are at most four such. Their valences are 2 (the pentagon), 3 (the Petersen graph), 7 (the Hoffman-Singleton graph) and possibly 57 (though the existence of this last remains an unsolved problem). The Hoffman-Singleton graph has 50 vertices and 175 edges, and like every Moore graph of diameter 2 its shortest cycles are pentagons so that its *girth* is 5. Its automorphism group has order $252000 = 2^5 3^2 5^3 7$. It is *arc-transitive*, that is it has an automorphism sending a particular edge to any of its 175 edges with either of 2 orientations. The stabilizer of an oriented edge thus has order $252000/(175 \cdot 2) = 720$, and indeed is isomorphic to S_6, as reflected in the following construction of the graph from our versatile TABLE 2.

To draw the Hoffman-Singleton graph, start with an edge joining vertices which we label ⋆ and G. Label the six other vertices adjacent to ⋆ with the letters A B C D E F and the other six adjacent to G with the symbols ∞ 0 1 2 3 4 as in FIGURE 15. The other 36 vertices are {Xn}, where X runs through the letters A B C D E F and n runs through the symbols ∞ 0 1 2 3 4, and there are the implied adjacencies, for example vertex C2 is adjacent to vertices C and 2. It remains to insert the other $175 - (1 + 12 + 36 + 36) = 90$ edges. Again, they correspond to our edge-colorings of K_6.

Recall the fifteen swaps of TABLE 3. They each provide six adjacencies, for example

$$(\text{CE}) \qquad (\infty 2)(30)(41)$$

provides the six adjacencies

C∞—E2 C2—E∞ C3—E0 C0—E3 C4—E1 C1—E4

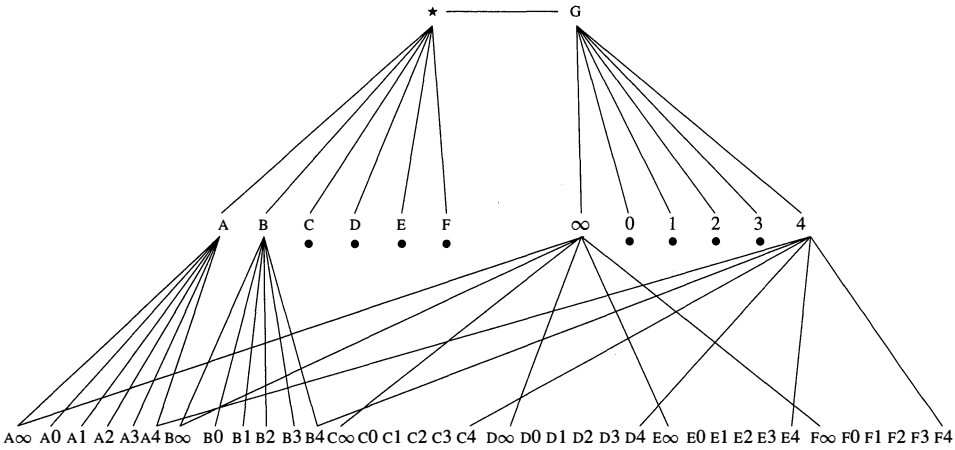


Figure 15 How to construct the Hoffman-Singleton graph

We can also succinctly describe the 6! automorphisms of the graph fixing the edge ★—G: they permute the vertices A B C D E F arbitrarily and the vertices ∞ 0 1 2 3 4 as dictated by construction.

Other constructions for the Hoffman-Singleton graph are given in [**3**, §13.1]. Conway showed us his perspective, which begins with a distinguished vertex rather than an edge. We'll choose ★ in FIGURE 15 as this vertex. Its neighbors are the six monads ABCDEF and G, and the other neighbors of G are the totals. This suggests that to place all seven neighbors of ★ on an equal footing we should recognize G as a seventh monad and interpret the other neighbors of an original monad I as the totals on the set of the six other monads, so that what we before called X*n* is reinterpreted as the total *n* with X replaced by G. Therefore the vertices adjacent to a numbered total *n* on ABCDEF are just the totals X*n* on ABCDEFG that differ from it only by a single-letter substitution. In fact this turns out to be true of any pair of totals, determining all remaining edges of the graph. The resulting picture of the Hoffman-Singleton graph is FIGURE 16.

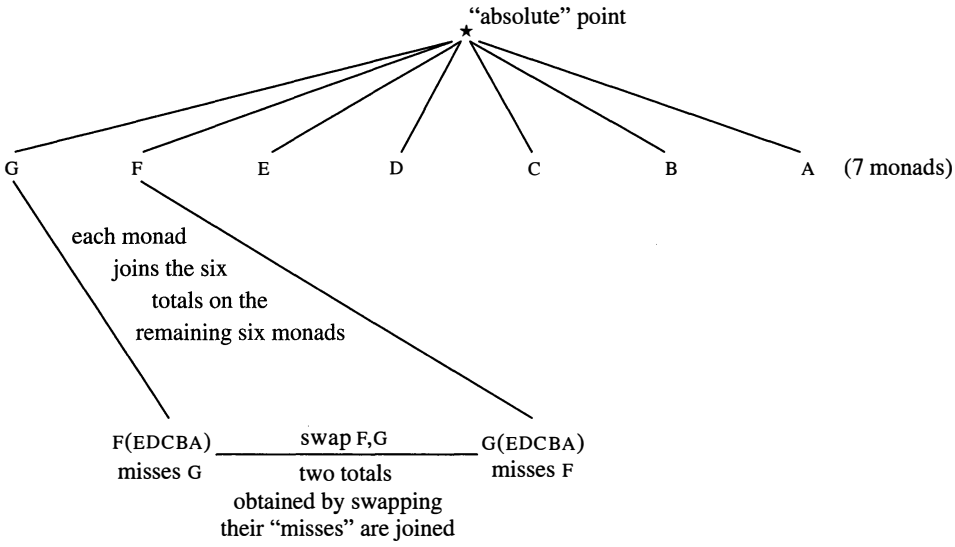


Figure 16 Conway's description of the Hoffman-Singleton graph

The Steiner system $S(5, 6, 12)$

The Steiner system $S(5, 6, 12)$ is a set of *blocks* of 6 elements, *hexads*, chosen from a set of 12 so that each *pentad*, or choice of 5 elements from the 12, occurs exactly once in a block. Hence the number of blocks is $\binom{12}{5}/\binom{6}{5} = 132$.

We use A B C D E F ∞ 0 1 2 3 4 for our 12 elements: in fact ABCDEF and ∞01234 will be two of the blocks. We get $15 \times 6 = 90$ blocks that contain four letters and two numbers, or two letters and four numbers, from the fifteen swaps of TABLE 3.

For example, the swap

$$(\text{FB})\ (\infty 2)(34)(01)$$

yields the six blocks

$$\text{A2CDE}\infty, \quad \text{A4CDE3}, \quad \text{A1CDE0}, \quad \text{F01B34}, \quad \infty\text{012FB}, \quad \infty\text{FB234}$$

where the pairs of numbers ∞2, 34, 01 have been substituted for the pair of letters FB in ABCDEF and, conversely, the letters FB have been substituted for the pairs of numbers in ∞01234.

The other 40 blocks have three letters and three numbers and may be generated in pairs from the $\binom{6}{3} = 20$ three-cycles of TABLE 4, by substitutions exchanging three letters and three digits. That table omits the three-cycles moving the monad A, but all we need here is the partition of the totals into the two three-cycles that these induce, and this partition is the same one that arises from the cycles on the other three monads. So for instance the cycles (BDE) and (BED) correspond to the permutations $(\infty 32)(014)$ and $(\infty 23)(041)$ while (ACF) and (AFC) correspond to $(\infty 23)(014)$ and $(\infty 32)(041)$.

For example, the cycle (BEF) associated with the permutation $(\infty 30)(214)$ gives rise to the four blocks

$$\text{A}\infty\text{CD30} \quad \text{A1CD42} \quad \infty\text{0BF3E} \quad \text{BF12E4}$$

How do we know that each pentad occurs exactly once? If a pentad consists of 5 letters, or 5 numbers, then the hexad is ABCDEF or ∞01234. If it consists of 4 letters and a number n the hexad will contain a second number. This is found in TABLE 3 which displays all $\binom{6}{2} = 15$ swaps of two vertices. Select the swap of the two letters which are not in the pentad and take the number paired with n. For example, given the pentad ACEF3, look at the entry (DB) $(\infty 1)(24)(30)$ where 3 is paired with 0, so that the pentad belongs to the unique hexad A3C0EF. If the pentad contains 4 numbers and a letter, for example, ∞024B, find the entries of TABLE 3 that contain the missing numbers 13, namely (AD), (CF), (BE). Here B is paired with E, so the hexad is ∞0B2E4. If the pentad contains 3 letters and 2 numbers, or 3 numbers and 2 letters, we use TABLE 4. For example, for BCF23 we find (FBC) $(\infty 41)(203)$ so that the hexad is completed with 0. But if the pentad were BCF24, with 2 and 4 in different triples, the hexad must be completed with a letter. In TABLE 3 the pair (24) occurs in the swaps (AE), (BD) and (FC), so the missing letter is D: 2BCD4F.

If the pentad were BCF02, then (02) occurs in (AC), (DF), (EB) with B C F in three different pairs: the pentad requires a number; TABLE 4 gives (FBC) $(\infty 41)(203)$; the missing number is 3.

A (12, 132, 4) binary code and the ternary Golay code $\mathcal{C}_{12}$

In a binary code, the letters of the codewords are zeroes and ones. The number of letters in a codeword is its length and the number of ones is its weight. The 132 hexads

of the Steiner system $S(5, 6, 12)$ form a basis for a binary code with words of length 12 and weight 6.

The blocks of the Steiner system indicate which letters of the 12-letter codewords are occupied by six ones or by six zeroes. In anticipation of the construction of the ternary Golay code $\mathcal{C}_{12}$ we will put the letters in the order

$$\text{A } 0\,1\,2\,3\,4\,\infty \text{ B C D E F}$$

and, for ease of reading, we will leave space round the 1st and 7th letters.

For example, our initial blocks ABCDEF and ∞01234 correspond to the codewords 1 00000 0 11111 and 0 11111 1 00000; the blocks

$$\text{A2CDE}\infty,\ \text{A4CDE3},\ \text{A1CDE0},$$

and their complements

$$\text{F01B34},\ \infty\text{012FB},\ \infty\text{FB234}$$

correspond respectively to the codewords

$$1\ 00100\ 1\ 01110, \quad 1\ 00011\ 0\ 01110, \quad 1\ 11000\ 0\ 01110,$$

and their complements

$$0\ 11011\ 0\ 10001, \quad 0\ 11100\ 1\ 10001, \quad 0\ 00111\ 1\ 10001,$$

while the blocks

$$\infty\text{B30EF},\ 1\text{B24EF},\ \text{AD12C4},\ \infty\text{0AC3D}$$

correspond to

$$0\ 10010\ 1\ 10011, \quad 0\ 01101\ 0\ 10011, \quad 1\ 01101\ 0\ 01100, \quad 1\ 10010\ 1\ 01100.$$

Each codeword differs from every other in at least four places, that is, the *Hamming distance* between any two words is at least 4.

Suppose that you received a codeword 0 01101 1 10101. This contains seven ones, so there is an error. Assume that the zeroes are correct. They correspond to the pentad A03CE. TABLE 4 has the (complementary to A C E) entry (FBD) (∞04)(231); 0 and 3 are in different triples, so the missing element is a letter. In TABLE 3 the pair (30) occurs in (DB), (EC), and (AF), so that the missing letter is F and the final 1 in the erroneous codeword should have been 0, making it 0 01101 1 10100.

We can pass from this binary code to a ternary code, which we now present in outline.

To incorporate the words of our binary code into a ternary code we will leave the zeroes as they are and endow the ones with signs. With a correct choice of signs the resulting 132 words of length 12 can be made to generate by addition a *linear code* of dimension 6, that is a 6-dimensional subspace of the ambient vector space $\mathbb{F}_3^{12}$ over the finite field $\mathbb{F}_3 = \{-1, 0, +1\}$. Our code will thus contain $3^6 = 729$ codewords.

Aside from the zero word 0 00000 0 00000, the words will come in pairs of opposite sign. In fact, we will obtain no nonzero codewords with more zeroes than the signed manifestations, two apiece, of our 132 words from the binary code. So the minimal distance of our code will increase to 6. The resulting code is known as the *ternary Golay code* and denoted as $\mathcal{C}_{12}$.

From [**5**, p.85] we learn that $\mathcal{C}_{12}$ may be obtained by appending a zero-sum check digit to $\mathcal{C}_{11}$, the quadratic residue code of length 11 over $\mathbb{F}_3$; that a generator matrix is

$$\begin{array}{c} \begin{array}{cccccccccccc} \text{A} & 0 & 1 & 2 & 3 & 4 & \infty & \text{B} & \text{C} & \text{D} & \text{E} & \text{F} \end{array} \\ \left[\begin{array}{cccccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 1 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & -1 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & -1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 & -1 & 1 & 0 \end{array}\right] \end{array}$$

that it has weight enumerator

$$x^{12} + 264x^6y^6 + 440x^3y^9 + 24y^{12}$$

that is it contains 1, 264, 440, and 24 words with respectively 0, 6, 9, and 12 nonzero letters, and that its automorphism group is $2.M_{12}$, that is it has the Mathieu group M_{12} as a normal subgroup with quotient (cyclic of order) 2.

But by now we've roved far enough from the Tricky Six puzzle, so we pursue codes no further and turn to the

Conclusion

Our favorite for an actual puzzle changes C into W and T into D, turning RECANT into REWAND. Manoeuvre #1 of the "Not much of a puzzle" section then gives the figure on the cover of this MAGAZINE, which should be read clockwise, starting from twelve o'clock. The solution: move the letters R E W A R E W A R E and read clockwise from noon again.

Acknowledgment. We are indebted to John Conway for several helpful insights, and to Ezra Brown, Richard Nowakowski, and two referees, all of whom read our first draft with great care and made many constructive suggestions for improvement.

REFERENCES

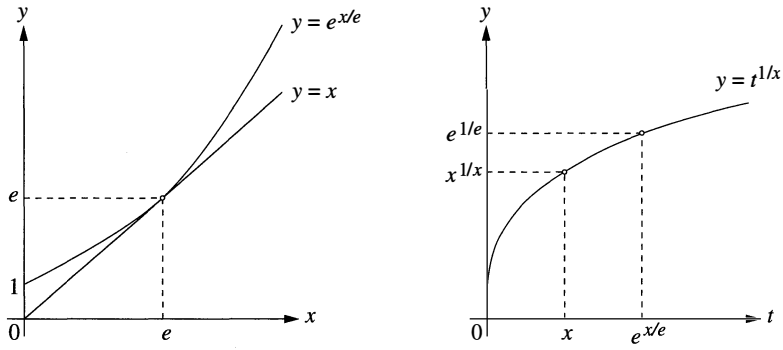
1. Albrecht Beutelspacher, 21 − 6 = 15: a connection between two distinguished geometries, *Amer. Math. Monthly* **93** (1986) 29–41; *MR* **87g**:51010.
2. Garrett Birkhoff and Saunders Mac Lane, *A Survey of Modern Algebra*, 3rd ed., Macmillan, New York, 1965.
3. A. E. Brouwer, A. M. Cohen, and A. Neumaier, *Distance-regular Graphs, Ergebnisse der Mathematik und ihrer Grenzgebiete* (3), **18**, Springer-Verlag, Berlin, 1989.
4. Ezra Brown and Keith A. Mellinger, Kirkman's schoolgirls wearing hats and walking through fields of numbers, this MAGAZINE **82** (2009) 1–14.
5. J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups, Grundlehren der mathematischen Wissenschaften* **290**, Springer-Verlag, 1988.
6. P. Diaconis, R. L. Graham, and W. M. Kantor, The mathematics of perfect shuffles, *Adv. in App. Math.* **4** (1983) 175–193; *MR* **84j**:20040.
7. Richard K. Guy, The unity of combinatorics, *Proc. 25th Iran. Math. Conf., Tehran* (1994), *Math. Appl.* **329** 129–159, Kluwer Acad. Publ., Dordrecht, 1995; *MR* **96k**:05001.
8. A. J. Hoffman and R. R. Singleton, On Moore graphs with diameters 2 and 3, *IBM J. Res. Develop.*, **4** (1960) 497–504; *MR* **25** #3857.
9. L. E. Hordern, *Sliding Piece Puzzles*, Oxford Univ. Press, 1986.
10. Ben Howard, John Millson, Andrew Snowden, and Ravi Vakil, A new description of the outer automorphism of S_6, and the invariants of six points in projective space, *J. Combin. Theory A* **115** (2008) 1296–1303.
11. Gerald Janusz and Joseph Rotman, Outer automorphisms of S_6, *Amer. Math. Monthly* **89** (1982) 407–410; *MR* **83g**:20002.
12. T. P. Kirkman, Note on an unanswered prize question, *Cambridge & Dublin Math. J.* **5** (1850) 255–262.

13. B. Polster, Pretty pictures of geometries, *Finite geometry and combinatorics* (Deinze 1997), *Bull. Belg. Math. Soc. Simon Stevin* **5** (1998) 417–425; *MR* **99f**:51022.
14. J. J. Rotman, *The Theory of Groups, An Introduction*, 2nd ed., Allyn & Bacon, Boston, 1973.
15. Daniel Scully, Perfect shuffles **77**(2004) 101–117; *MR* **2005g**:05008.
16. David Singmaster, *Sources in Recreational Mathematics*, 6th ed., Nov. 1993.
17. Jerry Slocum and Dic Sonneveld, *The 15 Puzzle: How it Drove the World Crazy. The Puzzle that Started the Craze of 1880. How America's Greatest Puzzle Designer, Sam Loyd, Fooled Everyone for 115 Years*, Slocum Puzzle Foundation, Beverly Hills, CA, 2006.
18. S. K. Stein, *Mathematics: The Man-Made Universe*, 2nd ed., Freeman, San Francisco, 1969.
19. W. E. Story, Notes on the '15' puzzle, II, *Amer. J. Math.* **2** (1879) 399–404.
20. J. J. Sylvester, Elementary researches in the analysis of combinatorial aggregation, *Phil. Mag.* **24** (1844) 285–296; *Collected Math. Papers*, Vol. I, 91–102.
21. J. J. Sylvester, Note on the historical origin of the unsymmetrical six-valued function of six letters, *Phil. Mag.* **21** (1861) 369–377; *Collected Math. Papers*, Vol. II, 264–271.
22. J. J. Sylvester, On a problem in tactic which serves to disclose the existence of a four-valued function of three sets of three letters each, *Phil. Mag.* **21** (1861) 515–520; *Collected Math. Papers*, Vol. II, 272–276.
23. J. J. Sylvester, Concluding paper on tactic, *Phil. Mag.* **22** (1861) 45–54; *Collected Math. Papers*, Vol. II, 277–285.
24. J. J. Sylvester, Remark on the tactic of nine elements, *Phil. Mag.* **22** (1861) 144–147; *Collected Math. Papers*, Vol. II, 286–289.
25. Richard M. Wilson, Graph puzzles, homotopy, and the alternating group, *J. Combin. Theory Ser. B* **16** (1974) 86–96; *MR* **48** #10882.

Proof Without Words: Steiner's Problem on the Number *e*

For what positive x is the xth root of x the greatest? [**1, 2**]

Solution. $x > 0 \Rightarrow \sqrt[x]{x} \le \sqrt[e]{e}$.



(a) $x \le e^{x/e}$ $\Rightarrow$ (b) $x^{1/x} \le e^{1/e}$

[In the right-hand figure, $x > 1$; the other case differs only in concavity.]

REFERENCES

1. R. M. Dimitrić, Using less calculus in teaching calculus: An historical approach, this MAGAZINE **74** (2001) 201–211.
2. H. Dörrie, *100 Great Problems of Elementary Mathematics,* Dover, New York, 1965.

—Roger B. Nelsen
Lewis & Clark College
Portland, OR 97219

The Geometry behind Paradoxes of Voting Power

MICHAEL A. JONES
Mathematical Reviews
Ann Arbor, MI 48103
maj@ams.org

Anneliese, Brian, and Carlos among them own all 1500 shares of a small company's stock. At the annual stockholders' meeting, each stockholder's vote counts the same as the number of shares that he or she owns. A measure at the meeting passes if stockholders accounting for 2/3 of all shares support the measure. Because controversial measures about moving manufacturing to Asia and replacing the CEO of the company are to be voted on at the upcoming meeting, Brian and Carlos each buy 100 shares of stock from Anneliese in an effort to gain more influence. Brian finds that he has more influence than before, while Carlos discovers that his vote cannot affect the outcome on any measure, when before it could! How can this happen?

I will introduce *simple weighted-voting games* to model the stockholder scenario, and other voting situations, and *power indices* to measure the effect voters have on the outcome of yes/no elections. Power indices have been used to analyze the simple weighted-voting game models of the International Monetary Fund [**9, 19**], the Electoral College [**21**], the European Union Council of Ministers [**13, 18**], and the Israeli Knesset [**17**]. Power index calculations have also been used in the debate on the design of institutions, as in articles about the effects of reforms on, and the introduction of new members into, the European Union [**32, 33**].

Counterintuitive results such as Carlos' predicament are often called paradoxes. The literature on power indices is full of paradoxes, as well as real-life institutions that exhibit them. The paradox of redistribution [**9, 25**], the donor and transfer paradoxes [**10**], the paradox of quarreling members [**15**], the paradox of a new member [**3, 4**], and the paradox of large size [**3, 28**] capture diverse aspects of counterintuitive behavior in simple weighted-voting games. Using geometry, I will explain and classify the causes of voting power paradoxes. Surprisingly, 3-voter examples are sufficient to understand the geometry. The low dimension and the inherent symmetry of the 3-voter examples are often enough to prove that all power indices are susceptible to a particular paradox. Bradberry [**2**] used the same geometric approach to examine paradoxes of apportionment methods.

To place a simple weighted-voting game in a geometric setting, let the weights of the voters (the number of shares in the stockholder example) represent a point in Euclidean space. The voting rule (2/3 for the stockholder game) defines hyperplanes that partition the space into different parts or equivalence classes so that the power of each player is fixed for all games in an equivalence class; for the 3-voter examples, the hyperplanes are lines. Three types of geometric phenomena describe the changes in a game that may result in counterintuitive outcomes. A change in the weights of the players (as when Brian and Carlos buy additional shares) may cause a point to pass a hyperplane. A change in a voting rule (for example, if 2/3 were changed to 3/4) may affect the size and number of parts in the hyperplane partition. A change in the number of players (for example, if Anneliese were to sell all of her shares or to sell some shares to a 4th person) results in a projection to or from a boundary of the space of games with a specific number of players.

Simple weighted-voting games, geometry, and power

Much of the language in simple weighted-voting games comes from the context in which a measure, sometimes called a bill, proposal, legislation, or candidate, is being compared to a *status quo*. Simple weighted-voting games model situations in which yes/no votes are not treated equally by assigning weights, often nonnegative integers, to "yes" votes. If the sum of the weights of the "yes" votes matches or exceeds a threshold, then the measure is passed, replacing the *status quo*. In this case, the subset of "yes" voters is called a *winning coalition*. If the sum weight of the "yes" votes is less than the threshold, then the *status quo* is retained and the subset of "yes" voters is called a *losing coalition*.

The stockholder game, where the number of shares of stock are the weights, is one situation that can be modeled by simple weighted-voting games, but there are others. In the Electoral College game, the weights are explicit from the context: the weight of a state is the number of its electoral votes. Decisions in the United Nations Security Council (UNSC) treat votes of the permanent members (United States, United Kingdom, Russia, France, and China) differently than votes of the 10 nonpermanent members. Even though the UNSC does not specify weights, it is possible to derive weights from the description of winning coalitions, as described in [**27**].

The threshold of the simple weighted-voting game is called the *quota*. In the Electoral College, the quota is 270 electoral votes, a simple majority of the possible 538 electoral votes, the sum of the electoral votes of all states. In the UNSC game, the quota is determined in conjunction with the weights from the ways in which subsets of members of the UNSC form winning coalitions. The weights and the quota are connected and together define a simple weighted-voting game. If the quota is less than a simple majority of the sum of the weights of the voters, then it would be possible for two coalitions to pass conflicting legislation, so we eliminate that possibility. Further, a measure should pass if all voters support it. These two conditions ensure that the simple weighted-voting game is well defined.

Let $[q; w_1, \ldots, w_n]$ represent the simple weighted-voting game where voter $i \in N = \{1, \ldots, n\}$ has weight w_i assigned to its "yes" votes and q is the quota. In this notation, the two conditions for a simple weighted-voting game to be well defined are $w/2 < q \leq w = w_1 + \cdots + w_n$. Let the value function v indicate whether a coalition $S \subseteq N$ is losing or winning by

$$v(S) = \begin{cases} 0 & \text{if } \sum_{i \in S} w_i < q, \\ 1 & \text{if } \sum_{i \in S} w_i \geq q. \end{cases}$$

It is often useful to keep track of the *minimal winning coalitions*, winning coalitions for which no proper subset is also winning. For a simple weighted-voting game, a voter influences the outcome of an election only if it is a member of a minimal winning coalition. This follows because if all other voters in a minimal winning coalition vote "yes," then the outcome of the election hinges on the yes/no vote of the remaining voter.

Simple weighted-voting games are a special case of the larger class of simple games. A *simple game* is simply a list of winning coalitions, subject to the requirements that the complement of a winning coalition is never a winning coalition, a superset of a winning coalition is always a winning coalition, and the set of all players is a winning coalition. Taylor and Zwicker [**31**] provided a vote-trading condition to determine whether or not a simple game can be represented as a simple weighted-voting game. I restrict attention to simple weighted-voting games because the geometric approach provides insight to the causes of paradoxes of voting power. These paradoxes

arise even in the most easily understood voting systems and do not require the added generality of simple games.

Stockholder game To flesh out the stockholder game from the introduction, assume that Anneliese, Brian, and Carlos have 800, 500, and 200 shares of stock, respectively. The quota is 2/3 of the 1500 shares or 1000 shares. The stockholder game is given by [1000; 800, 500, 200], where Anneliese, Brian, and Carlos are voters (or players) 1, 2, and 3, respectively. The minimal winning coalitions are $\{1, 2\}$ and $\{1, 3\}$, because $w_1 + w_2 \geq q$ and $w_1 + w_3 \geq q$. If player 1 votes "yes" and player 2 votes "no," then player 3 determines whether or not a measure passes. Despite having only 200 shares of stock, how Carlos votes may affect the outcome. Whenever a voter is a member of a minimal winning coalition, then his or her vote may make a difference.

Suppose that Brian and Carlos each buy 100 shares of stock from Anneliese. With the players ordered as before, the resulting game is [1000; 600, 600, 300]. Even though Carlos increases his number of shares to 300, now his vote cannot change the outcome of an election, regardless of how Anneliese and Brian vote. Because the quota is only reached if Anneliese and Brian vote in the affirmative, $\{1, 2\}$ is the sole minimal winning coalition. As before, $w_1 + w_2 \geq q$, but now, $w_1 + w_3 < q$, flipping the inequality. I will return to this example and will view the change in the relationship between $w_1 + w_3$ and the quota q from a geometric perspective.

Geometry of simple weighted-voting games We can normalize a simple weighted-voting game $[q; w_1, \ldots, w_n]$ by dividing by w to yield $[q/w; x_1, \ldots x_n]$ where $x_i = w_i/w$. The weights of normalized games are viewed as points on the $(n-1)$-dimensional simplex

$$S_{n-1} = \{(x_1, \ldots, x_n) \mid x_1 + \cdots + x_n = 1 \text{ and } x_i \geq 0 \text{ for all } i\}.$$

For each $S \subseteq N$, the hyperplane of the form $\sum_{i \in S} x_i = q/w$ divides the simplex into two regions: those games for which S is a winning coalition and those games for which S is a losing coalition. The collection of hyperplanes partitions the simplex into parts whose number and size depend on the normalized quota q/w.

Fortunately, only 3-voter examples are necessary to understand the paradoxes of voting power. For a game with 3 players, the normalized weights of the 3 players are viewed as a point on the 2-simplex, which is the intersection of the plane $x_1 + x_2 + x_3 = 1$ and the nonnegative octant, where $x_i \geq 0$ for all i. This forms the equilateral triangle shown in FIGURE 1 with a coordinate system in which each x_i^* measures the perpendicular distance from the point (x_1, x_2, x_3) to one of the sides of the triangle. An important consequence of this geometry is that one x_i is constant on any segment parallel to a side of the triangle and $x_i^* = \sqrt{3/2}\, x_i$. Because $x_1 + x_2 + x_3 = 1$, it follows that the sum of the x_i^*s is constant and equal to the height of the equilateral triangle, a result known as Viviani's Theorem. Three Proofs Without Words of this result have appeared in the MAGAZINE [**24, 30, 34**].

When considering all possible normalized simple weighted-voting games in the simplex S_2 for a fixed normalized quota $q \in (1/2, 1]$, the hyperplanes (or lines)

$$\begin{array}{ccc} x_1 = q & x_2 = q & x_3 = q \\ x_1 + x_2 = q & x_1 + x_3 = q & x_2 + x_3 = q \end{array}$$

partition the simplex. The hyperplane $x_i = q$ defines the segment $x_i^* = \sqrt{3/2}\, q$ parallel to one of the sides of the equilateral triangle. Because $x_1 + x_2 + x_3 = 1$, then

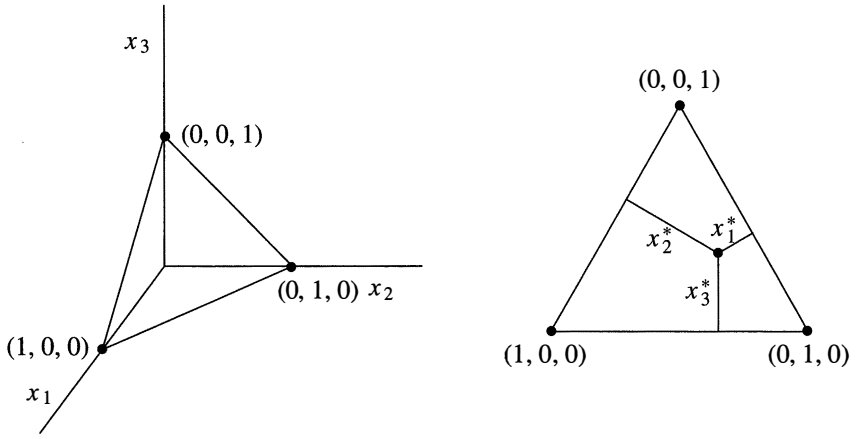


Figure 1 The plane $x_1 + x_2 + x_3 = 1$ and the 2-simplex $\{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 1, x_1 \geq 0, x_2 \geq 0, \text{and } x_3 \geq 0\}$ with distances $x_i^* = \left(\sqrt{3/2}\right) x_i$

$x_1 + x_2 = q$, $x_1 + x_3 = q$, and $x_2 + x_3 = q$ are equivalent to $x_3 = 1 - q$, $x_2 = 1 - q$, and $x_1 = 1 - q$, respectively. For $q < 1$, the hyperplanes define 6 line segments parallel to the sides of the equilateral triangle and partition the simplex into ten regions $R_1 - R_{10}$ (FIGURE 2 and TABLE 1). If a game lies *on* a hyperplane, its region can be determined from the winning coalitions in TABLE 1. For example, a game on the part of $w_2 = 1 - q$ between R_6 and R_7 belongs to R_7, because $w_1 + w_3 = q$ ensures that $\{1, 3\}$ is a winning coalition. When $q = 1$, R_{10} is the entire interior of the simplex because regions R_4, R_5, and R_6 become line segments, R_1, R_2, and R_3 reduce to points, and R_7, R_8, and R_9 are empty.

In regions R_1, R_2, and R_3, a *dictator* forms a singleton winning coalition, which is necessarily minimal, as seen in TABLE 1. As the name implies, a dictator's vote decides the outcome of the election because its weight is at least as large as the quota. A voter is called a *dummy voter* if he or she is not part of any minimal winning coalition. For example, in region R_6, voter 3 is a dummy voter, just like Carlos after he buys the additional shares of stock, because $3 \notin \{1, 2\}$, the only minimal winning coalition (as in TABLE 1).

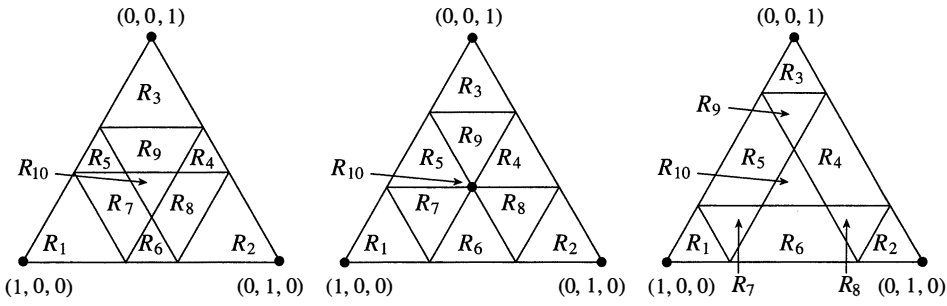


Figure 2 Shape of regions for various normalized quotas: $1/2 < q < 2/3$ (left), $q = 2/3$ (middle), and $2/3 < q < 1$ (right)

The games in each region form an equivalence class in which each game has the same set of minimal winning coalitions. Although the hyperplanes determine the minimal winning coalitions, it is useful to work backwards from the minimal winning coalitions of a specific equivalence class to understand the relationship to the hyperplanes.

For example, games in R_7 from FIGURE 2 have minimal winning coalitions $\{1, 2\}$ and $\{1, 3\}$. These coalitions not only ensure that $x_1 + x_2 \geq q$ and $x_1 + x_3 \geq q$, but also that $x_1 < q$, $x_2 < q$, $x_3 < q$ and $x_2 + x_3 < q$. Because $x_1 + x_2 + x_3 = 1$, the inequalities that involve sums are rewritten as $x_3 \leq 1 - q$, $x_2 \leq 1 - q$, and $x_1 > 1 - q$. It follows that a point (x_1, x_2, x_3) is in R_7 if and only if $1 - q < x_1 < q$, $x_2 < 1 - q < q$ and $x_3 < 1 - q < q$. A point's relationship to the 6 hyperplanes places it into one of the equivalence classes.

TABLE 1: Regions and their corresponding winning and minimal winning coalitions

Region	Winning Coalitions	Minimal Winning Coalitions
R_1	$\{1\}, \{1, 2\}, \{1, 3\}, \{1, 2, 3\}$	$\{1\}$
R_2	$\{2\}, \{1, 2\}, \{2, 3\}, \{1, 2, 3\}$	$\{2\}$
R_3	$\{3\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}$	$\{3\}$
R_4	$\{2, 3\}, \{1, 2, 3\}$	$\{2, 3\}$
R_5	$\{1, 3\}, \{1, 2, 3\}$	$\{1, 3\}$
R_6	$\{1, 2\}, \{1, 2, 3\}$	$\{1, 2\}$
R_7	$\{1, 2\}, \{1, 3\}, \{1, 2, 3\}$	$\{1, 2\}, \{1, 3\}$
R_8	$\{1, 2\}, \{2, 3\}, \{1, 2, 3\}$	$\{1, 2\}, \{2, 3\}$
R_9	$\{1, 3\}, \{2, 3\}, \{1, 2, 3\}$	$\{1, 3\}, \{2, 3\}$
R_{10} for $q \leq 2/3$	$\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}$	$\{1, 2\}, \{1, 3\}, \{2, 3\}$
R_{10} for $q > 2/3$	$\{1, 2, 3\}$:	$\{1, 2, 3\}$

Measuring the power of voters in simple weighted-voting games The weight of a voter coarsely measures how important, or how much power, an individual brings to a coalition. More refined measures of power—things called *power indices*—calculate a player's contribution to a political process and are used to determine the fairness of different political institutions under different assumptions about how coalitions form, the size of coalitions, a player's role in changing a coalition from losing to winning or winning to losing, etc. As opposed to looking at how individuals vote on a particular issue, power indices measure *a priori* power, due to the structure of the institution. This power is determined by which coalitions are winning or losing, assuming a distribution over all possible ways in which the voters may vote.

There are many specialized power indices (as introduced in [**1, 6, 8, 22, 29**]) that measure different aspects of power. But, in general, for a normalized quota $q \in (1/2, 1]$, a power index P_q is a discrete map from the $(n-1)$-simplex of normalized n-player, simple weighted-voting games to vectors in $\mathbb{R}^n$ where the ith entry of the vector represents the power of the ith player. If all voters, including i, in a coalition S vote "yes," then voter i influences the election outcome when $v(S) - v(S/\{i\}) = 1$. Summing over all possible coalitions of voters, the power of voter i is

$$P_q(\mathbf{x})_i = \sum_{S \subseteq N} \lambda_S [v(S) - v(S/\{i\})]$$

where the λ_S coefficients depend on the specific power index used. (Notice that if $i \notin S$, then $v(S) - v(S/\{i\}) = 0$.) Regardless of the definition of the coefficients λ_S, the geometry of the domain and the hyperplanes that slice the simplex into parts (that indicate the winning and losing coalitions) are the same, creating equivalence classes

of games with the same power. Often, the power $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ for a simple weighted-voting game is normalized so that $r_1 + \cdots + r_n = 1$ and is written as $r_1 : r_2 : \cdots : r_n$.

Two commonly used power indices that are taught in general education courses at the undergraduate level [**7**] and sometimes used in upper level courses [**12, 16**] are the Banzhaf and Shapley-Shubik power indices. In 1965, Banzhaf introduced his power index in a lawsuit while arguing that voting among the Nassau County (NY) Board of Supervisors was not fair [**1**]. For the Banzhaf power index, $\lambda_S = 1$ for all S. Hence, the Banzhaf index counts the number of times that a voter is necessary to be part of a coalition for a measure to pass. This necessary voter is referred to as a critical voter.

The Shapley-Shubik power index [**29**] applies the Shapley value, a solution concept from cooperative game theory [**26**], to simple weighted-voting games. For the Shapley-Shubik power index, $\lambda_S = (|S| - 1)!$ depends on the number of voters in S. Intuitively, the Shapley-Shubik power index measures the power of a voter given every sequence of "yes" votes. If $v(S) - v(S/\{i\}) = 1$, then the $|S|$ voters could join the coalition in any order. In $(|S| - 1)!$ of these orders, voter i joined the coalition last and changed the coalition from losing to winning. In this case, voter i is referred to as a pivotal voter.

For 3 voters, the Banzhaf and Shapley-Shubik power indices may assign different normalized powers to the same game, but they always agree on the relative ranking of the voters' powers [**23**]. It follows from the definitions of a dummy voter and a dictator that the power of a dummy is 0 while the normalized power of a dictator is 1. To get a better sense of how to calculate power for a 3-voter game, we return to the stockholder game.

Measuring the power of the stockholders The stockholder game in the introduction normalizes to $[10/15; 8/15, 5/15, 2/15]$. The 3 winning coalitions $\{1, 2, 3\}$, $\{1, 2\}$, and $\{1, 3\}$ yield the following positive differences, all equal to 1: $v(\{1, 2, 3\}) - v(\{2, 3\})$, $v(\{1, 2\}) - v(\{2\})$, $v(\{1, 3\}) - v(\{3\})$, $v(\{1, 2\}) - v(\{1\})$, and $v(\{1, 3\}) - v(\{1\})$.

From the general definition of a power index, the power associated with the stockholder game is

$$P_{10/15}\left(\frac{8}{15}, \frac{5}{15}, \frac{2}{15}\right) = \left(\lambda_{\{1,2,3\}} + \lambda_{\{1,2\}} + \lambda_{\{1,3\}}, \lambda_{\{1,2\}}, \lambda_{\{1,3\}}\right).$$

It follows that the Banzhaf power index for the stockholder game is $3 : 1 : 1$ because $\lambda_S = 1$ for all S. And, the normalized Banzhaf power index is $3/5 : 1/5 : 1/5$. The coefficients of the nonzero differences in the Shapley-Shubik power index are $\lambda_{\{1,2,3\}} = (3 - 1)! = 2$, $\lambda_{\{1,2\}} = (2 - 1)! = 1$ and $\lambda_{\{1,3\}} = (2 - 1)! = 1$. The Shapley-Shubik power index for the stockholder game is $4 : 1 : 1$, which normalizes to $4/6 : 1/6 : 1/6$. Notice that the Banzhaf and Shapley-Shubik power indices agree that voter 1 has the most power while voters 2 and 3 have equal power less than voter 1.

Geometry of paradoxes of voting power

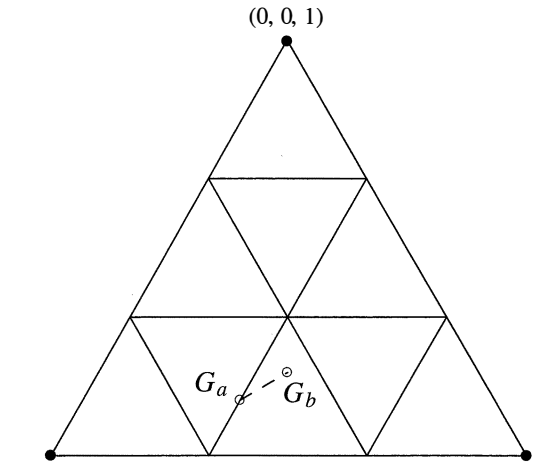
Because of the hyperplane partition, a game in the interior of a part or equivalence class may not change parts under a small perturbation of its weights. Only upon passing a hyperplane would we expect to see the power change. Keeping the weights fixed and changing the quota will result in the hyperplanes shifting, changing the size, shape, and number of equivalence classes. Adding or subtracting voters changes the dimension of

the simplex, as well as the possible power outcomes on the parts. These changes in geometry may result in counterintuitive behavior.

Domain effects The stockholder game of the introduction is an example of the *paradox of redistribution* in which a voter's weight increases and its power decreases or a voter's weight decreases and its power increases. The geometry of the simplex readily explains how the paradox arises. Because simple weighted-voting games are domain points on the simplex, a change in the weight of one voter (or coordinate) must be met with changes in the weight of at least one other voter, too. In fact, the paradox stems from the words "at least" in the last sentence. Often times, changes involve more than 2 players; the case of only 2 players is discussed later.

The paradox of redistribution was first noted by Fischer and Schotter [**11**]. Schotter [**25**] used simplices to determine the likelihood of the paradox for the Banzhaf and Shapley-Shubik power indices. However, the paradox is not an artifact of the particular power index used.

More is not always better: Carlos buys more shares in the stockholder game Recall that the normalized stockholder game from the introduction [10/15; 8/15, 5/15, 2/15], or G_a in FIGURE 3, has 3/5 : 1/5 : 1/5 as its normalized Banzhaf index. When Anneliese sells 100 shares to Brian and 100 shares to Carlos, the resulting game [1000; 600, 600, 300] normalizes to G_b = [10/15; 6/15, 6/15, 3/15], shown in FIGURE 3. (The nonreduced fractions allow the games to be compared easily.) For G_b, the normalized Banzhaf index is 1/2 : 1/2 : 0. (This follows from symmetry: Since {1, 2} is the only minimal winning coalition, voters 1 and 2 each have power $\lambda_{\{1,2\}} + \lambda_{\{1,2,3\}}$, before the power is normalized. Hence, voters 1 and 2 split the power equally.) FIGURE 3 shows how redistributing the weights from G_a to G_b passes a hyperplane with G_a on the border of region R_7 and G_b in region R_6 (with the regions defined as in FIGURE 2 and TABLE 1), resulting in an increase of power for



$G_a = [\frac{10}{15}; \frac{8}{15}, \frac{5}{15}, \frac{2}{15}]$ with normalized Banzhaf index $\frac{3}{5} : \frac{1}{5} : \frac{1}{5}$

$G_b = [\frac{10}{15}; \frac{6}{15}, \frac{6}{15}, \frac{3}{15}]$ with normalized Banzhaf index $\frac{1}{2} : \frac{1}{2} : 0$

Figure 3 The paradox of redistribution (in the stockholder example) as an effect of passing a hyperplane: When $G_a \rightarrow G_b$, player 3's weight increases but its power decreases. When $G_b \rightarrow G_a$, player 3's weight decreases but its power increases.

Brian and a decrease in power for Carlos. The paradoxical behavior present in the stockholder game holds for any power index (with positive λ_S): Although Carlos increases the number of his shares, he becomes a dummy voter because the weights of the other voters change too.

For 3-voter simple weighted-voting games, it is not possible to have both a player's weight increase and power decrease, and another player's weight decrease and power increase. However, this may happen in a 4-voter game, as this next example demonstrates. Further, the weights of the other two voters in the example change in a way that matches our intuition: one player has both her weight and power increase, while another player has his weight and power decrease. This shows that anything can happen! The effect of a change of the weight of a player depends on how the weights of other players change, too.

Anything can happen! The game [70; 40, 30, 20, 10] has normalized Banzhaf power index of 0.5 : 0.3 : 0.1 : 0.1. Suppose that players 1 and 4 have their weights increase to 46 and 18, respectively, while players 2 and 3 have their weights both decrease to 18. The resulting game is [70; 46, 18, 18, 18], which has 0.4 : 0.2 : 0.2 : 0.2 as its normalized Banzhaf power index. The changes to the weights of players 2 and 4 confirm our intuition: an increase (decrease) in weight resulted in an increase (decrease) in power. The changes to the weights of players 1 and 3 defy our intuition: an increase (decrease) in weight resulted in a decrease (increase) in power.

Felsenthal and Machover [**10**] considered an even more counterintuitive version of the paradox of redistribution called the *donation paradox*. They showed that if the power index doesn't satisfy a monotonicity condition, then it is possible for a voter to donate some of its weight to another voter (while all other weights remain the same) and the donor's power increases while the recipient's power decreases! Although this requires a nonmonotone power index, the geometry behind the paradox remains the same: a perturbation in the weights of the players causes the game to pass a hyperplane.

Partition effects So far we have considered the effect of changing the weights of the game. However, it is possible to achieve paradoxical outcomes by leaving the weights fixed and changing the quota. In general, the quota affects the size and number of parts in the partition of the simplex. FIGURE 2 demonstrates how the geometry of the parts change as the quota changes for 3-voter games. Of course, changing the quota may have consequences for institutions. For example, Dreyer and Schotter [**9**] considered the effect of changing the quota for the International Monetary Fund.

We might expect that lowering the quota benefits the voter with the largest weight. Winning coalitions from before will be retained. However, the critical voters may change. And, new winning coalitions may form. Perhaps lowering the quota hurts the voter with the largest weight! Because decreasing the quota does not always benefit the voter with the largest weight, I refer to this as the *quota paradox*. The following example demonstrates two scenarios in which the same weights are used to show how the quota affects the voter with the largest weight.

Changing the voting rule may be good or bad: the quota paradox Consider the effect on power as q decreases from 8/11 to 7/11 to 6/11 for the game $[q; 5/11, 4/11, 2/11]$. Under the Shapley-Shubik power index, these three games have power indices 1/2 : 1/2 : 0, 4/6 : 1/6 : 1/6, and 1/3 : 1/3 : 1/3, respectively. The voter with the largest weight initially benefits from a decrease

in the quota, but a further decrease in the quota lowers the power of voter 1. FIGURE 4 depicts this quota paradox.

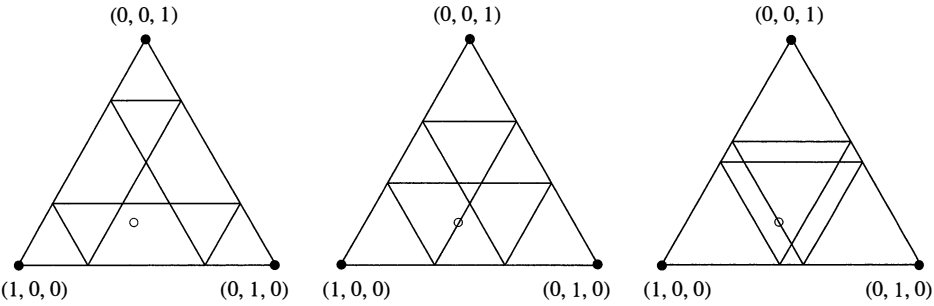


Figure 4 As the quota decreases from $q_1 = 8/11$ (left), $q_2 = 7/11$ (middle), and $q_3 = 6/11$ (right) games with weights 5/11, 4/11, 2/11 lie in regions R_6, R_7, and R_{10}, respectively

There are other ways to adjust the size and number of parts of the partition of the simplex. Kilgour [**15**] introduced the *paradox of quarreling members* in which quarreling members' powers may increase despite restricting the coalitions they may enter together. Specifically, if two voters quarrel, they will never both vote "yes" on a measure. Even though they refuse to be part of the same winning coalition, it is possible that the power of one of these voters increases. This is paradoxical because restricting the coalitions that can form hurts the quarreling members by restricting their freedom and decreasing their options. However, the restriction also eliminates options for the nonquarreling members, possibly making the quarreling members more powerful. The two quarreling members are not assumed to always be on opposite sides of a vote; they may both vote "no."

Measuring the power of the voters when certain coalitions cannot form requires modifications of the power indices. Clearly, certain sequences of voters necessary to compute the Shapley-Shubik power index would be impossible, as they would require quarreling members both to vote in the affirmative. The following example demonstrates how quarreling members reduces the number of regions in the partition. After the geometric interpretation, I provide an alternative view to compute the power index for a game with quarreling members.

The paradox of quarreling members: Fewer options may be better Suppose voters 1 and 3 quarrel. This means that coalitions {1, 2, 3} and {1, 3} will never be winning coalitions. In particular, because {1, 3} cannot form, then the hyperplane $x_1 + x_3 = q$ (or equivalently, $x_2 = 1 - q$) is not used to partition the simplex. Hence, when voters 1 and 3 quarrel, the simplex is partitioned into fewer regions. For a normalized quota of 3/4, the simplex is divided into 7 regions or equivalence classes, $Q_1 - Q_7$, as in FIGURE 5.

Consider the simple weighted-voting game [3/4; 1/4, 2/4, 1/4]. When voters 1 and 3 do not quarrel, this game is on the boundary of R_8 from FIGURE 2 and has a normalized Banzhaf power index of 1/5 : 3/5 : 1/5. When voters 1 and 3 do quarrel, then the only winning coalitions are {1, 2} and {2, 3}, because {1, 2, 3} cannot form. Voter 2 is critical twice while voters 1 and 3 are each critical once. Under the normalized Banzhaf power index, the game with quarreling has a power index of 1/4 : 1/2 : 1/4. In this case, despite having fewer options, both voters 1 and 3 had their powers increase because of the quarrel.

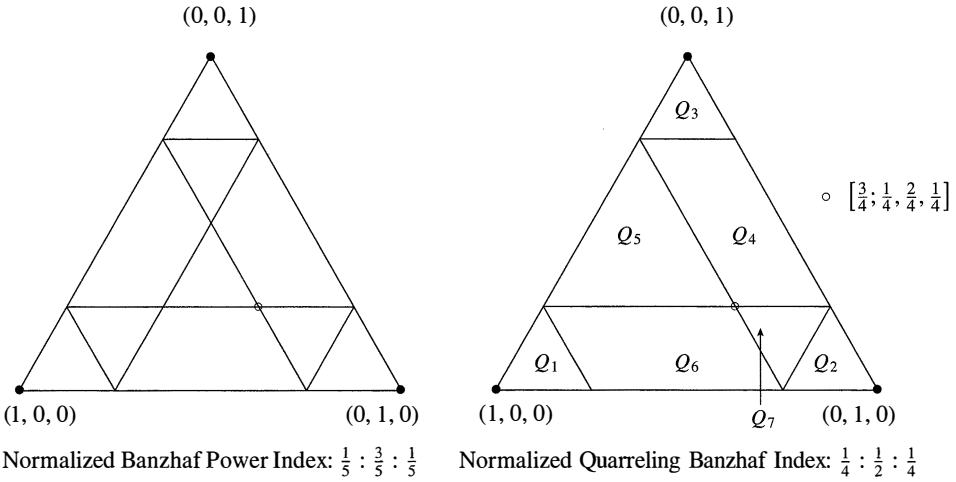


Figure 5 The paradox of quarreling members: Players 1 and 3 quarrel resulting in the removal of a hyperplane that decreases the number of parts in the partition

Another way to view the power index computation for a game with quarreling members i and j is to set $\lambda_S = 0$ if both i and j are in S. In the example, $\{1, 2, 3\}$ could still be considered a winning coalition, but $\lambda_{\{1,2,3\}}$ would be 0, giving the term $v(\{1, 2, 3\}) - v(\{1, 3\}) = 1$ no weight in the power calculation.

Dimensional effects Rich, and often paradoxical, behavior exists when voters are introduced or removed from a game. Because it is too difficult to visualize the geometry of a 4-voter game, I will focus on changing the number of voters between 2 and 3. The space of all 2-voter games is a 1-simplex or interval $I = \{(w_1, w_2) \mid w_1 + w_2 = 1 \text{ and } w_i \geq 0 \text{ for both } i\}$ with left endpoint (1, 0) and right endpoint (0, 1), as in FIGURE 6. For any quota q, the unit interval is partitioned into 3 regions in which the minimal winning coalitions are: $\{1\}$, $\{1, 2\}$ or $\{2\}$; these are the minimal winning coalitions for the equivalence classes from left-to-right on the unit interval in FIGURE 6. Power for 2-voter games is easily calculated because voter i is a dictator if $\{i\}$ is the only minimal winning coalition and both players share the power equally if the only minimal winning coalition is $\{1, 2\}$.

The first paradox focuses on the effect of the introduction of a new voter. The *paradox of a new member* was introduced by Brams and Affuso [**3, 4**]. They introduce a new voter into the game while keeping the relative weights of the other voters constant; that is, the weights of the original voters are proportional. Felsenthal and Machover [**10**] explained this paradox in a particularly succinct way. The paradox occurs when the game $[q; u_1, u_2, \ldots, u_n]$ changes to $[q; v_1, v_2, \ldots, v_n, v_{n+1}]$ where $v_{n+1} \in [0, 1]$ and $v_i = (1 - v_{n+1})u_i$ for $i = 1$ to n and the power of one of the original n voters increases. This seems paradoxical because introducing the new voter would seem to take power away from the other voters, but the following example shows otherwise.

Power may go up with more people sharing: paradox of a new member Consider the 2-voter game [0.75; 0.7, 0.3]. FIGURE 6 shows the partition of the 1-dimensional simplex into regions of games that have the same power. Clearly, the power under any index with $\lambda_{\{1,2\}} > 0$ is 1/2 : 1/2 as both voters are necessary for a coalition to be winning. The line in FIGURE 6 shows the possible games in which a third voter is added while keeping the ratio of the weights of voters 1 and 2 constant. For all the games on this line within region R_7, the power

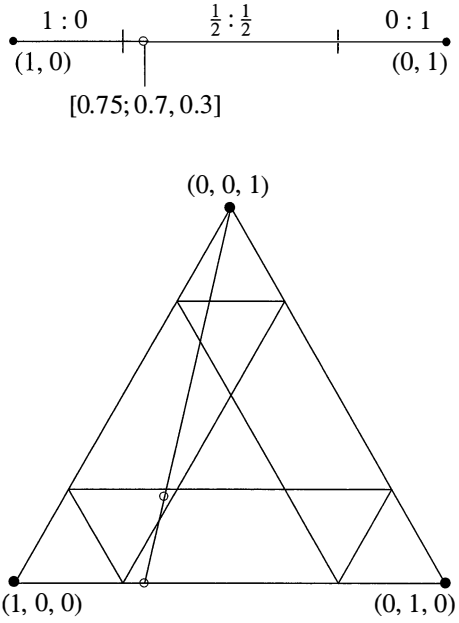


Figure 6 The paradox of a new member: Voter 1's power increases despite the introduction of a new voter

of voter 1 increases. As a representative game on the line and in R_7, consider [0.75; 7/13, 3/13, 3/13]. Under the normalized Shapley-Shubik power index, the power of voter 1 is 4/6 in the new 3-voter game while under the normalized Banzhaf power index the power of voter 1 is 3/5; both are greater than 1/2.

The paradox of a new member demonstrates that adding a voter to an organization may have unanticipated consequences. Researchers have applied power indices to see the effect of proposed expansion of the European Union [**32, 33**].

The *paradox of large size* considers another paradoxical outcome: that a voter who annexes another voter (and absorbs its weight) may have less power than the two voters would have if the merger did not occur. As we see in the next example, 3 voters are sufficient to demonstrate that this paradox is independent of the measurement of power.

Bigger and fewer is not always better: the paradox of large size By symmetry, the power of each player for the simple weighted-voting game $G_a = [3/4; 1/3, 1/3, 1/3]$ is 1/3. If voter 1 receives the entirety of the weight of voter 3, then the resulting game is $G_b = [3/4; 2/3, 1/3]$, as shown in FIGURE 7. As both voters are necessary to form a winning coalition, the resulting power is 1/2 for each of them. The sum of the power of the first two voters in G_a is 2/3 while the power of the merged players (player 1) in G_b has dropped to 1/2.

The paradox of large size is comparable to eliminating a voter and distributing its weight to one of the voters. Saari and Sieberg [**23**] showed that complete reversals of the power rankings of voters may occur when adding or subtracting a voter. The subtraction of a voter may be viewed as a projection from the $(n + 1)$-voter simplex to the n-voter simplex. Because there are many such projections which result in different powers in the projected games, it is not surprising that paradoxical outcomes may occur.

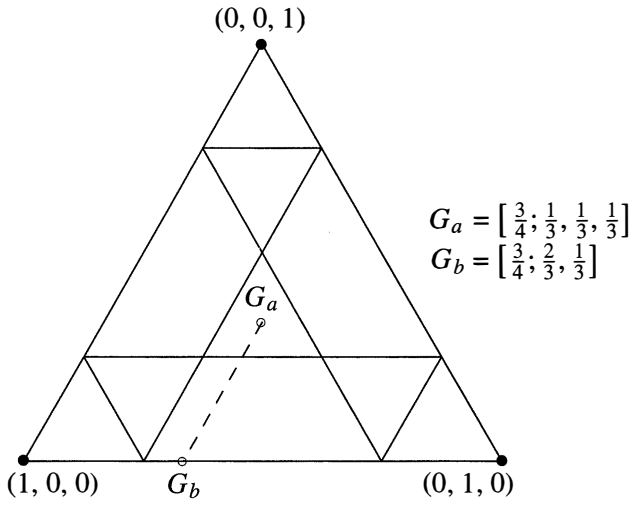


Figure 7 The paradox of large size: Player 3 coalesces with Player 1 and their cumulative power decreases

Combining geometric elements Additional paradoxes may combine more than one geometric element. Felsenthal and Machover [**10**] introduced the *fattening paradox* in which the weight of one voter is increased, while the (nonnormalized) weights of the other voters remain fixed, but the (un)lucky recipient of the extra weight has her power decrease. This is comparable to changing the position in the simplex of the weights at the same time as changing the quota. The following example demonstrates how two geometric elements combine in the fattening paradox.

> **Bigger may not be better, even if everyone else's weight stays the same: the fattening paradox** Suppose we start with the game $G_a = [8; 4, 4, 1, 1, 1]$ and increase the weight of voter 1 from 4 in to 5, producing the game $G_b = [8; 5, 4, 1, 1, 1]$ [**10**]. Under the normalized Banzhaf power index, the power of voter 1 is $1/2$ in G_a (due to symmetry, as voters 1 and 2 are the only two critical voters). In G_b, the power of voter 1 decreases to approximately 0.474 under the normalized Banzhaf power index. The power was calculated using the algorithms on Leech's website [**20**].

This paradox combines two geometric properties. Not only have the weights been changed, but the normalized quota has changed from $8/11$ to $8/12$. The decrease of the normalized quota changes the size and possibly even the number of parts in the partition. Geometrically, this may be viewed as moving the hyperplane while also redistributing the normalized weights. These two actions cause the game to pass a hyperplane. Writing our example in normalized form gives

$$\left[\frac{8}{11}; \frac{4}{11}, \frac{4}{11}, \frac{1}{11}, \frac{1}{11}, \frac{1}{11}\right] \to \left[\frac{8}{12}; \frac{5}{12}, \frac{4}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}\right].$$

The fattening paradox may be described for a more general, n-voter game. For example, if the weight of voter 1 increases by k from x_1 to $x_1 + k$, then the game changes from $[q; x_1, x_2, \ldots, x_n]$ to $[q; x_1 + k, x_2, \ldots, x_n]$. Normalizing gives

$$\left[\frac{q}{X}; \frac{x_1}{X}, \frac{x_2}{X}, \ldots, \frac{x_n}{X}\right] \to \left[\frac{q}{X+k}; \frac{x_1+k}{X+k}, \frac{x_2}{X+k}, \ldots, \frac{x_n}{X+k}\right]$$

where $X = \sum_{i=1}^{n} x_i$. Geometrically, the normalized quota has decreased from q/X to $q/(X + k)$, changing the hyperplanes, at the same time as the game moves proportionally in the direction of the $(1, 0, \ldots, 0)$-vertex of the simplex. Moving proportionally in the direction of a vertex is a process comparable to adding or subtracting a player. Hence, the fattening paradox has elements of each of the geometric properties.

Conclusion

The geometry that arises from the partition on the simplex of simple weighted-voting games is a natural way to classify paradoxical outcomes in voting power. Although the names of the paradoxes do not indicate the geometry behind the paradox, three geometric properties: passing a hyperplane, altering the number and/or size of the parts of a partition, and projecting to or from a boundary are the building blocks for the paradoxes. Not only does the geometry provide a tool to analyze paradoxes, but also a tool to construct new ones. Concentrating on paradoxes in low dimensions helps to visualize the geometry of the paradoxes. Of course, paradoxes occur in higher dimensions, too. Leech [**20**] provides online access to algorithms to compute power indices for large games, which will help you to create your own paradoxes for higher dimensional games.

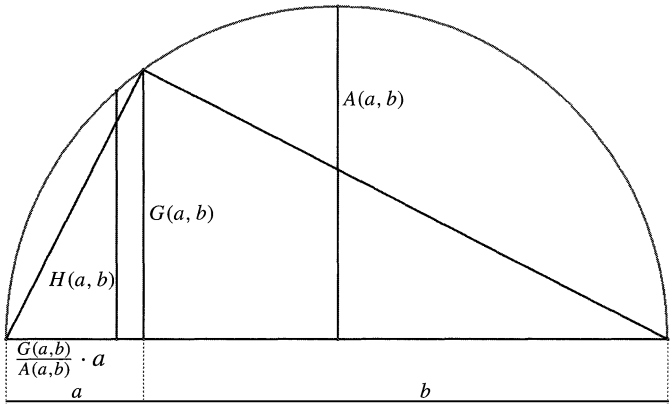
Acknowledgment. An early version of this article appeared in a conference proceedings [**14**]. I thank Don Saari, Jennifer Wilson, and participants of the DIMACS-LAMSADE Workshop on Voting Theory and Preference Modeling and the DIMACS Re-connect Conference on Mathematics of Elections and Decisions for helpful comments.

REFERENCES

1. J. Banzhaf, Weighted voting doesn't work; a mathematical analysis, *Rutgers Law Review* **19** (1965) 317–343.
2. B. A. Bradberry, A geometric view of some apportionment paradoxes, this MAGAZINE **65** (1992) 3–17.
3. S. J. Brams, *Game Theory and Politics*, The Free Press, New York, 1975.
4. S. J. Brams and P. J. Affuso, Power and size: a new paradox, *Theory and Decision* **7** (1976) 29–56.
5. S. J. Brams and P. J. Affuso, New paradoxes of voting power on the EC Council of Ministers, *Electoral Studies* **4**(2) (1985) 135–139.
6. J. S. Coleman, Control of collectivities and the power of a collectivity to act, pp. 269–300 in *Social Choice*, B. Lieberman, ed., Gordon and Breach, New York, 1971.
7. COMAP [Consortium for Mathematics and Its Applications], *For All Practical Purposes: Mathematical Literary in Today's World*, 5th ed., W. H. Freeman, New York, 2000.
8. J. Deegan and E. W. Packel, To the (minimal winning) victors go the (equally divided) spoils: A new power index for simple n-person games, pp. 239–255 in S. J. Brams, W. F. Lucas, and P. D. Straffin, eds., *Political and Related Models*, Springer Verlag, New York, 1982.
9. J. S. Dreyer and A. Schotter, Power relationships in the International Monetary Fund: The consequences of quota changes, *Review of Economic Studies* **62** (1980) 97–106.
10. D. S. Felsenthal and M. Machover, *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes*, Edward Elgar Publishers, 1998.
11. D. Fischer and A. Schotter, The paradox of redistribution in the allocation of voting weights, *Public Choice* **33** (1978) 49–67.
12. M. J. Haines and M. A. Jones, Integrating combinatorics, geometry, and probability through the Shapley-Shubik power index, pp. 143–162 in D. Maher, ed., *Innovative Methods in Classes Beyond Calculus*, Mathematical Association of America, Washington, DC, 2005.
13. R. J. Johnston, The conflict over qualified majority voting in the European Union Council of Ministers: An analysis of the UK negotiating stance using power indices, *British Journal of Political Science* **25** (1995) 245–288.
14. M. A. Jones, The geometry behind paradoxes of voting power, Proceedings of the DIMACS-LAMSADE Workshop on Voting Theory and Preference Modeling, *Annales du LAMSADE* **6** (2006) 193–209.
15. D. M. Kilgour, A Shapley value for cooperative games with quarreling, in *Game Theory as a Theory of Conflict Resolution*, A. Rapoport, ed., D. Reidel, Holland, 1974.
16. J. P. Lampert, Voting games, power indices, and presidential elections, *UMAP Module* (1988) 144–197.
17. A. Laruelle, Implementing democracy in indirect voting processes: the Knesset case," in *Power Indices and Coalition Formation*, M. J. Holler and G. Owen, eds., Kluwer Academic Publishers, Boston, 2001.

18. D. Leech, Designing the voting system for the Council of the European Union, *Public Choice* **113** (2002) 437–464.
19. D. Leech, Voting power in the governance of the International Monetary Fund, *Annals of Operations Research* **103** (2002) 375–397.
20. D. Leech, `http://www.warwick.ac.uk/~ecaae/` (2008).
21. I. Mann and L. S. Shapley, The *a priori* voting strength of the Electoral College, pages 151–164 in *Game Theory and Related Approaches to Social Behavior*, M. Shubik, ed., Wiley, New York, 1964.
22. L. S. Penrose, The elementary statistics of majority voting, *Journal of the Royal Statistical Society* **109** (1946) 53–57.
23. D. G. Saari and K. K. Sieberg, Some surprising properties of power indices, *Games and Economic Behavior*, **36**(2) (2001) 241–263.
24. H. Samelson, Proof without words: Viviani's Theorem with vectors, this MAGAZINE **76** (2003) 225.
25. A. Schotter, The paradox of redistribution: some theoretical and empirical results, in *Power, Voting, and Voting Power*, M. J. Holler, ed., Physica, Würzburg, Germany, 1981.
26. L. S. Shapley, A value for n-person games, in *Contributions to the Theory of Games II*, Annals of Mathematics Study 28, H. Kuhn and A. Tucker, eds., Princeton University Press, Princeton, NJ, 1953.
27. L. S. Shapley, Simple games: an outline of the descriptive theory, *Behavioral Science* **7** (1962) 59–66.
28. L. S. Shapley, Political science: voting and bargaining games, in *Notes of Lectures on Mathematics in Behavioral Sciences*, H. A. Selby, ed., Mathematical Association of America, Williamstown, MA, 1973.
29. L. S. Shapley and M. Shubik, A method for evaluating the distribution of power in a committee system, in *Game Theory and Related Approaches to Social Behavior*, M. Shubik, ed., Wiley, London, 1954.
30. J. Tanton, Proof without words: Equilateral triangle, this MAGAZINE **74** (2001) 313.
31. A. Taylor and W. Zwicker, A characterization of weighted voting, *Proceedings of the American Mathematical Society* **115** (1992) 1089–1094.
32. F. Turnovec, Weights and votes in the European Union: Extension and institutional reform, *Prague Economic Papers* **5** (1996) 161–174.
33. M. Widgrén, Voting power in the EC and the consequences on two different enlargements, *European Economic Review* **38** (1994) 1153–1170.
34. S. Wolf, Proof without words, this MAGAZINE **62** (1989) 190.

Proof Without Words: Ordering Arithmetic, Geometric, and Harmonic Means



The arithmetic mean $A(a, b) = (a + b)/2$, the geometric mean $G(a, b) = \sqrt{ab}$, and the harmonic mean $H(a, b) = 2ab/(a + b)$ can be seen from the picture to satisfy $H(a, b) \leq G(a, b) \leq A(a, b)$, with equality if and only if $a = b$.

—C. L. Frenzen
Naval Postgraduate School
Monterey, CA 93943

Counting on Chebyshev Polynomials

ARTHUR T. BENJAMIN
Harvey Mudd College
Claremont, CA 91711
benjamin@hmc.edu

DANIEL WALTON
University of California, Los Angeles
Los Angeles, CA 90095-1555
waltond@ucla.edu

It's hard to avoid Chebyshev polynomials. They appear in just about every branch of mathematics, including geometry, combinatorics, number theory, differential equations, approximation theory, numerical analysis, and statistics. (Rivlin [**6**] gives numerous examples.) Their significance can be immediately appreciated by the fact that the function $\cos n\theta$ is a Chebyshev polynomial function of $\cos\theta$. Specifically, for $n \geq 0$,

$$\cos(n\theta) = T_n(\cos(\theta)), \tag{1}$$

where T_n is the *Chebyshev polynomial of the first kind,* defined by $T_0(x) = 1$, $T_1(x) = x$, and for $n \geq 2$,

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x). \tag{2}$$

For example, $T_2(x) = 2x^2 - 1$, $T_3(x) = 4x^3 - 3x$, $T_4(x) = 8x^4 - 8x^2 + 1$. This generates the familiar trigonometric identity $\cos(2\theta) = 2\cos^2\theta - 1$, and the less familiar $\cos(3\theta) = 4\cos^3\theta - 3\cos\theta$ and $\cos(4\theta) = 8\cos^4\theta - 8\cos^2\theta + 1$.

If we change the initial conditions to be $U_0(x) = 1$ and $U_1(x) = 2x$, but keep the same recurrence

$$U_n(x) = 2xU_{n-1}(x) - U_{n-2}(x),$$

we get the *Chebyshev polynomials of the second kind.* For instance, $U_2(x) = 4x^2 - 1$, $U_3(x) = 8x^3 - 4x$, $U_4(x) = 16x^4 - 12x^2 + 1$.

The Chebyshev polynomials generate many fundamental sequences, including the constant sequence, the sequence of integers, and the Fibonacci numbers. It's easy to show that for all $n \geq 0$, $T_n(1) = 1$ and $U_n(1) = n + 1$, $T_n(-1) = (-1)^n$, $U_n(-1) = (-1)^n(n + 1)$. When we substitute complex numbers, such as $x = i/2$, the Fibonacci and Lucas numbers appear. Specifically,

$$i^{-n}U_n(i/2) = f_n \tag{3}$$

and

$$2i^{-n}T_n(i/2) = L_n, \tag{4}$$

where $f_n = f_{n-1} + f_{n-2}$, and $L_n = L_{n-1} + L_{n-2}$, with initial conditions $f_0 = f_1 = 1$, and $L_0 = 2$, $L_1 = 1$. (We note that the "classical" Fibonacci numbers are defined by $F_0 = 0$ and $F_1 = 1$, but $f_n = F_{n+1}$ is more natural for combinatorial purposes.) In fact, any sequence of numbers that satisfies a second order recurrence with constant coefficients can be expressed in terms of Chebyshev polynomials [**1**].

Here we list a few more intriguing identities satisfied by the Chebyshev polynomials. For $m, n \geq 0$,

$$U_n(x) = \sum_{j=0}^{n} x^j T_{n-j}(x) \tag{5}$$

$$T_m(T_n(x)) = T_{mn}(x) \tag{6}$$

Finally, if we define $\hat{U}_n(x) = U_{n-1}(x)$, then

$$\gcd(\hat{U}_m(x), \hat{U}_n(x)) = \hat{U}_{\gcd(m,n)}(x) \tag{7}$$

All of the identities above can be proved by induction and various algebraic methods. The point of this article is to show that these identities, and many others, can also be given elegant combinatorial proofs, once we understand what the Chebyshev polynomials are counting.

Combinatorial models for $U_n(x)$

So what do Chebyshev polynomials count? As motivation, consider the combinatorial model for the Fibonacci numbers. It's easy to show [**3, 4**], that the Fibonacci number f_n counts the ways to tile a $1 \times n$ strip using 1×1 squares and 1×2 dominoes of length two. For example, $f_4 = 5$ counts the five tilings of length four below.

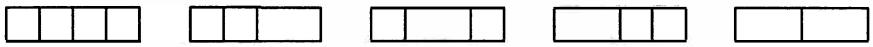


Figure 1 The fourth Fibonacci number $f_4 = 5$ is the number of square-domino tilings of length four

As it turns out, Chebyshev polynomials count the same objects as Fibonacci numbers, but we assign a *weight* to each tile. Specifically, we assign each square a weight of $2x$ and each domino a weight of -1, and define the *weight of a tiling* to be the product of the weights of its tiles. We provide the tilings of lengths two, three, and four, along with their respective weights, in FIGURE 2, and we see that their weights sum to Chebyshev polynomials, $U_2(x)$, $U_3(x)$, and $U_4(x)$.

This suggests the following theorem, originally due to Louis Shapiro [**7**].

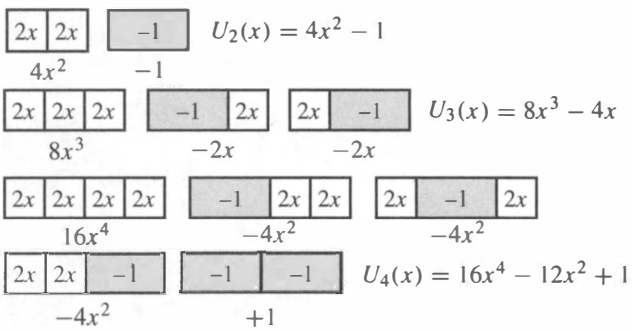


Figure 2 $U_2(x) = 4x^2 - 1$, $U_3(x) = 8x^3 - 4x$, $U_4(x) = 16x^4 - 12x^2 + 1$ is the total weight of all tilings of length two, three, and four, respectively

THEOREM 1. *For $n \geq 0$, $U_n(x)$ is the sum of the weights of all square-domino tilings of length n.*

Proof. Let w_n denote the total weight of all tilings of length n. It's easy to verify that $w_0 = 1 = U_0(x)$, $w_1 = 2x = U_1(x)$, and $w_2 = 4x^2 - 1 = U_2(x)$. Every tiling of length $n \geq 2$ comes from a tiling of length $n-1$ followed by a square (of weight $2x$) or comes from a tiling of length $n-2$ followed by a domino (of weight -1). Hence, $w_n = 2xw_{n-1} - 1w_{n-2}$. Then, by induction and the recurrence for U_n, $w_n = 2xU_{n-1}(x) - U_{n-2}(x) = U_n(x)$, as desired. ∎

Notice that a tiling of length n with k dominoes has exactly $n-2k$ squares and therefore has weight $(-1)^k(2x)^{n-2k}$. We leave it to the reader to show that the number of such tilings is $\binom{n-k}{k}$, which gives us the following closed form for $U_n(x)$.

IDENTITY 1. *For $n \geq 0$,*

$$U_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n-k}{k} (-1)^k (2x)^{n-2k}.$$

Moreover, if we let x take on the imaginary value $x = i/2$, we see that for $0 \leq k \leq n/2$, every length n tiling with k dominoes has weight $(-1)^k i^{n-2k} = i^n$, independent of k. Since there are f_n tilings of length n, we have $U_n(i/2) = i^n f_n$, and therefore we have the following Fibonacci identity

IDENTITY 2. *For $n \geq 0$, $i^{-n}U_n(i/2) = f_n$.*

The following "addition formula" is also easy to see, once we define the concept of breakability. We say that a tiling is *unbreakable* at cell m if a domino covers cells m and $m+1$; otherwise we say the tiling is *breakable* at cell m.

IDENTITY 3. *For all $m, n \geq 1$,*

$$U_{m+n}(x) = U_m(x)U_n(x) - U_{m-1}(x)U_{n-1}(x).$$

Proof. The total weight of length $m+n$ tilings that are breakable at cell m is $U_m(x)U_n(x)$ (by the distributive law). All tilings that are unbreakable at cell m consist of a tiling of length $m-1$ followed by a domino (with weight -1) followed by a tiling of weight $n-1$, and thus have total weight $-U_{m-1}(x)U_{n-1}(x)$. ∎

There is another way to interpret $U_n(x)$ combinatorially, which is a little more "colorful." Consider the set of *colored* tilings, where dominoes have just one color (light gray), but squares come in two colors (white or black). (Incidentally, the number of such tilings is the nth Pell number p_n, defined recursively by $p_0 = 1$, $p_1 = 2$ and for $n \geq 2$, $p_n = 2p_{n-1} + p_{n-2}$.) As in the previous model, we assign all dominoes a weight of -1, but since $2x = x + x$, we can assign each white square a weight of x and each black square a weight of x. As before, the weight of a tiling is the product of the weights of its tiles. In FIGURE 3, we list the five colored tilings of length 2 and the twelve colored tilings of length 3, along with their total weights.

Reasoning as before, we have the following theorem.

THEOREM 2. *For $n \geq 0$, $U_n(x)$ is the sum of the weights of all colored square-domino tilings of length n.*

Having two colors of squares to play with will allow us to prove many interesting facts about Chebyshev polynomials (especially of the first kind). Here is a simple identity that is easy to prove by induction, but the combinatorial technique we introduce will be useful to us later on.

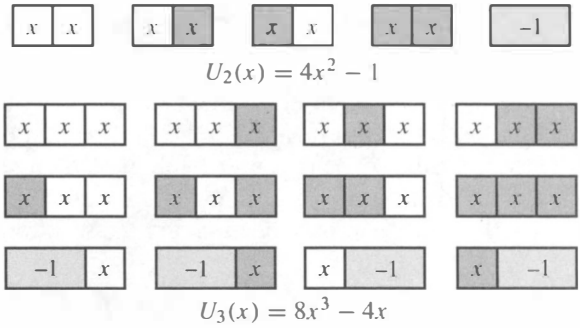


Figure 3 $U_n(x)$ using colored tilings, for $n = 2$ and $n = 3$

IDENTITY 4. *For $n \geq 0$, $U_n(1) = n + 1$.*

Proof. If we assign all squares (black or white) a weight of $x = 1$ and each domino, as usual, a weight of -1, then the weight of a colored tiling will be $(-1)^k$, where k is the number of dominoes. Consequently $U_n(1)$ counts the number of length n tilings with an even number of dominoes minus the number of length n tilings with an odd number of dominoes. Given any colored tiling X, we will try to pair it up with another tiling X' where the number of dominoes in X and X' have opposite parity. (Or put more poetically, we try to *find a mate of opposite weight.*)

Given a tiling X we look for the smallest number k where either (i) cells k and $k + 1$ are covered by a domino or (ii) cell k is covered by a white square and cell $k + 1$ is covered by a black square. If case (i) occurs, then we define X' to be the same tiling as X but with the first domino replaced by wb, where w denotes a white square and b denotes a black square. If case (ii) occurs, we replace the first wb with a domino. Thus X and X' have opposite weight. Notice that $(X')' = X$, whenever X' is defined.

When is X' undefined? Whenever X has no dominoes and no occurrence of wb, that is, whenever $X = b^j w^{n-j}$ (j black squares followed by $n - j$ white squares) for some $0 \leq j \leq n$. Thus, there are $n + 1$ exceptional tilings, all of which have positive weight (since they have no dominoes), and therefore $U_n(1) = n + 1$. ■

Combinatorial models for $T_n(x)$

Chebyshev polynomials of the first kind have at least four useful combinatorial interpretations using tilings. Since they satisfy the same recurrence as Chebyshev polynomials of the second kind, but with different initial conditions, then they only differ in how they weight the initial tile [**4**, Chapter 3]. As before, we define the weight of a length n tiling of squares and dominoes to be the product of the weights of its tiles, where each domino has weight -1 and each square has weight $2x$, *but* if the tiling begins with a square then that initial square has weight x. For example, the tilings for $T_n(x)$, with $n = 2, 3, 4$ are given in FIGURE 4.

The following "uncolored" interpretation has essentially the same proof as Theorem 1.

THEOREM 3. *$T_n(x)$ is the total weight of all uncolored tilings of length n, where an initial square has weight x, all other squares have weight $2x$, and all dominoes have weight -1.*

Reasoning as in Identity 3, we get an addition formula for $T_n(x)$.

IDENTITY 5. *For $m, n \geq 1$, $T_{m+n}(x) = T_m(x)U_n(x) - T_{m-1}(x)U_{n-1}(x)$.*

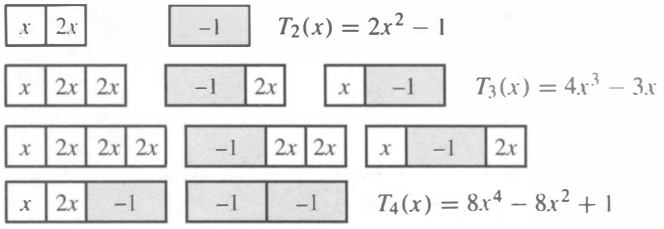


Figure 4 The weight of uncolored tilings as counted by $T_n(x)$

We note that Identities 3 and 5 remain true when $m = 0$ or $n = 0$ provided we extend the recurrence so that $T_{-1}(x) = x$ and $U_{-1}(x) = 0$.

Chebyshev polynomials of the first kind can also be thought of as counting circular tilings of *bracelets*. Specifically, if we take the previous model and multiply the weight of the initial tile by two, then all squares would receive a weight of $2x$, but now an initial domino has weight -2. We can think of this as counting two types of initial dominoes (each with weight -1). A domino of the first type will cover cells 1 and 2, as usual, but a domino of the second type will cover cells n and 1, as in FIGURE 5, giving us the following theorem.

THEOREM 4. *$2T_n(x)$ is the total weight of all uncolored circular tilings of length n, where each square has weight $2x$ and each domino has weight -1.*

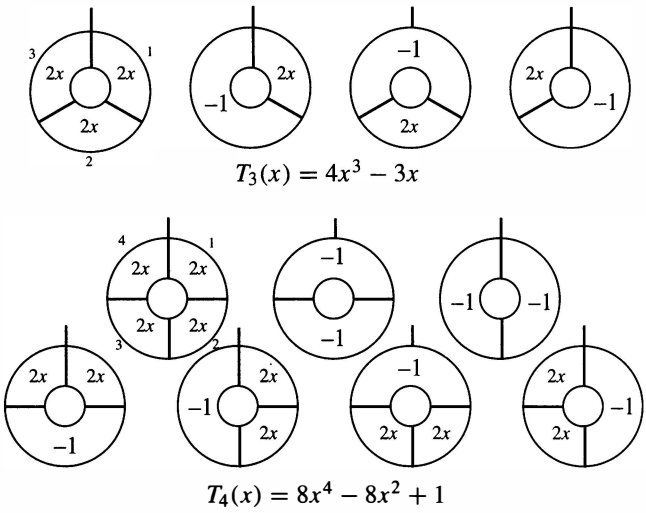


Figure 5 $2T_n(x)$ counts weighted bracelets

By considering whether or not a bracelet has a domino covering cells n and 1, we have

IDENTITY 6. *For $n \geq 1$, $2T_n(x) = U_n(x) - U_{n-2}(x)$.*

How many bracelets of length n have exactly k dominoes? By considering whether or not it has a domino covering cells n and 1, there are

$$\binom{n-k-1}{k-1} + \binom{n-k}{k} = \frac{n}{n-k}\binom{n-k}{k}$$

such bracelets. Thus by the reasoning that precedes Identity 1, we have a similar closed form for $T_n(x)$.

IDENTITY 7. *For $n > 0$,*

$$T_n(x) = \frac{1}{2}\sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n}{n-k}\binom{n-k}{k}(-1)^k(2x)^{n-2k}.$$

Since the number of length n bracelets is well known to be the Lucas number L_n [**4**, Chapter 2]. So just like with Identity 2, we have

IDENTITY 8. *For $n \geq 0$, $2i^{-n}T_n(i/2) = L_n$.*

Other useful identities are obtained by allowing our squares to come in two colors. As an immediate consequence of Theorem 3, we have

THEOREM 5. *$T_n(x)$ is the total weight of all colored tilings of length n, where an initial square has weight $x/2$, all other squares have weight x, and all dominoes have weight -1.*

Alternatively, we can allow all squares, to have weight x, but now we *restrict* the color of an initial square.

THEOREM 6. *$T_n(x)$ is the total weight of all colored tilings of length n, where all squares have weight x, all dominoes have weight -1, but the tiling may not begin with a black square. (Alternatively, $T_n(x)$ is the total weight of tilings that do not begin with a white square.)*

In FIGURE 6, we list the restricted colored tilings counted by $T_n(x)$, for $n = 2$ and $n = 3$.

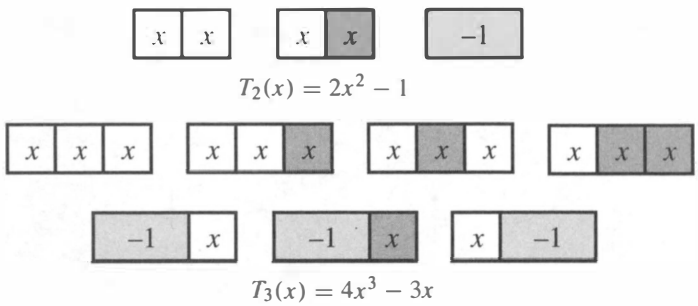


Figure 6 Examples of $T_n(x)$ using restricted colored tilings

This tiling interpretation of $T_n(x)$ was exploited combinatorially by Walton [**8**] to prove many Chebyshev polynomial identities. For example, by applying the same argument as done in Identity 4, we invite the reader to combinatorially prove

IDENTITY 9. *For all $n \geq 0$, $T_n(1) = 1$.*

Here is another quick proof of something we call the "string of lights" identity.

IDENTITY 10. *For $n \geq 1$, $T_n(x) = \sum_{j=1}^{n}(x^j T_{n-j}(x)) - U_{n-2}(x)$.*

Proof. A restricted tiling must either begin with a domino or a string of white squares. The total weight of tilings of length n that begin with exactly $j \geq 1$ white squares is $x^j T_{n-j}(x)$ since the tile that follows the first j white squares is restricted to

be a domino or a dark square. The total weight of those tilings that begin with a domino is $-U_{n-2}(x)$ since the initial domino has weight -1 and the remaining colored tiling is unrestricted. ■

We invite the reader to combinatorially prove

IDENTITY 11. *For* $n \geq 1$, $U_n(x) = T_n(x) + xU_{n-1}(x)$.

With more ingenuity, Walton [**8**] presents combinatorial proofs of trickier identities. For example, for $n \geq m \geq 0$,

$$T_n^2(x) + T_m^2(x) = 1 + T_{n+m}(x)T_{n-m}(x),$$

and for $m, n \geq 0$,

$$T_m(T_n(x)) = T_{mn}(x).$$

We will come back to this last identity in the next section.

At this point, we should expose the fact that some of the identities presented here are true for any sequence satisfying a second order recurrence with constant coefficients. If $u_{-1} = 0$, $u_0 = 1$, and u_n satisfies the recurrence $u_n = au_{n-1} + bu_{n-2}$, then u_n is the total weight of all tilings of length n where squares have weight a, dominoes have weight b, and the weight of a tiling is the product of its weights [**3**]. (Ironically, these are called Lucas sequences of the first kind, but they correspond to Chebyshev polynomials of the second kind.) Thus we immediately obtain generalizations of some of our earlier identities like

$$u_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n-k}{k} b^k a^{n-2k}$$

and

$$u_{m+n} = u_m u_n + b u_{m-1} u_{n-1}.$$

The constants a and b can be real or complex numbers, but they could also be polynomials. Moreover if a and b are relatively prime integers, and if we define $U_n = u_{n-1}$, then it can be shown by combinatorial argument [**3**] that

$$\gcd(U_m, U_n) = U_{\gcd(m,n)}.$$

The same line of reasoning will work when a and b are relatively prime polynomials like $2x$ and -1, which explains equation (7) in the introduction.

Combinatorial trigonometry

Finally, we come *full circle* and explain the trigonometric identity (1) at the beginning of the paper, namely

IDENTITY 12. *For* $n \geq 0$, $\cos(n\theta) = T_n(\cos\theta)$.

Readers may wish to prove this theorem by induction, using the definition of $T_{n+1}(x)$ and two applications of the angle addition formula for cosine. But the combinatorial proof, due to Benjamin, Ericksen, Jayawant, and Shattuck [**2**], is more fun and leads to other insights.

Proof. From Theorem 3, $T_n(\cos\theta)$ is the total weight of all tilings of length n where each domino has weight -1 and each square has weight $2\cos\theta$, except for an initial square, which has weight $\cos\theta$. But how do we combinatorialize $\cos\theta$? First, we use a formula from Euler

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}.$$

Then we assign to each square, except for an initial one, the weight $e^{i\theta} + e^{-i\theta}$, and half that weight to an initial square. Next, we introduce colored squares, but this time the white and black squares are given different weights: the weight of a white square is $e^{i\theta}$ and the weight of a black square is $e^{-i\theta}$ (except for an initial colored square, whose weight will be $\frac{1}{2}e^{i\theta}$ or $\frac{1}{2}e^{-i\theta}$). Thus, for example, the colored tiling in FIGURE 7 has weight $\frac{1}{2}e^{3i\theta}$.

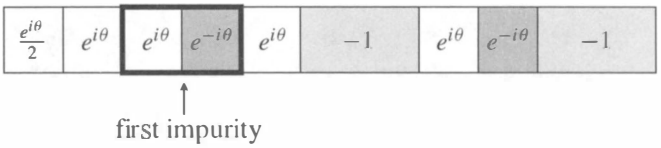


Figure 7 A colored tiling with weight $\frac{1}{2}e^{3i\theta}$

Reasoning as in Theorem 5, $T_n(\cos\theta)$ is the sum of the weights of all of these colored tilings. Our goal is to show that this sum is nearly zero. We say that a colored tiling is *impure* if it contains two consecutive square tiles of opposite color or at least one domino. In a colored tiling, we will call a domino or two consecutive squares of opposite color an *impurity*. For example, the tiling in FIGURE 7 has its first impurity at cells 3 and 4.

Next we claim that the sum of the weights of all impure tilings is zero. Let X be an impure tiling with its first impurity on cells k and $k+1$. We consider two cases.

First consider the case where $k \geq 2$. If cells k and $k+1$ are squares of opposite color, then we "find a mate of opposite weight" X' by replacing those two squares with a domino and leave all other tiles the same. If cells k and $k+1$ are covered by a domino, then we form X' by replacing the domino with two squares of opposite color where the color of the square on cell k is the same as the color of the square on cell $k-1$. Thus $(X')' = X$. Moreover, since two squares of opposite color have a weight $e^{i\theta}e^{-i\theta} = 1$ and a domino has weight -1, then it is clear that X and X' have opposite weight, so their weights sum to zero, as in FIGURE 8.

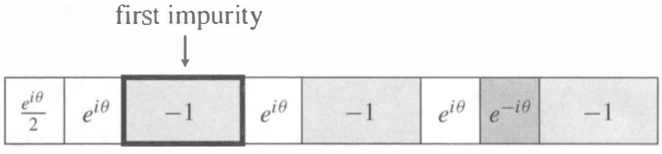


Figure 8 This tiling X' is the mate of the previous one and has weight $-\frac{1}{2}e^{3i\theta}$

On the other hand, if $k = 1$, then we "find a trio that sums to zero" by creating tilings X_1 and X_2, which are identical to X except for the first two cells. Among X, X_1, and X_2, one of them begins with wb (with weights that multiply to $1/2$) one of them begins with bw (with weights that multiply to $1/2$) and the other begins with a domino (with weight -1). Thus the weights of X, X_1, and X_2 sum to zero.

Consequently, every impure tiling belongs to a pair or trio with weights that sum to zero. Thus $T_n(\cos\theta)$ is just the sum of the weights of the pure tilings. But there are only two pure tilings, namely the tiling consisting of n white squares, with weight $e^{in\theta}/2$, and the tiling consisting of all black squares, with weight $e^{-in\theta}/2$. Thus

$$T_n(\cos\theta) = \frac{e^{in\theta} + e^{-in\theta}}{2} = \cos(n\theta),$$

as desired. ∎

By the same logic, we can generalize the last identity as follows.

IDENTITY 13. *For $n \geq 0$ and any real or complex number $z \neq 0$,*

$$T_n\left(\frac{z + 1/z}{2}\right) = \frac{z^n + 1/z^n}{2}.$$

We note that once the theorem is expressed in this form, it can then be proved easily by induction, but the combinatorial proof allows us to anticipate and appreciate this generalization.

By a slightly different argument and using $\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$, we can prove

$$\sin((n+1)\theta) = U_n(\cos\theta)\sin\theta$$

and its generalization

$$\left(\frac{z - 1/z}{2}\right) U_n\left(\frac{z - 1/z}{2}\right) = \frac{z^n - 1/z^n}{2}.$$

These and other trigonometric identities can also be given combinatorial proofs [**2**].

Identity 12 also leads to a quick proof of the composition theorem mentioned in the introduction.

IDENTITY 14. *For $m, n \geq 0$, $T_m(T_n(x)) = T_{mn}(x)$.*

Proof. When $x = \cos\theta$, we have

$$T_m(T_n(\cos\theta)) = T_m(\cos(n\theta)) = \cos(mn\theta) = T_{mn}(\cos\theta).$$

Since these polynomials agree at an infinite number of points (namely for all points $\cos\theta$), then they must be the same polynomial. ∎

Using a similar argument, there is a composition theorem for the Chebyshev polynomials of the second kind, namely

$$U_{m-1}(T_n(x))U_{n-1}(x) = U_{mn-1}(x).$$

For the combinatorial proof enthusiast, both of these composition theorems can also be proved using "tilings of tilings" [**5, 8**], but some might say that this is going a little "overboard."

Acknowledgment. The authors are grateful to Mark Shattuck, Doron Zeilberger, and the referees for their many helpful suggestions.

REFERENCES

1. D. Aharonov, A. Beardon, and K. Driver, Fibonacci, Chebyshev, and orthogonal polynomials, *Amer. Math. Monthly* **112** (2005) 612–630.
2. A. T. Benjamin, L. Ericksen, P. Jayawant, and M. Shattuck, Combinatorially composing Chebyshev polynomials, to appear in *Journal of Statistical Planning and Inference, Proceedings of the 6th International Conference on Lattice Path Combinatorics*, 2008.

3. A. T. Benjamin and J. J. Quinn, The Fibonacci numbers—exposed more discretely, this MAGAZINE **76** (2003) 182–192.
4. A. T. Benjamin and J. J. Quinn, *Proofs That Really Count: The Art of Combinatorial Proof*, Mathematical Association of America, Washington, DC, 2003.
5. A. T. Benjamin and D. Walton, Combinatorially composing Chebyshev polynomials, to appear in *Journal of Statistical Planning and Inference, Proceedings of the 6th International Conference on Lattice Path Combinatorics*, 2008.
6. T. J. Rivlin, *Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory*, John Wiley, New York, 1990.
7. L. W. Shapiro, A combinatorial proof of a Chebyshev polynomial identity, *Discrete Math.* **34** (1981) 203–206.
8. D. Walton, A tiling approach to Chebyshev polynomials, senior thesis, Harvey Mudd College, Claremont, CA, 2007.

Math Bite: Sums of Sines and Cosines

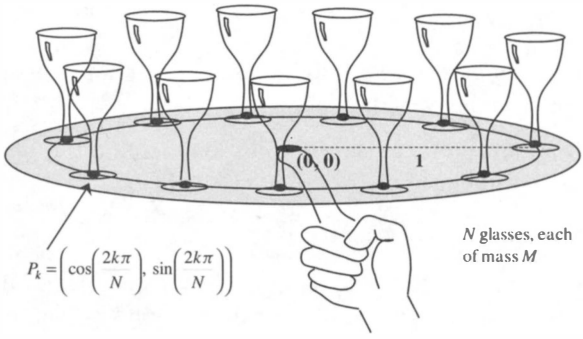
The sums of the title are, for N a natural number greater than 1,

$$\sum_{k=1}^{N} \cos\left(\frac{2\pi k}{N}\right) = 0 \quad \text{and} \quad \sum_{k=1}^{N} \sin\left(\frac{2\pi k}{N}\right) = 0.$$

The proofs of these two identities make a good exercise in complex algebra: Using DeMoivre's Theorem and Euler's formula, $e^{i(2\pi)} = (e^{i2\pi/N})^N = 1$, we derive the identities from the real and imaginary parts of the following equation:

$$\sum_{k=1}^{N} \left(\cos\left(\frac{2\pi k}{N}\right) + i \sin\left(\frac{2\pi k}{N}\right)\right) = \sum_{k=0}^{N-1} (e^{i2\pi/N})^k = \frac{(e^{i2\pi/N})^N - 1}{e^{i2\pi/N} - 1} = 0.$$

It seems that one must enter the realm of the complex numbers to prove this result, yet the validity of the identities becomes absolutely clear from the picture that follows.



$$\begin{aligned}(\bar{x}, \bar{y}) &= \left(\left(\textstyle\sum_{k=1}^{N} M\cos(2k\pi/N)\right)/NM, \left(\textstyle\sum_{k=1}^{N} M\sin(2k\pi/N)\right)/NM\right) \\ &= \left(\tfrac{1}{N}\textstyle\sum_{k=1}^{N}\cos(2k\pi/N), \tfrac{1}{N}\textstyle\sum_{k=1}^{N}\sin(2k\pi/N)\right) \\ &= (0, 0)\end{aligned}$$

—Judy A. Holdener
Kenyon College
Gambier, OH 43022

NOTES

Long Days on the Fibonacci Clock

EDWARD DUNNE
American Mathematical Society
Providence, RI 02904
egd@ams.org

The rule generating the Fibonacci sequence is extraordinarily simple, but its repeated application produces rich mathematics. If we leave the rule alone and change the starting values, we again find sequences with interesting properties—such as the Lucas numbers. If we introduce modular arithmetic, there are new questions to answer. In what follows, I study Fibonacci sequences in $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, the integers mod p, where p is a prime. I like to call these the "$\mathbb{F}_p$ibonacci numbers." Regardless of the starting pair, the sequence will repeat [**10**]. The question I want to answer is "What is the maximum period for any Fibonacci sequence in $\mathbb{Z}/p\mathbb{Z}$?"

In what follows, I present a particular point of view about the Fibonacci sequence in a way that gives some insight into both the standard sequence and its variations. Specifically, the Fibonacci sequence is interpreted in terms of a matrix acting on a finite set, an idea that is related to group actions and to (discrete) dynamical systems. The underlying set is the two-dimensional vector space $\mathbb{F}_p{}^2$; the matrix M from (2) provides the rule for the process. Iterations of the system correspond to powers of the matrix. Periods in the system are related, then, to powers of M that are equivalent modulo p to the identity matrix. The point of view works for all cases, even for the generalized $\mathbb{F}_p$ibonacci numbers, where weights are allowed in the recursion formula. For a thorough look at dynamical systems and number theory, Silverman's book [**8**] is an excellent source.

Most of what is contained here is not new. Searching *Mathematical Reviews* turns up dozens of articles about periods of Fibonacci numbers in $\mathbb{Z}/m\mathbb{Z}$, including many where m doesn't even have to be a prime or a power of a prime. The most-referenced article is Wall's article [**10**] in the MONTHLY in 1960. Wall established many fundamental results, and posed some tantalizing questions. In particular, he showed that the period divides $(p-1)$ when 5 is a quadratic residue mod p and divides $(2p+2)$ when it is not, but he did not find the maximal periods. Wall's investigation was motivated by a search for methods of generating pseudorandom numbers. Later, Brent [**1**], also motivated by pseudorandom numbers, considered the special properties of Fibonacci sequences modulo a power of 2. The story, however, begins even before the days of *Mathematical Reviews*. In the 1930s, Ward [**11**] considered periods, both minimal and maximal, and other characteristics of sequences arising from rather general recurrence relations, not just the Fibonacci relation. Kalman and Mena's article in an earlier issue of this MAGAZINE [**5**] examines many of the famous properties of the Fibonacci numbers as specific instances of properties of general second-order recurrences. Ward's results built on even earlier work by Carmichael [**2**] and others. For the early history of the subject, the curious reader should consult Dickson's history [**3**], particularly Volume I, Chapter XVII, where elements of the problem are traced back to Gauss and Lagrange. Earlier in the MAGAZINE, Vella and Vella [**9**] looked at possible periods in

the generalized Fibonacci sequence modulo a prime. Their approach emphasized recursive formulas and led to similar results to those here, but are somewhat less precise when applied to the standard Fibonacci numbers.

This investigation stems from a homework assignment from my daughter's fourth-grade mathematics class. The students were taught the Fibonacci recursion relation and how to reduce mod 100. The assignment was to find two starting numbers that gave the longest sequence before it repeated mod 100. Being the child of a mathematician, my daughter tried to *solve* the problem. It turns out the teacher imagined that the students would try some numbers and make some guesses about what would work best. This article shows what to do when the reduction is modulo a prime. If you would like to complete the fourth-grade assignment, you may apply the prime case and a little extra work to find the longest possible period for reduction modulo a composite.

Some examples Let $p = 19$. Since order matters and repeats are allowed, there are $19^2 = 361$ possible choices for the starting pair a_0 and a_1. The standard sequence, which starts with $a_0 = 1$ and $a_1 = 1$, becomes

$$1, 1, 2, 3, 5, 8, 13, 2, 15, 17, 13, 11, 5, 16, 2, 18, 1, 0, 1, 1, \ldots \pmod{19},$$

which has period 18. Using *Maple* to try all possible starting pairs shows that three hundred forty-two of them have period 18, eighteen of them have period 9, and one has period 1. In this case the maximum period is 18.

For $p = 23$, there are $23^2 = 529$ possible starting pairs. Direct computation shows that, other than the trivial sequence with $a_0 = a_1 = 0$, all the sequences have period equal to 48, making 48 the maximal period.

For $p = 29$, there are $29^2 = 841$ possible starting pairs. The standard sequence has period 14, as do eight hundred eleven other sequences. Twenty-eight sequences have period 7, and the trivial sequence has period 1.

Comment Since the sequences are periodic, it is a bit unnatural to say that, in the first example, the sequence starting with $a_0 = 1$ and $a_1 = 1$ is different from the sequence starting with $a_0 = 5$ and $a_1 = 8$, since they eventually come around to match up with each other. However, for these examples this is a convenient way to count.

At first glance, it seems that the periods are all over the map: Sometimes the period is $p - 1$, sometimes it is much less. Sometimes it is even bigger than p. However, by analyzing the sequences, in particular by examining the matrix that generates them all, certain features emerge that allow us to divide the problem into cases where the pattern becomes clear.

The problem

The sequences in question are

$$a_0, a_1, \ldots, a_{k-1}, a_k, a_{k+1}, \ldots,$$

$$a_{k+1} = a_k + a_{k-1} \tag{1}$$

$$\text{with } a_N \equiv a_0 \pmod{p} \text{ and } a_{N+1} \equiv a_1 \pmod{p}$$

and N is some number we don't know in advance. We will not restrict to the standard starting values of $a_0 = 0$ and $a_1 = 1$, which makes the question more interesting. We will assume that $p \neq 2$, since we often have to divide by 2. Most of the time, we will also assume that $p \neq 5$ to avoid a similar complication, as you will see. A few basic

facts from number theory are used, and can be found in the classic texts by Hardy and Wright [4] or Niven and Zuckerman [7].

A convenient way of generating the sequence is to call on linear algebra. The standard trick is to write the recursion relation as:

$$\begin{pmatrix} a_k \\ a_{k+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_{k-1} \\ a_k \end{pmatrix}. \tag{2}$$

For convenience, let $M = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Note that this is the companion matrix of the polynomial $x^2 - x - 1$, which is important for Fibonacci numbers. We can now write:

$$\begin{pmatrix} a_k \\ a_{k+1} \end{pmatrix} = M^k \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}. \tag{3}$$

In this formulation, the period of the Fibonacci sequence starting with a_0 and a_1 is the smallest positive integer k such that

$$M^k \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \equiv \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \pmod{p}. \tag{4}$$

Our problem, then, is to find a_0 and a_1 so that k is as large as possible. From (4), we see that k will always be less than any n such that

$$M^n \equiv I \pmod{p}, \tag{5}$$

where I is the 2×2 identity matrix. Moroever, the period of any Fibonacci sequence will be a divisor of n, meaning our k must divide n. As a result, the smallest such n, denoted $n(p)$, is an upper bound on the longest period.

Standard trick, part II: diagonalize the matrix:

$$M = A^{-1} D A. \tag{6}$$

The eigenvalues of M are $\mu = (1 + \sqrt{5})/2$ and $\overline{\mu} = (1 - \sqrt{5})/2$. Thus $D = \operatorname{diag}(\mu, \overline{\mu})$ and

$$A = \begin{pmatrix} 1 & 1 \\ \mu & \overline{\mu} \end{pmatrix}.$$

Then, $M^n \equiv I$ exactly when $\mu^n \equiv 1 \pmod{p}$ and $\overline{\mu}^n \equiv 1 \pmod{p}$. It's enough to figure out the minimum n for one of the two eigenvalues. I pick μ.

The solution

There are three cases to examine. We leave the case $p = 5$ to the end, as it is very different from the others. In each case, there are two tasks: compute $n(p)$ and determine whether any sequences of this maximal period occur.

Case 1: Suppose 5 is quadratic residue of p In this case, 5 has a square root in $\mathbb{F}_p$. By quadratic reciprocity, 5 is a quadratic residue of p when $p = 5N \pm 1$. Since the interesting values of p are odd, we actually have $p = 10N \pm 1$. Moreover, in this case, both μ and $\overline{\mu}$ are also elements of $\mathbb{F}_p$.

A *primitive element* of $\mathbb{F}_p^*$ is an element that generates $\mathbb{F}_p^*$ as a multiplicative group. If x is a primitive element of $\mathbb{F}_p^*$, then group theory tells us that $x^{p-1} \equiv 1$ and $p - 1$

is the least such exponent. Here $p-1$ is the the number of elements of $\mathbb{F}_p^*$. If x is not primitive, then the least such exponent is the order of x in the multiplicative group $\mathbb{F}_p^*$, which is necessarily a divisor of $p-1$.

PROPOSITION 1. *If 5 is a quadratic residue of p, then the smallest $n = n(p)$ satisfying (5) is the order of μ in $\mathbb{F}_p^*$. Moreover, there is at least one sequence with period $n(p)$.*

Proof. We have already seen that $n(p)$ equals the order of μ. For the second statement, we note that any nonmaximal period k corresponds to a nontrivial solution to (4), which means that $M^k - I$ has a nontrivial nullspace as a linear transformation on ${\mathbb{F}_p}^2$, the two-dimensional vector space over $\mathbb{F}_p$. The only way this nullspace can be nontrivial is for k to be a divisor of $n(p)$. We can now see that there won't be any sequences of maximal possible length if and only if the nullspaces of $M^k - I$, running over all *proper* divisors k of $n(p)$, exhaust ${\mathbb{F}_p}^2$. Now, since the nullspace is a vector space over $\mathbb{F}_p$, it will have either 1 or p elements. (We are already assuming that it's not the whole space.) However, $n(p)$ is either $p-1$ or a divisor of $p-1$, and there are fewer than p proper divisors of $p-1$. So by multiplying and counting, we see that the union of these nullspaces has fewer than p^2 elements, and cannot be all of ${\mathbb{F}_p}^2$. Hence, there must be at least one Fibonacci sequence with maximal period, $n(p)$. ■

Case 2: Suppose 5 is not a quadratic residue of p In this case, μ and $\overline{\mu}$ are not elements of $\mathbb{F}_p$. It is necessary, then, to work over the field $\mathbb{F}_p(\sqrt{5}) \cong \mathbb{F}_{p^2}$. The problem becomes finding the order of μ in $\mathbb{F}_p(\sqrt{5})^*$.

Write out

$$\mu^{p+1} = \left(\frac{1+\sqrt{5}}{2}\right)^{p+1} = \frac{1}{2^{p+1}}(1+\sqrt{5})^{p+1}$$

and reduce mod p. Reducing the denominator as $2^{p+1} = 2^p \cdot 2 \equiv 2 \cdot 2 \equiv 4$, since $x^p \equiv x \pmod{p}$ for all x, gives $1/2^{p+1} \equiv 1/4$. The second factor reduces as:

$$\begin{aligned}(1+\sqrt{5})^{p+1} &= \left((\sqrt{5})^{p+1} + (p+1)(\sqrt{5})^p + \frac{p(p+1)}{2}(\sqrt{5})^{p-1} + \cdots \right.\\ &\qquad \left. + \frac{p(p+1)}{2}(\sqrt{5})^2 + (p+1)\sqrt{5} + 1\right)\\ &\equiv \left((\sqrt{5})^{p+1} + (\sqrt{5})^p + 0 + \cdots + 0 + \sqrt{5} + 1\right) \pmod{p}.\end{aligned}$$

Now, compute $(\sqrt{5})^{p-1} = 5^{(p-1)/2}$: In general, if a is *any* quadratic nonresidue of p, then

$$(p-1)! \equiv a^{(p-1)/2} \pmod{p}, \tag{7}$$

which can be seen by multiplying together the (necessarily unequal) pairs of elements x and x' such that $x \cdot x' \equiv a \pmod{p}$. On the other hand, by doing a similar thing for $a = -1$ (and handling separately the cases when -1 is and is not a quadratic residue) we get Wilson's Theorem,

$$(p-1)! \equiv -1 \pmod{p}. \tag{8}$$

By combining (7) and (8), we see

$$(\sqrt{5})^{p-1} = 5^{(p-1)/2} \equiv -1 \pmod{p}.$$

Comment This formula was already known to Euler, but we will want to recall the method when considering generalized sequences in the last section. Now, substituting this into the expansion of $(1+\sqrt{5})^{p+1}$, we obtain

$$\begin{aligned}\mu^{p+1} &\equiv \frac{1}{2^{p+1}}\left((\sqrt{5})^{p+1}+(\sqrt{5})^{p}+\sqrt{5}+1\right)\\ &\equiv (1/4)\left((5\cdot 5^{(p-1)/2}+1)+\sqrt{5}(5^{(p-1)/2}+1)\right)\\ &\equiv (1/4)\left((5(-1)+1)+\sqrt{5}(-1+1)\right)\equiv (1/4)(-4)\equiv -1\end{aligned}$$

Then, $\mu^{2(p+1)}=1$ in $\mathbb{F}_p(\sqrt{5})$. Moreover, since $\mu^{p+1}=-1$ in $\mathbb{F}_p(\sqrt{5})$, we see that $2(p+1)$ is the least such exponent.

PROPOSITION 2. *If 5 is not a quadratic residue of p, then the smallest $n=n(p)$ satisfying* (5) *is $n(p)=2(p+1)$. There is at least one sequence with period $n(p)$.*

Proof. We have already computed the value of $n(p)$. The proof of the second statement is essentially the same sort of counting argument as in the first case, which shows that there aren't enough "short periods" to exhaust the $\mathbb{F}_p{}^2$ of possible sequences. Therefore, the maximum is attained in this case, too. ∎

Case 3: $p=5$ Since $5\equiv 0 \pmod 5$, we have $\mu\equiv(1+0)/2\equiv 1/2\equiv 3 \pmod 5$ and $\overline{\mu}\equiv(1-0)/2\equiv 1/2\equiv 3=\mu$. Thus, the matrix A in our earlier analysis is

$$A=\begin{pmatrix}1 & 1\\ \mu & \overline{\mu}\end{pmatrix}\equiv\begin{pmatrix}1 & 1\\ 3 & 3\end{pmatrix},$$

which is singular, i.e., there is no A^{-1}. The point is that we cannot diagonalize the Fibonacci matrix M in this case. However, $\mathbb{Z}/(5)$ is small, and it is not too hard to run through all the possibilities. Starting with $a_0=1$ and $a_1=1$ leads to the sequence

$$1, 1, 2, 3, 0, 3, 3, 1, 4, 0, 4, 4, 3, 2, 0, 2, 2, 4, 1, 0, \text{ repeat}$$

which has period 20. This is the period for most choices of initial values. However, for $a_0=1$ and $a_1=3$, the sequence is just

$$1, 3, 4, 2, \text{ repeat}$$

which has a period of just 4. The only other sequence is the trivial sequence $0, 0, 0, \ldots$.

Summary for Fibonacci sequences If 5 is a quadratic residue of p, then the maximal period of any Fibonacci sequence in $\mathbb{F}_p$ is the order of μ in $\mathbb{F}_p^*$. This is the maximal value of $p-1$ when μ is a primitive element. If μ is not primitive, then the order is some divisor of $p-1$, which needs to be determined by a direct calculation. By quadratic reciprocity, the primes p are of the form $p=10N\pm 1$.

If 5 is a quadratic nonresidue of p, then the maximal period of a Fibonacci sequence in $\mathbb{F}_p$ is $2(p+1)$. By quadratic reciprocity, these primes are of the form $p=10N+7$ and $p=10N+3$.

For $p=5$, the possible nontrivial periods are 4 and 20.

Generalized Fibonacci numbers

The Fibonacci recurrence relation can be generalized to allow for weights: $b_n=\alpha b_{n-1}+\beta b_{n-2}$, where α and β are integers. We can then ask for the maximal periods

of sequences of generalized Fibonacci numbers, modulo a prime p:

$$\begin{gathered} b_0, b_1, \ldots, b_{k-1}, b_k, b_{k+1}, \ldots, b_n = \alpha b_{n-1} + \beta b_{n-2} \\ b_N \equiv b_0 \pmod{p} \text{ and } b_{N+1} \equiv b_1 \pmod{p}. \end{gathered} \tag{9}$$

The weights α and β need not be positive integers. However, in order to avoid accidental multiplication by zero, we should make sure the α and β are both relatively prime to p.

The arguments used for the standard Fibonacci now carry over, but become harder. The matrix becomes $M = \begin{pmatrix} 0 & 1 \\ \beta & \alpha \end{pmatrix}$. The eigenvalues become

$$\mu = \frac{\alpha + \sqrt{\alpha^2 + 4\beta}}{2} \text{ and } \overline{\mu} = \frac{\alpha - \sqrt{\alpha^2 + 4\beta}}{2}.$$

Let $D = \alpha^2 + 4\beta$ be the discriminant of the polynomial $x^2 - \alpha x - \beta$, which plays the role of $x^2 - x - 1$ from the standard Fibonacci numbers. Assume, for the time being, that $D \neq 0$. If D is a quadratic residue of p, then the argument in Case 1 goes through *mutatis mutando.*

If D is not a quadratic residue of p, then we need to be more careful. We again need to work in an extension field of $\mathbb{F}_p$, this time the field is $\mathbb{F}_p(\sqrt{D})$. The essential problem is to determine the orders of μ and $\overline{\mu}$ in $\mathbb{F}_p(\sqrt{D})^*$. As before, the order of this multiplicative group is $p^2 - 1$, which factors as $(p-1)(p+1)$. Computing μ^{p+1} is a little more difficult now, since α and β are not explicit, meaning we don't have a tidy expression for μ. However, we can use the isomorphism $\mathbb{F}_p(\sqrt{D}) \cong \mathbb{F}_{p^2}$ and a standard fact about the map $x \mapsto x^p$ in a field of characteristic p. This map, denoted Fr, is called the *Frobenius map* or the Frobenius endomorphism, and it is very useful in number theory. In our setting, the Frobenius map has the following useful properties, as can be found in the book by Mullen and Mummert [**6**]:

$$\begin{aligned} \mathrm{Fr}(xy) &= \mathrm{Fr}(x)\mathrm{Fr}(y) \\ \mathrm{Fr}(x+y) &= \mathrm{Fr}(x) + \mathrm{Fr}(y) \\ \mathrm{Fr}(x) &= x \quad \text{if and only if } x \in \mathbb{F}_p \\ \mathrm{Fr}^2 &= \mathrm{Id}. \end{aligned}$$

The first two properties are just another way of saying that Fr is an endomorphism, that is, a homomorphism from $\mathbb{F}_p(\sqrt{D})$ to itself. The last property is special to the case of a quadratic extension, and can be deduced using reasoning similar to the computation presented in Case 2 above. The last two properties combine to imply $\mathrm{Fr}(\sqrt{D}) = -\sqrt{D}$.

We can now set about computing μ^{p+1}. Write out:

$$\mu^{p+1} = \left(\frac{\alpha + \sqrt{D}}{2}\right)^{p+1} = \frac{1}{2^{p+1}}(\alpha + \sqrt{D})^{p+1}$$

and reduce mod p. As before, the denominator reduces as $2^{p+1} = 2^p 2 \cong 2 \cdot 2 = 4$. The second factor reduces as:

$$\begin{aligned} (\alpha + \sqrt{D})^{p+1} &= (\alpha + \sqrt{D})^p(\alpha + \sqrt{D}) = \mathrm{Fr}(\alpha + \sqrt{D})(\alpha + \sqrt{D}) \\ &= (\alpha - \sqrt{D})(\alpha + \sqrt{D}) = \alpha^2 - D \end{aligned}$$

But $D = \alpha^2 + 4\beta$, so the second factor reduces to $\alpha^2 - (\alpha^2 + 4\beta) = -4\beta$ and

$$\mu^{p+1} \cong -\beta.$$

For the standard Fibonacci numbers, $\beta = 1$. So we knew that $\beta^2 = 1$ and could conclude that $\mu^{2(p+1)} \cong 1$. Now, however, we need to know the order d of $-\beta$ in $\mathbb{F}_p^*$. This is not an easy problem in general. All we really know is that d must divide $(p-1)$, the order of $\mathbb{F}_p^*$. Thus, the best we can conclude is only that the maximum period of the generalized Fibonacci sequence is $d(p+1)$, and we are left with separate computations for every case. Again, a counting argument verifies that the maximal period $n(p)$ does occur. For $\alpha = 3$, $\beta = 7$, and $p = 13$, the maximum *possible* period is $n(p) = p^2 - 1 = 168$. Using the starting values $b_0 = 0$ and $b_1 = 1$, a computation using *Maple* shows that this maximum period does indeed occur.

If the discriminant $D = \alpha^2 + 4\beta$ is zero, the situation is rather different. Observe that $\beta = -(\alpha/2)^2$ and, for notational convenience, let $\lambda = \alpha/2$. The recurrence relation now becomes:

$$b_n = 2\lambda b_{n-1} - \lambda^2 b_{n-2}$$

and the matrix becomes

$$M = \begin{pmatrix} 0 & 1 \\ -\lambda^2 & 2\lambda \end{pmatrix}.$$

Unfortunately, M is not diagonalizable. It has Jordan form

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \text{ so that } J^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}.$$

If k is a period for this generalized Fibonacci sequence, then we want to find k such that $J^k \equiv I$, which means

$$\begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{p}.$$

Since $\mathbb{Z}/p\mathbb{Z}$ is a field, the equality of the (2, 2) entries tells us that $\lambda^k \equiv 1 \pmod{p}$. This implies that k is a multiple of the order of λ in $(\mathbb{Z}/p\mathbb{Z})^*$, which is a divisor of $(p-1)$. Comparison of the (1, 2) entries tells us that $k\lambda^{k-1} \equiv 0$. Again using that $\mathbb{Z}/p\mathbb{Z}$ is a field, we deduce that one of the factors must be zero. But a power of λ is zero only if λ itself is zero, so we are left with $k \equiv 0 \pmod{p}$, implying that k is a multiple of p. So the longest possible period is kp, where k is the order of λ in $(\mathbb{Z}/p\mathbb{Z})^*$. When λ is primitive, the longest possible period is $p(p-1)$.

Acknowledgment. I am grateful to Peter Trapa who suggested the counting argument that is used in the proofs of the second statements in the propositions.

REFERENCES

1. R. P. Brent, On the periods of generalized Fibonacci recurrences, *Math. of Comp.* **63** (1994) 389–401.
2. R. D. Carmichael, On sequences of integers defined by recurrence relations, *Quart. J. Math.* **48** (1920) 343–372.
3. L. E. Dickson, *History of the Theory of Numbers*, Amererican Mathematical Society, Providence, RI, 1999. (First edition: Carnegie Institution of Washington, 1919–1923.)
4. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 4th ed., Oxford University Press, Oxford, 1960.
5. D. Kalman and R. Mena, The Fibonacci numbers—exposed, this MAGAZINE **76** (2003) 167–181.

6. G. L. Mullen and C. Mummert, *Finite Fields and Applications*, American Mathematical Society, Providence, RI, 2007.
7. I. Niven and H. S. Zuckerman, *An Introduction to the Theory of Numbers*, 3rd ed., John Wiley, New York 1972.
8. J. Silverman, *The Arithmetic of Dynamical Systems*, Graduate Texts in Mathematics, vol. 241, Springer, New York, 2007.
9. D. Vella and A. Vella, Cycles in the generalized Fibonacci sequence modulo a prime, this MAGAZINE **75** (2002) 294–299.
10. D. D. Wall, Fibonacci series modulo m, *Amer. Math. Monthly* **67** (1960) 525–532.
11. M. Ward, The arithmetical theory of linear recurring series, *Trans. Amer. Math. Soc.* **35** (1933) 600–628.

Fooling Newton's Method as Much as One Can

JORMA K. MERIKOSKI
Department of Mathematics and Statistics
University of Tampere
FI-33014 Tampere, Finland
jorma.merikoski@uta.fi

TIMO TOSSAVAINEN
Department of Teacher Education
University of Joensuu
FI-57101 Savonlinna, Finland
timo.tossavainen@joensuu.fi

We enjoyed reading how Horton [**1**] "fooled Newton's method" with an example where the sequence

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

converges but its limit does not satisfy $f(x) = 0$. Indeed, if

$$f(x) = \begin{cases} \pi - 2x \sin \frac{\pi}{x} & \text{for } x \neq 0, \\ \pi & \text{for } x = 0, \end{cases} \tag{1}$$

then the Newton sequence is

$$x_{n+1} = x_n - \frac{1}{2} \frac{\pi x_n - 2x_n^2 \sin \frac{\pi}{x_n}}{\pi \cos \frac{\pi}{x_n} - x_n \sin \frac{\pi}{x_n}},$$

and, starting from $x_1 = 1/2$, we have $x_2 = 1/4$, $x_3 = 1/8, \ldots, x_n = 1/2^n \to 0$, although $f(0) = \pi \neq 0$.

Can f be differentiable? Note that the function in (1) is not differentiable at $x = 0$. Since we thought that a differentiable function would fool the method even better, we wanted to know if such a function exists. Simply modifying Horton's function, we found an example that readers might find even more surprising:

$$f(x) = \begin{cases} \pi - x^2 \sin \frac{\pi}{x^2} & \text{for } x \neq 0, \\ \pi & \text{for } x = 0. \end{cases} \tag{2}$$

This function is differentiable and its Newton sequence is

$$x_{n+1} = x_n - \frac{1}{2}\frac{\pi x_n - x_n^3 \sin\frac{\pi}{x_n^2}}{\pi\cos\frac{\pi}{x_n^2} - x_n^2\sin\frac{\pi}{x_n^2}}.$$

Again, if $x_1 = 1/2$, then $x_n = 1/2^n \to 0$, but $f(0) = \pi \neq 0$.

Can f be continuously differentiable? Because f in (2) is not continuously differentiable at $x = 0$, our next question was: Can we fool Newton's method with a continuously differentiable function?

The answer is negative. More generally, f' cannot be bounded near the limit point x_0. (If f' were continuous, then it would be bounded there.) For, assume that $x_n \to x_0$ and f is continuous at x_0. Since

$$x_n - x_{n+1} = \frac{f(x_n)}{f'(x_n)}$$

and the left-hand side has limit zero while the numerator of the right-hand side has limit $f(x_0)$, it follows that

$$f(x_0) = 0 \quad \text{or} \quad |f'(x_n)| \to \infty.$$

(Note that the existence of $f'(x_0)$ is not needed. It is enough that $f'(x)$ exists when $x \neq x_0$ is near x_0.)

Acknowledgment. We thank one referee for suggestions that simplified the original manuscript considerably. We also thank the other referee for valuable suggestions.

REFERENCE

1. P. Horton, No fooling! Newton's method can be fooled, this MAGAZINE **80** (2007) 383–387.

Ramanujan's 6–8–10 Equation and Beyond

MARC CHAMBERLAND
Grinnell College
Grinnell, IA 50112
chamberl@math.grinnell.edu

Among Ramanujan's many beautiful formulas is the 6–8–10 equation

$$\begin{aligned}
&64[(a+b+c)^6 + (b+c+d)^6 - (c+d+a)^6 \\
&\quad - (d+a+b)^6 + (a-d)^6 - (b-c)^6] \\
&\times [(a+b+c)^{10} + (b+c+d)^{10} - (c+d+a)^{10} \\
&\quad - (d+a+b)^{10} + (a-d)^{10} - (b-c)^{10}] \\
&= 45[(a+b+c)^8 + (b+c+d)^8 - (c+d+a)^8 \\
&\quad - (d+a+b)^8 + (a-d)^8 - (b-c)^8]^2
\end{aligned}$$

when $ad = bc$. Berndt and Bhargava [**2**] cite this as "one of the most fascinating identities we have ever seen." Letting

$$f_m = (1+x+y)^m + (-x-y-xy)^m - (-y-xy-1)^m \\ - (xy+1+x)^m + (-1+xy)^m - (-x+y)^m \tag{1}$$

and $b = ax$, $c = ay$, $d = axy$, Ramanujan's equation may be compactly stated as

$$45 f_8^2 = 64 f_6 f_{10}. \tag{2}$$

Proofs of equation (2) may be found in references [**3**], [**5**], and [**6**]. How Ramanujan found this identity (as with many of his results) remains shrouded in mystery. He also discovered $f_2 = 0$ and $f_4 = 0$. Only many decades later was a similar equation found by Hirschhorn [**5**], specifically

$$21 f_5^2 = 25 f_3 f_7. \tag{3}$$

The goal of this note is to show how these identities and others may be found.

Berndt [**1**] showed that the polynomials f_m associated with (2) factor nicely:

$$\begin{aligned} f_6 &= 3x(x-1)(2x+1)(x+2)(x+1)y(y-1)(2y+1)(y+2)(y+1), \\ f_8 &= 8x(x-1)(2x+1)(x+2)(1+x)(x^2+x+1) \\ &\quad \cdot y(y-1)(2y+1)(y+2)(y+1)(y^2+y+1), \\ f_{10} &= 15x(x-1)(2x+1)(x+2)(1+x)(x^2+x+1)^2 \\ &\quad \cdot y(y-1)(2y+1)(y+2)(y+1)(y^2+y+1)^2. \end{aligned}$$

These factorizations are easily replicated by *Maple*, but one can prove these formally by showing that the polynomials on each side have the same zeros (with multiplicity) and scaling factor. Ramanujan's identity (2) now becomes straightforward to demonstrate. Similarly, the polynomials f_m in Hirschhorn's equation (3) may be factored to obtain

$$\begin{aligned} f_3 &= -3(x-1)(2x+1)(x+2)y(y+1), \\ f_5 &= -5(x-1)(2x+1)(x+2)(x^2+x+1)y(y+1)(y^2+y+1), \\ f_7 &= -7(x-1)(2x+1)(x+2)(x^2+x+1)^2 y(y+1)(y^2+y+1)^2, \end{aligned}$$

allowing an easy proof of equation (3).

Are there other relationships among the polynomials f_m? Scanning the factors, one finds

$$5 f_3 f_8 = 8 f_5 f_6 \tag{4}$$

and

$$15 f_6 f_7 = 7 f_3 f_{10}. \tag{5}$$

In an attempt to find other identities, one may factor f_m for other values of m. Unfortunately, the factorizations don't yield any obvious treasures. *Maple* produces, for

example,

$$\begin{aligned} f_9 = -3(x-1)(2x+1)(x+2)y(y+1)(3 &+ 27xy + 9x + 9y + 21x^2 + 19y^2 \\ &+ 63x^2y + 57xy^2 + 27x^3 + 23y^3 + 105x^2y^2 + 81x^3y + 115x^3y^2 + 69xy^3 \\ &+ 105x^2y^3 + 95x^3y^3 + 21x^4 + 19y^4 + 63x^4y + 57xy^4 + 105x^4y^2 + 105x^2y^4 \\ &+ 27x^5y + 27xy^5 + 9x^5 + 9y^5 + 3x^6 + 3y^6 + 105x^4y^3 + 57x^5y^2 + 115x^3y^4 \\ &+ 9x^6y + 63x^2y^5 + 9xy^6 + 105x^4y^4 + 69x^5y^3 + 81x^3y^5 + 19x^6y^2 + 21x^2y^6 \\ &+ 57x^5y^4 + 63x^4y^5 + 27x^5y^5 + 23x^6y^3 + 27x^3y^6 + 19x^6y^4 + 21x^4y^6 \\ &+ 9x^6y^5 + 9x^5y^6 + 3x^6y^6) \end{aligned}$$

and

$$\begin{aligned} f_{11} = -11(x-1)(2x+1)(x+2)(x^2+x+1)y(y+1)(y^2+y+1) \\ (1 + 3x + 3y + 9xy + 9x^2 + 7y^2 + 27x^2y + 21xy^2 + 39x^3y + 33x^2y^2 \\ + 27xy^3 + 13x^3 + 9y^3 + 9x^4 + 7y^4 + 3x^5 + 3y^5 + x^6 + y^6 + 27x^4y \\ + 31x^3y^2 + 21x^2y^3 + 21xy^4 + 33x^4y^2 - 3x^3y^3 + 33x^2y^4 + 9x^5y + 9xy^5 \\ + 21x^5y^2 + 21x^4y^3 + 31x^3y^4 + 3x^6y + 27x^2y^5 + 3xy^6 + 7x^6y^2 + 27x^5y^3 \\ + 33x^4y^4 + 39x^3y^5 + 9x^2y^6 + 9x^6y^3 + 21x^5y^4 + 27x^4y^5 + 13x^3y^6 \\ + 7x^6y^4 + 9x^5y^5 + 3x^6y^5 + 9x^4y^6 + 3x^5y^6 + x^6y^6). \end{aligned}$$

By considering smaller values of m, one finds that f_{-1} and f_{-2} also have tidy factorizations, namely

$$f_{-1} = \frac{(x-1)(2x+1)(x+2)(x^2+x+1)y(y+1)(y^2+y+1)}{(1+x+y)(x+y+xy)(y+xy+1)(xy+1+x)(-1+xy)(x-y)}$$

and

$$f_{-2} = \frac{x(x-1)(2x+1)(x+2)(x+1)(x^2+x+1)^2y(y-1)(2y+1)(y+2)(y+1)(y^2+y+1)^2}{(1+x+y)^2(x+y+xy)^2(y+xy+1)^2(xy+1+x)^2(xy-1)^2(x-y)^2}.$$

Not surprisingly, there are relationships between f_{-1}, f_{-2}, and the other nicely factored terms:

$$f_{-2}f_3^2 = -3f_{-1}^2f_6. \tag{6}$$

Just as with f_9 and f_{11}, f_m does not seem to factor extensively for integers $m \leq -3$.

One may think this is the end of the road; however, note that the indices for each term in (2) sum to 16, to 10 in (3), 11 in (4), and 13 in (5). Searching for a similar equation where each term involves two f_ms whose indices sum to 14, one could look for constants a, b, c, and d for which

$$af_3f_{11} + bf_5f_9 + cf_6f_8 + df_7^2 = 0.$$

Note that any other possible terms are vacuous since f_1, f_2, and f_4 are each identically zero. By evaluating this equation at four points (x, y) and using linear algebra, one arrives at the conjecture

$$245f_3f_{11} - 539f_5f_9 + 330f_7^2 = 0. \tag{7}$$

How can one prove this equation is valid? Entering the expression on the left into *Maple* and simplifying produces zero. Alternatively, one may also use Hirschhorn's generating function approach [**5**]. Defining

$$a_1 = 1 + x + y, \quad b_1 = -x - y - xy, \quad c_1 = -1 + xy,$$
$$a_2 = -y - xy - 1, \quad b_2 = xy + 1 + x, \quad c_2 = -x + y,$$
$$q = (x^2 + x + 1)(y^2 + y + 1), \quad p_1 = a_1 b_1 c_1, \quad p_2 = a_2 b_2 c_2,$$

one finds

$$f_3 = 2(p_1 - p_2), \quad f_5 = 5q(p_1 - p_2), \quad f_6 = 3(p_1^2 - p_2^2), \quad f_7 = 7q^2(p_1 - p_2),$$
$$f_8 = 8q(p_1^2 - p_2^2), \quad f_9 = 3(p_1 - p_2)(p_1^2 + p_1 p_2 + p_2^2 + 3q^3),$$
$$f_{10} = 15q^2(p_1^2 - p_2^2), \quad f_{11} = 11q(p_1 - p_2)(p_1^2 + p_1 p_2 + p_2^2 + q^3).$$

While one now clearly obtains not only (2) and (3) but also (7), this approach has its limitations. Though the polynomials f_m may be expressed in a more compact form using p_1, p_2, and q, these representations of f_m will also become unwieldy for modest values of m.

Yet another approach to establish (7) involves difference equations. Since f_m is a linear combination of six mth powers, f_m satisfies a sixth order difference equation whose characteristic equation has the six bases as its roots. With the factorizations noted earlier, the characteristic equation becomes

$$\begin{aligned}
0 &= (r - 1 - x - y)(r + x + y + xy)(r + y + xy + 1) \\
&\quad (r - xy - 1 - x)(r + 1 - xy)(r + x - y) \\
&= r^6 - 2(x^2 + x + 1)(y^2 + y + 1)r^4 \\
&\quad + x(x + 1)(y - 1)(2y + 1)(y - 2)r^3 + (x^2 + x + 1)^2(y^2 + y + 1)^2 r^2 \\
&\quad - x(x + 1)(x^2 + x + 1)(y - 1)(2y + 1)(y - 2)(y^2 + y + 1)r \\
&\quad - (xy + 1 + y)(1 + x + y)(xy + 1 + x)(xy + x + y)(x - y)(xy - 1) \\
&= r^6 - \frac{6}{5}\frac{f_5}{f_3}r^4 - \frac{f_6}{f_3}r^3 + \frac{9}{25}\frac{f_5^2}{f_3^2}r^2 + \frac{3}{5}\frac{f_5 f_6}{f_3^2}r + \frac{1}{5}\frac{f_5}{f_{-1}}
\end{aligned}$$

thus yielding

$$0 = f_m - \frac{6}{5}\frac{f_5}{f_3}f_{m-2} - \frac{f_6}{f_3}f_{m-3} + \frac{9}{25}\frac{f_5^2}{f_3^2}f_{m-4} + \frac{3}{5}\frac{f_5 f_6}{f_3^2}f_{m-5} + \frac{1}{5}\frac{f_5}{f_{-1}}f_{m-6} \qquad (8)$$

for all m. Specific choices of m in (8) give some known formulas: $m = 4$ produces (6), $m = 7$ gives Hirschhorn's equation (3), $m = 8$ gives (4), and $m = 10$ (with help from (3) and (4)) yields Ramanujan's equation (2). Indeed, (8) may be used recursively to generate many formulas. To obtain (7), take the $m = 11$ equation multiplied by f_3, the $m = 9$ equation multiplied by f_5, then subtract. This eliminates the f_{-1} terms and, combined with previously discovered identities, yields the desired result.

The linear algebra approach used to find equation (7) may be used to find many identities. Other equations found include

$$\begin{aligned}
308f_{10}^2 &= 525f_8f_{12} - 300f_6f_{14},\\
1763580f_{11}^2 &= 2735810f_9f_{13} - 1172490f_7f_{15} + 144837f_5f_{17} + 71995f_2f_{19},\\
6395400f_{14}^2 &= 10445820f_{12}f_{16} - 5448212f_{10}f_{18} + 1460151f_8f_{20} + 49980f_6f_{22}.
\end{aligned}$$

These equations were mentioned in reference [**4**]. Upon further reflection, one realizes that limiting each term to the product of two f_ms is unnecessary; one may use partitions of integers to find even more possibilities. This produces more identities than we know what to do with. A small sampling includes

$$\begin{aligned}
-35f_3^4 - 945f_6^2 - 972f_5f_7 + 1260f_3f_9 &= 0\\
-88f_3^3f_5 - 1485f_6f_8 - 1584f_5f_9 + 2160f_3f_{11} &= 0\\
3375f_3f_6^2 - 4500f_3^2f_9 + 2916f_5^3 + 125f_3^5 &= 0\\
7776f_5^2f_7 + 4725f_3f_6f_8 - 10080f_3f_5f_9 + 280f_3^4f_5 &= 0\\
-35f_3^3f_8 - 630f_8f_9 - 108f_7f_{10} + 540f_3f_{14} &= 0\\
-35f_3^3f_8 - 630f_8f_9 - 1296f_7f_{10} + 1512f_5f_{12} &= 0\\
-2187f_5^2f_8 - 1800f_3f_6f_9 - 100f_3^4f_6 + 2700f_3^2f_{12} &= 0
\end{aligned}$$

A word should be said about the *Maple* code used to produce these examples. *Maple*'s built-in partition capabilities make the code relatively short. However, since the partition function grows very quickly, even *Maple*'s power will get bogged down after some time. For example, there are 627 partitions of the number 20. To reduce the amount of computing, recall that $f_m = 0$ for $m = 1, 2, 4$. This reduces the number of relevant partitions to 27, a much more manageable number. Lastly, each identity is factored to weed out those which are a multiplicative combination of others.

As a final exploration, one may discover combinations involving f_m with negative m:

$$\begin{aligned}
12f_5f_{-1}f_{-2} - 5f_6f_{-1}^4 + 5f_6f_{-2}^2 &= 0\\
18f_5f_{-1}f_{-2} - 5f_6f_{-1}^4 + 5f_3^2f_{-4} + 20f_6f_{-1}f_{-3} &= 0\\
-42f_5f_{-1}f_{-2} + 5f_7f_{-1}^3f_{-2} - 15f_7f_{-1}f_{-4}f_7f_{-2}f_{-3} &= 0\\
-36f_5f_{-1}^3 - 5f_3^2f_{-1}^4 + 5f_3^2f_{-2}^2 &= 0\\
-168f_5f_{-2} - 36f_7f_{-1}^2f_{-2} + 280f_6f_{-1}^3 + 21f_8f_{-5} - 21f_8f_{-1}f_{-2}^2 + 60f_7f_{-4} &= 0\\
45f_7f_{-1}^4 + 700f_3^2f_{-1}^3 + 84f_3f_5f_{-5} - 84f_3f_5f_{-1}f_{-2}^2 - 225f_7f_{-2}^2 + 3780f_5f_{-1}^2 &= 0\\
1440f_7f_{-1}^4 + 5600f_3^2f_{-1}^3 + 2688f_3f_5f_{-5}&\\
- 168f_3f_5f_{-1}f_{-2}^2 + 3600f_7f_{-2}^2 + 4725f_8f_{-1}f_{-4} &= 0\\
360f_7f_{-1}^4 + 1400f_3^2f_{-1}^3 + 672f_3f_5f_{-5}&\\
+ 168f_3f_5f_{-1}f_{-2}^2 + 3600f_7f_{-2}^2 + 1575f_8f_{-2}f_{-3} &= 0\\
120f_7f_{-1}^2f_{-2} - 840f_6f_{-1}^3 + 105f_8f_{-1}f_{-2}^2 + 56f_3f_5f_{-2}f_{-3} &= 0
\end{aligned}$$

$$-135f_7f_{-1}^4 - 350f_3^2f_{-1}^3 - 42f_3f_5f_{-5}$$
$$+ 42f_3f_5f_{-1}f_{-2}^2 - 225f_7f_{-2}^2 + 450f_7f_{-1}f_{-3} = 0$$
$$27f_8f_{-1}^2f_{-2} - 16f_3f_6f_{-2}f_{-3} + 4f_3f_6f_{-1}^3f_{-2} + 12f_3f_6f_{-1}f_{-4} = 0$$
$$112f_3f_5f_{-1}f_{-3} - 105f_8f_{-1}^2f_{-2} + 1680f_6f_{-2} + 960f_7f_{-3} + 315f_8f_{-4} = 0$$

REFERENCES

1. B. Berndt, *Ramanujan's Notebooks, Part IV*, New York, Springer, 1994.
2. B. Berndt and S. Bhargava, Ramanujan—For lowbrows, *Amer. Math. Monthly* **100** (1993) 644–656.
3. B. Berndt and S. Bhargava, A remarkable identity found in Ramanujan's third notebook, *Glasgow Math. J.* **34** (1992) 341–345.
4. M. Chamberland, Using integer relations algorithms for finding relationships among functions, "Tapas in Experimental Mathematics," Tewodros Amdeberhan and Victor Moll, eds., *Contemp. Math.* **257** (2008) 127–133.
5. M. Hirschhorn, Two or three identities of Ramanujan, *Amer. Math. Monthly* **105** (1998) 52–55.
6. T. Nanjundiah, A note on an identity of Ramanujan, *Amer. Math. Monthly* **100** (1993) 485–487.

Classifying α-Almost-Squares

BOBBE COOPER
University of Georgia
Athens, GA 30605
bcooper@math.uga.edu

Greg Martin's paper "Farmer Ted Goes Natural" [**1**]* addresses the question: "Given a positive integer N, find the dimensions of the rectangle with *integer* side lengths and area at most N whose area-to-perimeter ratio is largest among all such rectangles." (These numbers will be the best integer solutions to the common calculus problem of finding the least amount of fencing necessary for a given area.) He defines an *almost-square* as any number that can be factored into two terms in such a way that the rectangle with sides equal to the factors has area-to-perimeter ratio as large as any rectangle with smaller area. The main theorem of his paper gives a formula for listing the almost-squares in order, without factoring any integers.

In Martin's conclusion, he mentions the variation on the original calculus problem where some of the fencing costs more. This suggests the general problem: "Given a positive integer N, find the dimensions of the rectangle with integer side lengths and area at most N whose area-to-cost ratio is largest among all such rectangles, where two parallel sides are weighted by an integer cost α." Let's help the farmer solve this problem for $\alpha = 5$ and $N = 172$.

Our farmer needs to fence 172 square feet of pasture land. The south and north sides can be fenced in plain barbed wire that costs \$1 per foot, but the east and west sides must be fenced with windproof barbed wire that costs \$5 per foot. What dimensions would minimize the cost of the fence?

Easy calculus shows that each south-north side should be $\sqrt{172 \times 5} \approx 29.3$ ft, and each east-west side should be $\sqrt{172/5} \approx 5.9$ ft, for a total cost of about \$117.30.

*Editor's Note: Another follow-up to Martin's paper sparked a lively exchange of letters in our June 2006 issue about the tension between questions that address real-world problems and questions that spark mathematical interest. We admit that this material falls in the latter category and hope that readers will enjoy it in that spirit.

What if the farmer cannot buy fractional feet of fencing, let alone irrational lengths? She could use the dimensions 43×4, for a cost of \$126, which is better than, say, 86×2 for \$190. However, she could do better if she fenced off just a little less land. If she built her fence with dimensions 34×5, she would get 170 ft^2 for just \$118. This would be about 1.44 ft^2 per dollar, which is cheaper than the 1.37 ft^2 per dollar she would pay to fence the whole 172 ft^2. So 34×5 is more cost-effective—but are there other integer dimensions that give even more area per dollar?

Exploring the problem

Factoring $n = xy$ naturally leads to a rectangle with sides x and y. For fixed $\alpha \geq 1$, the cost—actually half the cost, for convenience—of fencing that rectangle will be considered the cost of the factorization, namely $x + \alpha y$. We want to factor integers so they have small costs.

We define a *best-factored form with respect to* α of a positive integer n, for $\alpha \geq 1$, to be the ordered pair (x, y) such that $xy = n$ and for all other factorizations $wz = n$, $x + \alpha y \leq w + \alpha z$. We denote this by $n = x \times_\alpha y$. (Where possible, the words "with respect to α" will be suppressed for brevity.) Note that in best-factored form, the order of the factors matters and the first factor will always be greater than or equal to the second factor when $\alpha \geq 2$.

If $x \times_\alpha y$ is a best-factored form of n, then the *least-cost* of n is defined as $s_\alpha(n) = x + \alpha y$. We seek integers that represent rectangles with large areas for small costs. So we define the *area-to-cost ratio* of a positive integer n to be $F_\alpha(n) = n/s_\alpha(n)$.

Following Martin, we define a positive integer n to be an *α-almost-square* if and only if $F_\alpha(k) \leq F_\alpha(n)$ for all $k < n$. The α-almost-squares (or α-squares for short) are the integers that beat (or tie) every lower integer for area-to-cost ratio.

Given a constant α, we can calculate the α-squares in order by brute force, calculating every possible factorization of each integer. For instance, 13 is *not* a 5-square, because $F_5(13) = 13/18$, and $F_5(12) = 3/4$ is greater. By brute force, the first fifty 5-squares, are

$$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 16, 18, 20, 22, 24, 26, 28,$$
$$30, 32, 33, 36, 39, 42, 45, 48, 51, 54, 56, 57, 60, 63, 64, 66, 68,$$
$$72, 76, 80, 84, 88, 92, 95, 96, 100, 104, 105, 108, 110 \ldots\}$$

Every integer of the form $(5m)(m)$ for some m is on the list, which is good, since those numbers are precisely the integer calculus solutions to the problem of minimizing cost with a given area. Their close cousins, the numbers $(5m)(m-1)$, also appear in the list for every m. Any other pattern is difficult to see in a list like this—when the numbers are small compared to α, α overpowers the factors. As we shall see, there are patterns that hold for every α, but they are hard to pick out when the numbers are small.

What patterns do you see in the 5-squares between 800 to 900, shown in TABLE 1? Notice that the least-costs of the 5-squares (the right-hand columns) form a nondecreasing sequence. In this nondecreasing sequence of least-costs, consecutive entries are either equal or differ by 1. (This holds true for all lists of α-squares.) After Martin, let's call a set of 5-squares with equal least-cost a *flock* of 5-squares. Generalizing, for any fixed α, a flock will be a set of α-squares with the same least-cost.

If you look closely, you can see that the largest 5-square in each flock has one of two forms: $(5m - \beta)(m)$ or $(5m - \beta)(m-1)$, where β is a nonnegative integer less than 5. We will use the name *flock leaders* for numbers that can be factored this way. For example, $806 = (5 \times 13 - 3)(13)$, so it is a flock leader; the name seems apt, because

TABLE 1: 5-squares between 800 and 900, their best-factored forms, and least-costs

$804 = 67 \times_5 12$	$s_5(804) = 127$	$852 = 71 \times_5 12$	$s_5(852) = 131$
$806 = 62 \times_5 13$	$s_5(806) = 127$	$854 = 61 \times_5 14$	$s_5(854) = 131$
$816 = 68 \times_5 12$	$s_5(816) = 128$	$858 = 66 \times_5 13$	$s_5(858) = 131$
$819 = 63 \times_5 13$	$s_5(819) = 128$	$868 = 62 \times_5 14$	$s_5(868) = 132$
$826 = 59 \times_5 14$	$s_5(826) = 129$	$871 = 67 \times_5 13$	$s_5(871) = 132$
$828 = 69 \times_5 12$	$s_5(828) = 129$	$882 = 63 \times_5 14$	$s_5(882) = 133$
$832 = 64 \times_5 13$	$s_5(832) = 129$	$884 = 68 \times_5 13$	$s_5(884) = 133$
$840 = 60 \times_5 14$	$s_5(840) = 130$	$896 = 64 \times_5 14$	$s_5(896) = 134$
$845 = 65 \times_5 13$	$s_5(845) = 130$	$897 = 69 \times_5 13$	$s_5(897) = 134$

806 is the largest 5-square in its flock. The next flock leader is $819 = (5 \times 13 - 2)(13)$. We take a moment to mention that $(5m)(m)$ and $(5m)(m-1)$ are flock leaders with $\beta = 0$.

You can see these patterns continue down the table. In Lemma 2, we confirm the patterns, by proving that every flock leader is an α-square, and that it is the largest number in its flock.

But these flock leaders are not the only 5-squares. A bit of thought shows that if you take a flock leader, add 5 to the first factor, and subtract 1 from the second factor, you will produce a smaller integer that has the same cost and belongs to the same flock (as long as it is a 5-square). Of course the same thing happens if you take a flock leader and subtract 5 from the first factor, and add 1 to the second factor. This process can be repeated to produce every integer in a particular flock.

It may seem at first that as long as one of these integers is greater than the previous flock leader, it will be a 5-square. After all, it represents a rectangle of larger area than the previous flock leader, and its least-cost is only one larger. But there are exceptions to this: For example let's take $819 = (5 \cdot 13 - 2) \times 13$ as our flock leader. If we subtract 5 from the first factor and add 1 to the second, we get $58 \times 14 = 812$. This number has the right cost and has a greater area than the previous flock leader, 806. But its area-to-cost ratio is 812/128, which is smaller than that of the previous flock leader, 806/127. A fenced area with dimensions 58×14 is not as cost-effective as the fenced area with dimensions 62×13, so 812 is not a 5-square. Is there any way to characterize which integers in a particular flock will be 5-squares, without extending our brute force list? And what about 3-squares, 9-squares, and 4739-squares?

The solution

Martin's solution to the original Farmer Ted problem was elegant; ours is just a little less so. According to our Theorem, the α-squares can be arranged *consecutively* in orderly flocks, each led by a flock leader, and the other numbers in the flock can be obtained by appropriate addition to the first factor and subtraction from the second factor (or vice versa.) There's only one snag: How many times can we continue that process and still obtain α-squares? The numbers p_m and q_m defined in the Theorem answer this question. For convenience, the Theorem is stated in two parts: the first part describes flocks whose leaders have the form $(\alpha m - \beta) \times_\alpha (m)$, and the second part describes flocks whose leaders have the form $(\alpha m - \beta) \times_\alpha (m-1)$. We use $\lfloor x \rfloor$ to mean the greatest integer less than x.

THEOREM PART 1. *For any integer $\alpha \geq 1$, any integer β satisfying $0 \leq \beta < \alpha$, and for all $m \geq 1$, the α-squares between $(\alpha m - \beta - 1)(m) + 1$ and $(\alpha m - \beta)(m)$ (inclusive) belong to a single flock. Define*

$$p_m = \left\lfloor \frac{-\beta}{2\alpha} + \sqrt{\left(\frac{-\beta}{2\alpha}\right)^2 + \frac{m^2}{2\alpha m - \beta - 1}} \right\rfloor \text{ and } k = \alpha(p_m + 1)^2 - \beta(p_m + 1).$$

Then the flock consists of $(\alpha(m + j) - \beta) \times_\alpha (m - j)$ for $j = 0, 1, \ldots, S$, where

$$S = \begin{cases} p_m + 1 & \text{if } m \geq k\left(1 + \sqrt{1 - (\beta + 1)/(\alpha k)}\right) \\ p_m & \text{if } m < k\left(1 + \sqrt{1 - (\beta + 1)/(\alpha k)}\right). \end{cases}$$

THEOREM PART 2. *For any $m \geq 2$ the α-squares between $(\alpha m - \beta - 1) \times (m - 1) + 1$ and $(\alpha m - \beta)(m - 1)$ (inclusive) belong to a single flock. Define*

$$q_m = \left\lfloor \frac{\beta - \alpha}{2\alpha} + \sqrt{\left(\frac{\beta - \alpha}{2\alpha}\right)^2 + \frac{(m-1)^2}{2\alpha m - \beta - \alpha - 1}} \right\rfloor$$

and $k = \alpha(q_m + 1)^2 + \beta(q_m + 1) - \alpha(q_m + 1)$. Then the flock consists of

$$(\alpha(m - j) - \beta) \times_\alpha (m - 1 + j) \quad \text{for} \quad j = 0, 1, \ldots, S,$$

where

$$S = \begin{cases} q_m + 1 & \text{if } m \geq k\left(1 + \sqrt{1 - (\beta + 1)/(\alpha k)}\right) \\ q_m & \text{if } m < k\left(1 + \sqrt{1 - (\beta + 1)/(\alpha k)}\right). \end{cases}$$

All about herding: a close examination of flocks We'll prove the Theorem after a sequence of four lemmas, each verifying a different aspect of flock behavior.

One thing that could make the flocks behave badly, if it ever happened, is this: Suppose one of the numbers on the list, say, $(\alpha(m - p_m + 1) - \beta) \times_\alpha (m + p_m - 1)$, could actually be factored in a different way, so as to have a smaller cost. Then it would belong to a different flock. Lemma 1 assures us that when we have a good factorization of an integer, it's actually the best factorization. It implies that the integers listed in the Theorem are in best-factored form (which is necessary if they are to be in the same flock!). We also see in the proof that the flock leaders have the greatest area in their flock.

LEMMA 1. (WHEN A GOOD FACTORIZATION IS GOOD ENOUGH.) *If an integer satisfying $(\alpha m - \beta - 1)(m - 1) < n \leq (\alpha m - \beta)(m - 1)$ for some m has the form $n = (\alpha(m + d) - \beta)(m - d - 1)$ or $n = (\alpha(m - d) - \beta)(m + d - 1)$, then this is a best-factored form of n and $s_\alpha(n) = 2\alpha m - \alpha - \beta$. Similarly, if an integer satisfying $(\alpha m - \beta - 1)(m) < n \leq (\alpha - \beta)(m)$ for some m has the form $n = (\alpha(m + d) - \beta) \times (m - d)$ or $n = (\alpha(m - d) - \beta)(m + d)$, then this is a best-factored form of n and $s_\alpha(n) = 2\alpha m - \beta$.*

Proof. Either factorization, $n = (\alpha(m + d) - \beta)(m - d - 1)$ or $n = (\alpha(m - d) - \beta)(m + d - 1)$, leads to $s_\alpha(n) \leq 2\alpha m - \alpha - \beta$. If there is a better way to factor n, then $s_\alpha(n) \leq 2\alpha m - \alpha - \beta - 1$. To show this is impossible, we can use calculus to show that the largest integer for which $s_\alpha(n) = 2\alpha m - \alpha - \beta - 1$ is the previous flock leader, $(\alpha m - \beta - 1)(m - 1)$.

We leave to the reader the easy calculus problem of maximizing the area $A = xy$ of the rectangle subject to the cost constraint, $x + \alpha y = 2\alpha m - \alpha - \beta - 1$. The y-value of the best rectangle is $y = m - (\beta + 1)/(2\alpha) - 1/2$. Of course, this is probably not an integer, but it is surely between $y = m - 1$ and $y = m$. The first option yields an area of $A = (\alpha m - \beta - 1) \times (m - 1)$, which can be seen to be larger than the area from the second option, $A = (\alpha m - \alpha - \beta - 1) \times (m)$. Thus, the largest rectangle with a cost of $2\alpha m - \alpha - \beta - 1$ has dimensions $(\alpha m - \beta - 1) \times (m - 1)$. This number is the previous flock leader. Since n is greater than the previous flock leader, $s_\alpha(n) = 2\alpha m - \alpha - \beta$, and there is no better way to factor n.

The other case is similar. ∎

Conversely, if the least-cost of an integer n is $2\alpha m - \beta$ for some m, then n can be written as $n = xy$ with $x \geq y$ and $x + \alpha y = 2\alpha m - \beta$. Letting $d = y - m$ gives us $n = xy = (\alpha(2m - y) - \beta)(y) = (\alpha(m - d) - \beta)(m + d)$. When y is less than m, d will be negative, yielding the form where d is subtracted from the second factor. A very similar argument holds when $s_\alpha(n) = 2\alpha - \alpha - \beta$. Thus every member of a given flock can be written in the form given in the Theorem.

While discussing the table above, we mentioned that numbers of the form $(5m)(m)$ and $(5m)(m - 1)$ always seemed to be 5-squares, and now we can prove it. In the next, we confirm that all the flock leaders are α-squares.

LEMMA 2. (FLOCK LEADERS ARE α-SQUARES) *If* $n = (\alpha m - \beta)(m)$ *or* $n = (\alpha m - \beta)(m - 1)$ *for some* m, *then* n *is an* α*-square.*

Proof. We know from the proof of Lemma 1 that the flock leaders have the greatest area in their flocks, so they have a greater area-to-cost ratio than every other integer with equal cost. We can show that each flock leader has a greater area-to-cost ratio than the previous flock leader, as follows: Suppose $n = (\alpha m - \beta)(m)$. If $\beta < \alpha - 1$, the previous flock leader is $(\alpha m - \beta - 1)(m)$; we need to confirm that

$$\frac{(\alpha m - \beta)(m)}{2\alpha m - \beta} \geq \frac{(\alpha m - \beta - 1)(m)}{2\alpha m - \beta - 1},$$

which is easy. If $\beta = \alpha - 1$, the previous flock leader is $(\alpha m)(m - 1)$; after verifying the inequality analogous to the one above, we conclude that n has a greater area-to-cost ratio than the previous flock leader. The algebra looks very similar when $n = (\alpha m - \beta)(m - 1)$. (Note: when $\beta = \alpha - 1$, the previous flock leader is $(\alpha m - \alpha) \times (m - 1)$.) All this means that each flock leader has a greater area-to-cost ratio than all smaller integers. Therefore, each flock leader is an α-almost-square. ∎

Our next lemma says that every α-square between two flock leaders belongs to the same flock as the larger leader. This means that the least-costs of consecutive α-squares are either the same or differ by 1, and when they differ, the smaller α-square is a flock leader. The proof proceeds by contradiction and we leave it to the reader.

LEMMA 3. (THE FLOCKS DON'T MINGLE) *Every* α*-square* n *satisfying* $(\alpha m - \beta - 1)(m - 1) < n \leq (\alpha m - \beta)(m - 1)$ *has least cost* $s_\alpha(n) = 2\alpha m - \alpha - \beta$. *Likewise, every* α*-square* n *satisfying* $(\alpha m - \beta - 1)(m) < n \leq (\alpha m - \beta)(m)$ *has least cost* $s_\alpha(n) = 2\alpha m - \beta$.

Finally, we'll look at the order in which the α-squares are arranged within a flock. For notational convenience, fix α, β, and m and define

$$n_{+d} = (\alpha(m + d) - \beta)(m - d) \text{ and } n_{-d} = (\alpha(m - d) - \beta)(m + d).$$

Likewise, define

$$r_{+d} = (\alpha(m+d) - \beta)(m-d-1) \text{ and } r_{-d} = (\alpha(m-d) - \beta)(m+d-1).$$

LEMMA 4. (HOW TO ARRANGE α-SQUARES) *For a given m and an integer $d \geq 0$, $n_{-d} \leq n_{+d} < n_{-(d-1)}$, and $r_{+d} < r_{-d} \leq r_{+(d-1)}$, with equality only when $\beta = 0$.*

Instead of a proof, let's just see an example of the algebra involved in this: To show that $n_{-d} \leq n_{+d}$, we look at the inequality

$$(\alpha(m-d) - \beta)(m+d) \leq (\alpha(m+d) - \beta)(m-d).$$

Multiplying each side out as binomials, the first terms cancel, leaving us with $-\beta(m+d) \leq -\beta(m-d)$. This is true because $m+d \geq m-d$. The other cases are very similar.

Proving the theorem Fix integers $\alpha \geq 1$, $0 \leq \beta < \alpha$, and $m \geq 1$. We know from Lemma 3 that every α-square n satisfying $(\alpha m - \beta - 1)(m) < n \leq (\alpha m - \beta)(m)$ belongs to the same flock as our flock leader, $(\alpha m - \beta)(m)$. Thus, by the second statement of Lemma 1, each has the form $n_{+d} = (\alpha(m+d) - \beta)(m-d)$ or $n_{-d} = (\alpha(m-d) - \beta)(m+d)$. We need to know which integers of these forms will be α-squares.

Note that by Lemma 4, the integers $n_{-d}, n_{+d}, n_{-(d-1)}, n_{+(d-1)}, \ldots$ form a non-decreasing sequence of integers in the same flock. This means that if n_{-d} is an α-square, then so are $n_{+d}, n_{-(d-1)}, n_{+(d-1)}, \ldots$. (Since all these integers have the same least-cost, they form an increasing sequence of α-squares.) Let's find the greatest d (corresponding to the smallest n_{-d}) for which n_{-d} is an α-square.

We know by Lemma 2 that the previous flock leader, $(\alpha m - \beta - 1)(m)$, is an α-square. This means that n_{-d} must have a greater cost ratio than the previous flock leader. Setting the area-to-cost ratio of n_{-d} greater than (or equal to) the area-to-cost ratio of the previous flock leader, we have:

$$\frac{(\alpha m - \beta - 1)(m)}{2\alpha m - \beta - 1} \leq \frac{\alpha((m-d) - \beta)(m+d)}{2\alpha m - \beta}.$$

Cross multiplying leads to a quadratic inequality in d. Solving for d gives us

$$\frac{-\beta}{2\alpha} - \sqrt{\left(\frac{\beta}{2\alpha}\right)^2 + \frac{m^2}{2\alpha m - \beta - 1}} \leq d \leq \frac{-\beta}{2\alpha} + \sqrt{\left(\frac{\beta}{2\alpha}\right)^2 + \frac{m^2}{2\alpha m - \beta - 1}}.$$

This means that the largest integer d for which n_{-d} is an α-square is

$$d = \left\lfloor \frac{-\beta}{2\alpha} + \sqrt{\left(\frac{\beta}{2\alpha}\right)^2 + \frac{m^2}{2\alpha m - \beta - 1}} \right\rfloor = p_m.$$

By Lemma 4, there's only one possible α-square greater than the previous flock leader, but less than n_{-p_m}, which is $n_{+(p_m+1)}$. When will $n_{+(p_m+1)}$ be an α-square? Comparing with the previous flock leader yields

$$\frac{(\alpha m - \beta - 1)(m)}{2\alpha m - \beta - 1} \leq \frac{(\alpha(m + (p_m+1)) - \beta)(m - (p_m+1))}{2\alpha m - \beta}.$$

Solving the quadratic inequality in m gives us:

$$m \geq (\alpha(p_m+1)^2 - \beta(p_m+1))\left(1+\sqrt{1-\frac{\beta+1}{\alpha(\alpha(p_m+1)^2-\beta(p_m+1))}}\right).$$

When m meets this condition, our list of α-squares will begin with $n_{+(p_m+1)}$. Otherwise, n_{-p_m} will be the least α-square within the bounds of the lemma.

The proof of the Theorem, Part 2 follows the same reasoning as the first part; the main difference comes from Lemma 4. Since $r_{-d} > r_{+d}$, we end up looking for the largest d such that r_{-d} is an α-square. Similar calculations show this to be q_m. The only possible α-square lower than r_{-q_m} (within the limits of the lemma) is $r_{+(q_m+1)}$. Again, similar calculations show that this will be an α-square only when m meets the condition listed in the first half of the second part of the Theorem. ■

Conclusion

We can now solve the fencing problem of the farmer at the beginning. To fence 172 square feet of land, she needs to know the least flock leader greater than 172. By rounding the noninteger optimum dimensions of 29.33×5.87, she knows that $m = 5$ or $m = 6$, and the larger dimension is either 30 or 29. Checking at most four products shows that $174 = (5 \times 6 - 1)(6)$ is the least flock leader greater than 172, so $m = 6$ and $\beta = 1$, and we are dealing with the first part of the Theorem. Evaluation yields $p_m = 0$; more calculator work reveals that $m < k\left(1+\sqrt{1-(\beta+1)/(\alpha k)}\right)$, so the flock leader is the only member of this flock. The greatest 5-square less than 172, then, must be the previous flock leader, which is $(5 \times 6 - 2)(6) = 168$. Thus, she should fence 168 square feet with dimensions 28×6. This gives her 2.90 square feet per dollar, which is certainly better than any of the other options she had considered.

A possible extension of the problem is to remove the restriction that α is an integer. This would allow us to solve the problem where just one side of the fence, not two parallel sides, is more expensive (this case would correspond to α being a half-integer); also, it could handle more complex situations where, for example, north-south fencing is \$3 and east-west fencing is \$5.

An initial analysis of numerical data (staring at lists of numbers produced by a brute-force α-squares program) suggests that a noninteger α changes the problem considerably. For example, let $\alpha = 3/2$. Brute force calculations show that $27 = 9 \times 3$ is a 3/2-square, with a cost of 27/2; and $30 = 6 \times 5$ is a 3/2-square with a cost of 27/2, so it belongs to the same flock. However, $28 = 7 \times 4$ is a 3/2-square between 27 and 30, and its cost perimeter is only 26/2, so it does not belong to the same flock. It would be very interesting to find a way to characterize these numbers precisely.

Finally, if anyone wishes to see lists of 3-squares, 9-squares, and 4739-almost-squares, they can be found at the author's website [2]. Also at that site is the Perl script used to generate these lists.

Acknowledgment. I would like to thank Dr. Matt DeLong for his guidance and encouragement; my brother Andy for providing the Perl scripts; and the Truth. This research was supported by the Taylor University Science Research Training Program.

REFERENCES

1. Greg Martin, Farmer Ted goes natural, this MAGAZINE **72** (1999) 259–276.
2. The α-squares home page, `http://www.css.tayloru.edu/~mdelong/alpha_squares.html`.

PROBLEMS

PROPOSALS

To be considered for publication, solutions should be received by September 1, 2009.

1816. *Proposed by Mehmet Sahin, Ankara University of Science, Ankara, Turkey.*

Let ABC be a triangle with $a = BC$, $b = CA$, and $c = AB$. Let $A'B'C'$ be another triangle with $B'C' = \sqrt{a}$, $C'A' = \sqrt{b}$, and $A'B' = \sqrt{c}$. Prove that

$$\sin\left(\frac{1}{2}A\right)\sin\left(\frac{1}{2}B\right)\sin\left(\frac{1}{2}C\right) = \cos A' \cos B' \cos C'.$$

1817. *Proposed by Marcos Donnantouni, La Plata, Argentina and José H. Nieto, Maracaibo, Estado Zulia, Venezuela.*

A TV game show has a format in which contestants are asked questions and give answers. Each contestant starts with a score of 0 points. A contestant's score is then calculated as follows: after giving a correct answer, the score is increased by 1; after a wrong answer the score is divided by 2. If a contestant responds to n questions, how many different scores are possible? (As an example, for $n = 3$ there are seven possible scores : 0, 1/4, 1/2, 1, 3/2, 2, and 3.)

1818. *Proposed by Cosmin Pohoata, Tudor Vianu National College of Informatics, Bucharest, Romania.*

Let $n, k, i, i_1, i_2, \ldots, i_k$ be positive integers with $n \geq i = i_1 + i_2 + \cdots + i_k$. Prove that 2^{n-i} is a factor of

$$\sum_{j=0}^{n}\left(\binom{n}{j}\prod_{r=1}^{k}\binom{j}{i_r}\right).$$

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.

Solutions and new proposals should be mailed to Elgin Johnston, Problems Editor, Department of Mathematics, Iowa State University, Ames IA 50011, or mailed electronically (ideally as a LaTeX file) to `ehjohnst@iastate.edu`. All communications, written or electronic, should include **on each page** the reader's name, full address, and an e-mail address and/or FAX number.

1819. *Proposed by Jody M. Lockhart and William P. Wardlaw, U.S. Naval Academy, Annapolis, MD.*

An element a of a ring R is reducible in R if there are elements b and c in R, neither of which are units in R, such that $a = bc$. If a is not reducible then we say a is irreducible. For each integer $m > 1$, let $\mathbb{Z}_m[x]$ denote the ring of polynomials over the ring $\mathbb{Z}_m$ of integers modulo m. For which integers $m > 1$ is the polynomial x irreducible in $\mathbb{Z}_m[x]$?

1820. *Proposed by Christopher J. Hillar, Texas A&M University, College Station, TX.*

A real positive semidefinite matrix is a symmetric matrix with all eigenvalues nonnegative. Prove that if P and Q are real positive semidefinite $n \times n$ matrices with $\operatorname{tr}(PQ) = 0$, then $PQ = 0$.

Quickies

Answers to the Quickies are on page 152.

Q989. *Proposed by Ovidiu Furdui, Campia Turzii, Cluj, Romania.*

Prove that for $-1 \leq x < 1$,

$$\sum_{n=1}^{\infty} \frac{1}{n} \left(\sum_{k=1}^{n} \frac{x^k}{k} - \ln\left(\frac{1}{1-x} \right) \right) = -\frac{1}{2} (\ln(1-x))^2 .$$

Q990. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

Let f be a differentiable real valued function defined on (a, ∞), and suppose that $\lim_{x\to\infty} f(x) = A$, with A finite.

a. Is it necessary that $\lim_{x\to\infty} f'(x) = 0$?

b. If the answer to part a. is no, is it necessary that $\lim_{x\to\infty} f'(x) = 0$ if this limit is assumed to exist?

Solutions

An equilateral condition — April 2008

1791. *Proposed by Mowaffaq Hajja, Yarmouk University, Irbid, Jordan.*

Let ABC be a triangle with circumcenter O, perimeter P, and area K. Prove that if

$$\frac{BC}{P} = \frac{[OBC]}{K} = \frac{1}{3},$$

then ABC is equilateral. (Here $[XYZ]$ denotes the area of triangle XYZ.)

I. *Solution by G.R.A.20 Problem Solving Group, Rome, Italy*
Let $a = BC$, $b = CA$, $c = AB$, $p = P/2$, and let R denote the circumradius of triangle ABC. Noting that $3[OBC] = K$, it follows from Heron's formula that

$$9\left(R + \frac{a}{2}\right)\left(R - \frac{a}{2}\right) \cdot \frac{a}{2} \cdot \frac{a}{2} = p(p-a)(p-b)(p-c).$$

Substituting $R = abc/(4K)$ and rearranging, this becomes

$$9a^4b^2c^2 = 64p^2(p-a)^2(p-b)^2(p-c)^2 + 36a^4p(p-a)(p-b)(p-c). \quad (1)$$

Substituting $p = 3a/2$ into (1) results in

$$4b^2c^2 = (3a-2b)^2(3a-2c)^2 + 3a^2(3a-2b)(3a-2c). \quad (2)$$

Noting that $a = (b+c)/2$ and substituting into (2) then gives

$$2bc = b^2 + c^2.$$

Hence $a = b = c$ and the triangle is equilateral.

II. *Solution by Philip Benjamin, Highland Park High School, Highland Park, NJ.*
Let $a = BC$. Put the triangle in the plane with $B = (-a/2, 0)$, $C = (a/2, 0)$, and $A = (x, y)$ with $y > 0$. Because $BA + AC = 2a$, point A lies on the ellipse with foci at B and C and major semi-axis of length a. This ellipse has equation

$$\frac{x^2}{a^2} + \frac{y^2}{3a^2/4} = 1.$$

Because $[OBC] = K/3$, we conclude that $O = (0, y/3)$. Since $OA = 0B$

$$x^2 + 4y^2/9 = \frac{a^2}{4} + \frac{y^2}{9}, \qquad \text{from which} \qquad \frac{4x^2}{a^2} + \frac{4y^2}{3a^2} = 1.$$

Because (x, y) also lies on the ellipse, it follows that $x = 0$ and $y = \sqrt{3}a/2$, so triangle ABC is equilateral.

Also solved by George Apostolopoulos (Greece), Herb Bailey, Michel Bataille (France), Berry College Dead Poets Society, Jany C. Binz (Switzerland), Bruce S. Burdick, Robert Calcaterra, Minh Can, Chip Curtis, Robert L. Doucette, John Ferdinands, Dmitry Fleischman, Marty Getz and Dixon Jones, Michael Goldenberg and Mark Kaplan, Peter Gressis and Dennis Gressis, Eugen J. Ionascu, Richard A. Jacobson, Victor Y. Kutsenok, Kee-Wai Lau (China), Charles McCraken, Evangelos Mouroukos (Greece), Ruthven Murgatroyd, Gabriel T. Prăjiturá, Kevin A. Roper, Toufic Saad, Jawad Sadek, C. R. Selvaraj, Seshadri Sivakumar, Earl A. Smith, Albert Stadler (Switzerland), H. T. Tang, Hansun To, Michael Vowe (Switzerland), Stuart V. Witt, and the proposer. There was one solution with no name.

Prime divisors — April 2008

1792. *Proposed by H. A. ShahAli, Tehran, Iran.*

Let N be a positive integer. Prove that there is a positive integer n such that $n^2 + 3$ is divisible by at least N distinct primes.

Solution by John H. Smith, Needham, MA.

The result is true if $n^2 + 3$ is replaced by any nonconstant polynomial $f(n) = a_m n^m + a_{m-1}n^{m-1} + \cdots + a_0$ with integer coefficients. We may assume that $a_m > 0$, and hence that there exists an $n_0 > 0$ such that $f(n)$ is positive and increasing on the interval (n_0, ∞).

It is sufficient to show that if for some $n_1 > n_0$, $f(n_1) = p_1^{r_1} \cdots p_k^{r_k}$ is divisible by exactly k distinct primes then, for some $n_2 > n_1$, $f(n_2)$ is divisible by more than k primes. Given such an n_1, let $n_2 = n_1 + p_1^{r_1+1} \cdots p_k^{r_k+1}$. Then

$$f(n_2) \equiv p_1^{r_1} \ldots p_k^{r_k} \pmod{p_1^{r_1+1} \cdots p_k^{r_k+1}}.$$

Thus for each j, $1 \le j \le k$, we have $p_j^{r_j} | f(n_2)$ but $p_j^{r_j+1} \nmid f(n_2)$. Because $f(n_2) > f(n_1) = p_1^{r_1} \ldots p_k^{r_k}$ it follows that $f(n_2)$ has at least $k+1$ prime divisors.

Also solved by Anurag Agarwal, Brian D. Beasley, Robert Calcaterra, John Christopher, Knut Dale (Norway), Gabriel Dospinescu (France) and Marian Tetiva (Romania), Robert L. Doucette, Dmitry Fleischman, Kenneth W. Fogarty, Marty Getz and Dixon Jones, G.R.A.20 Problem Solving Group (Italy), Christopher Hill, Benjamin Hirsch, Eugen J. Ionascu and Alin A. Stancu, Tom Jager, Victor Y. Kutsenok, Harris Kwong, Peter W. Lindstrom, Arturo Magidin, David E. Manes, Kim McInturff, José H. Nieto, Gabriel T. Prăjiturá, Nicholas C. Singer, Albert Stadler (Switzerland), H. T. Tang, Marian Tetiva (Romania), Bob Tomper, and the proposer. There was one solution with no name.

When $A^2 = A^*$ **April 2008**

1793. *Proposed by Götz Trenkler, University of Dortmund, Dortmund, Germany*

Let A be an $n \times n$ matrix with complex entries such that $A^2 = A^*$, where A^* denotes the conjugate transpose of A. Show that

a. $\operatorname{rank}(A + A^*) = \operatorname{rank}(A)$

b. $I_n + A$ is nonsingular.

Solution by Eugene A. Herman, Grinnell College, Grinnell, IA.

b. If $(I_n + A)\mathbf{u} = \mathbf{0}$, then $A\mathbf{u} = -\mathbf{u}$. Hence, after multiplying on the left by A^*, we have

$$A^*A\mathbf{u} = A^2(-\mathbf{u}) = -\mathbf{u}$$

Since all eigenvalues of A^*A are nonnegative, we conclude that $\mathbf{u} = \mathbf{0}$ and therefore $I_n + A$ is nonsingular.

a. Since $A + A^* = A + A^2 = (I_n + A)A$ and $I_n + A$ is nonsingular, $A + A^*$ and A have the same rank.

Note: In fact, since $A + A^* = (I_n + A)A$, $A + A^*$ and A have the same null space, not just the same nullity.

Also solved by Michael Andreoli, Michel Bataille, Paul Budney, Robert Calcaterra, Knut Dale (Norway), Luz M. DeAlba, Michael Goldenberg and Mark Kaplan, Jeffrey M. Groah, Eugen J. Ionascu and Alin A. Stancu, Tom Jager, Victor Y. Kutsenok, Charles Lindsey, Éric Pité (France), Vadim Ponomarenko, Gabriel T. Prăjiturã, Yanir A. Rubenstein, Nicholas C. Singer, John H. Smith, Albert Stadler (Switzerland), and the proposers. There was one solution with no name.

An exponential inequality **April 2008**

1794. *Proposed by Dorin Marghidanu, Colegiul National "A. I. Cuza," Corabia, Romania*

Let $x_1, x_2, \ldots, x_n \geq e$. Prove that

$$x_1^{\frac{x_1+x_2+\cdots+x_n}{x_1}} + x_2^{\frac{x_2+\cdots+x_n}{x_2}} + \cdots + x_{n-1}^{\frac{x_{n-1}+x_n}{x_{n-1}}} + x_n \geq x_1 + 2x_2 + \cdots + (n-1)x_{n-1} + nx_n.$$

Solution by Michel Bataille, Rouen, France

The function f defined by $f(x) = (\log x)/x$ is decreasing on $[e, \infty)$. Thus if $a_1, a_2, \ldots, a_k \geq e$, then

$$\frac{\log(a_1)}{a_1} \geq \frac{\log(a_1 + a_2 + \cdots + a_k)}{a_1 + a_2 + \cdots + a_k},$$

and exponentiation gives

$$a_1^{\frac{a_1+a_2+\cdots+a_k}{a_1}} \geq a_1 + a_2 + \cdots + a_k.$$

Thus

$$\begin{aligned} x_1^{\frac{x_1+x_2+\cdots+x_n}{x_1}} &\geq x_1 + x_2 + \cdots + x_n \\ x_2^{\frac{x_2+x_3+\cdots+x_n}{x_2}} &\geq x_2 + x_3 + \cdots + x_n \\ &\vdots \\ x_{n-1}^{\frac{x_{n-1}+x_n}{x_{n-1}}} &\geq x_{n-1} + x_n \\ x_n &\geq x_n. \end{aligned}$$

Adding these inequalities gives the desired result.

Also solved by Robert Calcaterra, Minh Can, Chip Curtis, Knut Dale (Norway), Robert L. Doucette, Dmitry Fleischman, Marty Getz and Dixon Jones, Eugen J. Ionascu, Tom Jager, Hidefumi Katsuura, Evangelos Mouroukos (Greece), Paolo Perfetti, Gabriel T. Prăjitură, Phillip P. Ray, Toufic Saad, C. R. Selvara and Suguna Selvaraj, Albert Stadler (Switzerland), and the proposer.

A many-to-one function **April 2008**

1795. *Proposed by Jeff Groah, Montgomery College, Conroe, TX.*

Find a function $f : [0, 1] \to [0, 1]$ such that for each nontrivial interval $I \subseteq [0, 1]$, we have $f(I) = [0, 1]$.

I. *Solution by Vadim Ponomarenko, San Diego State University, San Diego, CA.*
Each $x \in [0, 1]$, can be expressed in base 3: $x = [x_0.x_1x_2x_3 \ldots]_3$, where each $x_i \in \{0, 1, 2\}$ and the representation does not end in an infinite string of 2s. If the base three representation has no digits 1 or infinitely many digits 1 among $x_1, x_2, \ldots$ then define $f(x) = 1$. (so $f(1) = 1$.) If there are a positive finite number of digits 1 among $x_1, x_2, \ldots$, find d so that $x_d = 1$ and $x_k \neq 1$ for $k > d$, and define

$$f(x) = [0.x'_{d+1}x'_{d+2}x'_{d+3} \ldots]_2 \qquad \text{where} \qquad 0' = 0,\ 2' = 1,$$

and we consider the result as a number in base 2. Each element of $[0, 1)$ has a preimage in the interval $[a, b]$, for any $a = 0.x_1 \ldots x_d$ and $b = a + 3^d$, with $x_d = 1$. In fact, each element of $[0, 1]$ has a preimage in any interval of this type. (Note that $f(a + 3^d/2) = f(0.x_1 \ldots x_d 11 \ldots) = 1$.) Now, let I be a nontrivial interval. Then there is a $k > 0$ so that $[c, c + 3^k] \subseteq I$, for some $c = [0.y_1 \ldots y_k]_3$, where each $y_i \in \{0, 1, 2\}$. We set $a = c + 3^{-k-1}$ and $b = a + 3^{-k-1}$. Then $[a, b]$ is an interval of the desired type, with $I \subseteq [a, b] \subset [c, c + 3^{-k}] \subseteq I$. This completes the construction.

II. *Solution by Jerrold W. Grossman, Oakland University, Rochester, MI.*
Let $\aleph$ be the cardinality of $U = [0, 1]$. The set J of nontrivial subintervals of U has cardinality $\aleph \cdot \aleph = \aleph$. Viewing $\aleph$ as an ordinal number (and noting that this implies that every initial segment of $\aleph$ has strictly smaller cardinality), we have a well-ordering $y_1, y_2, \ldots, y_\omega, \ldots$ of U and a well-ordering $I_1, I_2, \ldots, I_\omega, \ldots$ of J. Now by double transfinite induction define f as follows. For each $i \in \aleph$, for each $j \in \aleph$, choose a number $x_j \in I_j$ that has not yet been assigned a value and set $f(x_j) = y_i$; it is possible to find such a number because at each stage in the process the set of numbers that have so far been assigned a value has cardinality smaller than $\aleph$ and so does not exhaust I_j. For any $x \in U$ that has not been assigned a value upon completion of this process, arbitrarily set $f(x) = x$. The desired property of f is clear.

Also solved by Michael Andreoli, Michel Bataille, Tom Beatty, Michael W. Botsko, Paul Budney, Bruce S. Burdick, Robert Calcaterra, Elliott Cohen, A. K. Desai and K. V. Thaker, Marty Getz and Dixon Jones, Michael Goldenberg and Mark Kaplan, Eugen J. Ionascu, Jean-Christophe Laugier (France), Evangelos Mouroukos, Stephen

Noltie, Northwestern University Math Problem Solving Group, Paolo Perfetti (Italy), Gabriel T. Prăjitură, Nicholas C. Singer, Albert Stadler (Switzerland), Tony Tam, Marian Tetiva (Romania), Dave Trautman, Stuart V. Witt, and the proposer.

Answers

Solutions to the Quickies from page 148.

A989. Because $1 + t + \cdots + t^{n-1} = \frac{1-t^n}{1-t}$, we have, by integration,

$$\sum_{k=1}^{n} \frac{x^k}{k} = -\int_0^x \frac{t^n}{1-t}\,dt - \ln(1-x).$$

Thus

$$\sum_{n=1}^{\infty} \frac{1}{n}\left(\sum_{k=1}^{n} \frac{x^k}{k} - \ln\left(\frac{1}{1-x}\right)\right) = -\sum_{n=1}^{\infty} \frac{1}{n}\int_0^x \frac{t^n}{1-t}\,dt = -\int_0^x \frac{1}{1-t}\left(\sum_{n=1}^{\infty} \frac{t^n}{n}\right)dt$$

$$= \int_0^x \frac{\ln(1-t)}{1-t}\,dt = -\frac{1}{2}(\ln(1-x))^2.$$

Note that the interchange of the order of summation and integration is valid because $\sum_{n=1}^{\infty} t^n/n = -\ln(1-t)$ converges on $-1 \le t < 1$, so the series converges uniformly on any closed interval in $(-1, 1)$. For $x = -1$, the desired equality follows from the Abel summation by parts formula.

A990.

a. The answer is no. As an example, let

$$f(x) = \frac{\sin(2x^2)}{x}$$

and note that $\lim_{x\to\infty} f(x) = 0$. On the other hand,

$$f'(x) = \frac{4x^2\cos(2x^2) - \sin(2x^2)}{x^2},$$

so $\lim_{x\to\infty} f'(x)$ does not exist.

b. The answer is yes. To prove this, let $x > a$. By the Mean Value Theorem there is a $c_x \in (x, x+1)$ such that

$$f(x+1) - f(x) = f'(c_x).$$

Taking the limit of both sides as $x \to \infty$, and noting that $c_x \to \infty$ as $x \to \infty$, we find

$$0 = A - A = \lim_{x\to\infty} f'(c_x) = \lim_{x\to\infty} f'(x),$$

where the last equality holds because $\lim_{x\to\infty} f'(x)$ exists.

REVIEWS

PAUL J. CAMPBELL, *Editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles, books, and other materials 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.

Vaaler, Leslie Jane Federer, and James W. Daniel, *Mathematical Interest Theory*, 2nd ed., MAA, 2009; xvii + 475 pp, $89.95 ($71.95 to MAA members). ISBN 978-0-88385-754-0. Vaaler, Leslie Jane Federer, *Student Solutions Manual for Mathematical Interest Theory*, MAA, 2009; ix + 107 pp, $34 ($26.95 to MAA members). ISBN 978-0-88385-755-7.

In my December 2008 column, I asked about the role of mathematics in the economic crisis. Now the question is, What will the economic crisis do to student interest in studying financial mathematics? In Spring 2008 I taught a semester course on that topic from the first edition of this textbook. The book deals with the comforting part of financial mathematics that is deterministic: It ignores uncertainty, variability, and risk, assuming that all payments of an annuity or a loan will be made and be made on time (don't laugh too hard—that's the way that it's supposed to happen). Yet the subject, and its presentation in this book, is not for the faint of heart. Solving the problems requires working systematically from first principles to arrive at expressions to evaluate; instructor and students must recognize and remember a large number of symbols and their corresponding formulas, as well as develop facility with TI's Business Analyst II Plus calculator; and calculus is used for a few topics. The book does not cover fully the current syllabus for the Society of Actuaries exam in Financial Mathematics.

Benjamin, Arthur T., and Ezra Brown (eds.), *Biscuits of Number Theory*, MAA, 2009; xiii + 313 pp, $62.50 (members: $49.95). ISBN 978-0-88385-340-5.

A biscuit: "not too big, easily digested, and makes you feel all warm and fuzzy when you're through." This "box" of "biscuits" of number theory contains 40 articles reprinted mostly from MAA journals (plus one original contribution). The articles are grouped by topic, each introduced by a short essay: arithmetic; primes; irrationality and continued fractions; sums of squares and polygonal numbers; Fibonacci numbers; number-theoretic functions; and elliptic curves, cubes, and Fermat's Last Theorem. Fortunately, undergraduate courses in number theory are not burdened by a preordained syllabus. The articles here are rich in ideas without being demanding of background; they could be either starting points for such a course or departure points from it to other exciting destinations.

Brummelen, Glen Van, *The Mathematics of the Heavens and the Earth: The Early History of Trigonometry*, Princeton University Press, 2009; xvii + 329 pp, $39.50. ISBN 978-0-691-12973-0.

Apart from a work in German a century ago and one in Russian in 1990, this is the first modern book-length history of trigonometry. This volume covers developments through 1550 (Copernicus's heliocentric model is the breakpoint); a projected sequel will continue from then. Because trigonometry arose in part from astronomy, some acquaintance with the concepts and terminology of spherical astronomy is needed; author Van Brummelen provides a short introduction before embarking on his tale. Egypt and Babylonia come in for brief discussion before major

sections on Greece, India, Islam, and the West, including methods to make trigonometric tables. The book is replete with figures and short translations (with explanations) of original sources.

Hopkins, Brian (ed.), *Resources for Teaching Discrete Mathematics: Classroom Projects, History Modules, and Articles*, MAA, 2009; xiv + 323 pp, $54.95 (P) (members: $44.95). ISBN 978-0-88385-184-5.

Courses in discrete mathematics have not become as popular as calculus nor have they had the 300 years to settle into a fixed syllabus. Desire to include specific topics for one major constituency, computer science students, has distinguished courses in "discrete structures" from others with more specifically mathematical content, such as combinatorics and graph theory. Meanwhile, other discrete mathematics courses are designed mainly to introduce students to proof, so that the actual mathematical content is secondary. This book strives to offer something for every kind of discrete mathematics course. It contains 19 classroom-tested learning modules, each designed for one to four class periods. Each module includes notes for the instructor (with references), reproducible worksheets for students, solutions, and sometimes open questions. Some modules introduce a topic, others extend one, and some go into applications and the use of technology. Additionally, there are 11 projects on the history of a topic, each intended to extend over a longer period of time, plus five further articles, including two on pedagogy.

Smullyan, Raymond M., *Logical Labyrinths*, A K Peters, 2009; viii + 327 pp, $49. ISBN 978-0-56881-4343-8.

Author Smullyan is known for his research in mathematical logic and his exposition of the subject, but more widely for his books of logical puzzles and paradoxes: *What Is the Name of This Book?* (1978), *The Lady or the Tiger?* (1982), *To Mock a Mockingbird...* (1985), and several more since. Those books feature liars and truth-tellers, knights and knaves, and lots of self-referentialism. This new book combines the recreational aspect with serious mathematical logic: It starts with puzzles, tours propositional logic, gets to predicate logic one-third of the way through the book, and then takes a detour to infinity and König's lemma. In the last third, it gets into serious results in first-order logic, ending with Gödel incompleteness. Most of the development is through problems, whose solutions are provided in the back. The author asserts, "After having read this book, you will have the knowledge of a typical one-semester graduate course in symbolic logic, and you will then have the preparation to read ... much of the general literature in the field." That is likely true for a motivated reader who reasons carefully, solves most of the problems, and internalizes the necessary notation.

Shortz, Will, A new puzzle challenges math skills, *New York Times* (9 February 2009) C6, `http://www.nytimes.com/2009/02/09/arts/09ken.html`. Gaffney, Matt, I was told there would be no math: Will KenKen be the next Sudoku or a passing puzzle fad?, `http://www.slate.com/id/2211595`. KenKen: The world's most ADDictive puzzle!, `http://www.kenken.com/`.

Inspiration for mathematics can come from unexpected sources, but seizing the opportunity is another matter. In October 2004, a fellow American visitor at the University of Augsburg in Germany, a colleague in American literature, showed me a German magazine with a single page of puzzles that involved filling numbers into grids of various sizes. Did I know where he could get more such puzzles, and was there mathematical theory that would help solve them? I couldn't help him on either count. One month later, unknown to either of us, similar puzzles began to appear daily in *The Times* of London. I had missed an opportunity to get in on the ground floor of research in sudoku(not to mention lucrative puzzle and how-to books). Who knew that it would become a craze? Now an heir apparent has emerged in KenKen (Japanese for "cleverness squared"), invented by Japanese teacher Tetsuya Miyamoto. Again pioneered by *The Times*, KenKen now appears in two sizes in the *New York Times* next to the crossword puzzle. Like sudoku, KenKen requires filling in digits without repetition in a row or column; unlike sudoku, certain boxes must contain digits fulfilling an arithmetic condition (e.g., their product must be 60). As commentator Gaffney says, "The marketing wheels, greased by the promise of Sudoku-style riches, are already in motion." Indeed, puzzle editor Shortz has written a large proportion of the 50 or so books on KenKen that are already available.

37th United States of America Mathematical Olympiad

CECIL ROUSSEAU
University of Memphis
Memphis, TN 38152-3240
rousseac@msci.memphis.edu

STEVEN R. DUNBAR
MAA American Mathematics Competitions
University of Nebraska-Lincoln
Lincoln, NE 68588-0658
sdunbar@maa.org

Problems

1. Prove that for each positive integer n, there are pairwise relatively prime integers $k_0, k_1, \ldots, k_n$, all strictly greater than 1, such that $k_0 k_1 \cdots k_n - 1$ is the product of two consecutive integers.
2. Let ABC be an acute, scalene triangle, and let M, N, and P be the midpoints of $\overline{BC}$, $\overline{CA}$, and $\overline{AB}$, respectively. Let the perpendicular bisectors of $\overline{AB}$ and $\overline{AC}$ intersect ray AM in points D and E respectively, and let lines BD and CE intersect in point F, inside of triangle ABC. Prove that points A, N, F, and P all lie on one circle.
3. Let n be a positive integer. Denote by S_n the set of points (x, y) with integer coordinates such that

$$|x| + \left| y + \frac{1}{2} \right| < n.$$

A *path* is a sequence of distinct points $(x_1, y_1), (x_2, y_2), \ldots, (x_\ell, y_\ell)$ in S_n such that, for $i = 2, \ldots, \ell$, the distance between (x_i, y_i) and (x_{i-1}, y_{i-1}) is 1 (in other words, the points (x_i, y_i) and (x_{i-1}, y_{i-1}) are neighbors in the lattice of points with integer coordinates).

Prove that the points in S_n cannot be partitioned into fewer than n paths (a partition of S_n into m paths is a set $\mathcal{P}$ of m nonempty paths such that each point in S_n appears in exactly one of the m paths in $\mathcal{P}$).

4. Let $\mathcal{P}$ be a convex polygon with n sides, $n \geq 3$. Any set of $n - 3$ diagonals of $\mathcal{P}$ that do not intersect in the interior of the polygon determine a *triangulation* of $\mathcal{P}$ into $n - 2$ triangles. If $\mathcal{P}$ is regular and there is a triangulation of $\mathcal{P}$ consisting of only isosceles triangles, find all the possible values of n.
5. Three nonnegative real numbers r_1, r_2, r_3 are written on a blackboard. These numbers have the property that there exist integers a_1, a_2, a_3, not all zero, satisfying $a_1 r_1 + a_2 r_2 + a_3 r_3 = 0$. We are permitted to perform the following operation: find two numbers x, y on the blackboard with $x \leq y$, then erase y and write $y - x$ in

its place. Prove that after a finite number of such operations, we can end up with at least one 0 on the blackboard.

6. At a certain mathematical conference, every pair of mathematicians are either friends or strangers. At mealtime, every participant eats in one of two large dining rooms. Each mathematician insists upon eating in a room which contains an even number of his or her friends. Prove that the number of ways that the mathematicians may be split between the two rooms is a power of two (i.e., is of the form 2^k for some positive integer k).

Solutions Following are solution sketches with the essential ideas for each problem. For interested readers, detailed solutions with figures and multiple approaches are at the website of the MAA American Mathematics Competitions: `http://www.unl.edu/amc/e-exams/e8-usamo/archiveusamo.shtml`. The website has solutions and generalizations developed by the USAMO Committee and the contestants.

1. We proceed by induction. The case $n = 1$ is clear with $k_0 = 3$ and $k_1 = 7$. Assume now that for $n > 1$ there are pairwise relatively prime integers $1 < k_0 < k_1 < \cdots < k_n$ such that $k_0k_1 \cdots k_n - 1 = a_n(a_n - 1)$, for some positive integer a_n. Then choosing $k_{n+1} = a_n^2 + a_n + 1$ yields

$$k_0k_1 \cdots k_{n+1} = (a_n^2 - a_n + 1)(a_n^2 + a_n + 1) = a_n^4 + a_n^2 + 1,$$

so $k_0k_1 \cdots k_{n+1} - 1$ is the product of consecutive integers a_n^2 and $a_n^2 + 1$. Moreover,

$$\gcd(k_0k_1 \cdots k_n, k_{n+1}) = \gcd(a_n^2 - a_n + 1, a_n^2 + a_n + 1) = 1,$$

hence $k_0, k_1, \ldots, k_{n+1}$ are pairwise relatively prime. ■

Titu Andreescu suggested this problem.

2. Invert the figure about a circle centered at A. Let X' denote the image of the point X under this inversion. Find point F_1' so that $AB'F_1'C'$ is a parallelogram Let Z' denote the center of this parallelogram. Note that $\triangle BAC \sim \triangle C'AB'$ and $\triangle BAD \sim \triangle D'AB'$. Because M is the midpoint of BC and Z' is the midpoint of $B'C'$, we also have $\triangle BAM \sim \triangle C'AZ'$. Thus

$$\angle AF_1'B' = \angle F_1'AC' = \angle Z'AC' = \angle MAB = \angle DAB = \angle DBA = \angle AD'B'.$$

Hence quadrilateral $AB'D'F_1'$ is cyclic. By a similar argument, quadrilateral $AC'E'F_1'$ is cyclic. Because the images under the inversion of lines BDF and CFE are circles that intersect in A and F', it follows that $F_1' = F'$.

Next note that B', Z', and C' are collinear and are the images of P', F', and N', respectively, under a homothety centered at A and with ratio $1/2$. It follows that P', F', and N' are collinear, and then that the points A, P, F, and N lie on a circle. ■

Zuming Feng suggested this problem. Gabriel Carroll suggested the given solution.

3. Color the points in S_n as follows:
   (a) if $y \geq 0$, color (x, y) white if $x + y - n$ is even and black if $x + y - n$ is odd;
   (b) if $y < 0$, color (x, y) white if $x + y - n$ is odd and black if $x + y - n$ is even.

Consider a path $(x_1, y_1), (x_2, y_2), \ldots, (x_\ell, y_\ell)$ in S_n. A pair of successive points (x_{i-1}, y_{i-1}) and (x_i, y_i) in the path is called a pair of successive black points if both points in the pair are colored black.

Suppose now that the points of S_n are partitioned into m paths and the total number of successive pairs of black points in all paths is k. By breaking the paths

at each pair of successive black points, we obtain $k+m$ paths in each of which the number of black points exceeds the number of white points by at most one. Therefore the total number of black points in S_n cannot exceed the number of white points by more than $k+m$. On the other hand, the total number of black points in S_n exceeds the total number of white points by exactly $2n$ (there is exactly one more black point in each row of S_n). Therefore $2n \le k+m$. There are exactly n adjacent black points in S_n (call two points in S_n *adjacent* if their distance is 1), namely the pairs $(x, 0)$ and $(x, -1)$, for $x = -n+1, -n+3, \ldots, n-3, n-1$. Therefore $k \le n$ (the number of successive pairs of black points in the paths in the partition of S_n cannot exceed the total number of adjacent pairs of black points in S_n) and so $n \le m$. ∎

Gabriel Carroll suggested this problem.

4. The answer is $n = 2^{m+1} + 2^k$, where m and k are nonnegative integers.

LEMMA. *Let $\mathcal{Q} = Q_0Q_1 \ldots Q_t$ be a convex polygon with $Q_0Q_1 = Q_1Q_2 = \cdots = Q_{t-1}Q_t$. Suppose that $\mathcal{Q}$ is cyclic and its circumcenter does not lie in its interior. If there is a triangulation of $\mathcal{Q}$ consisting only of isosceles triangles, then $t = 2^a$, where a is a positive integer.*

Let $\mathcal{P} = P_1P_2 \ldots P_n$ denote the regular polygon. There is an isosceles triangle in the triangulation such that the center of $\mathcal{P}$ lies within the boundary of the triangle. Without loss of generality, we may assume that $P_1P_iP_j$, with $P_1P_i = P_1P_j$ (that is, $P_j = P_{n-i+2}$), is this triangle. Applying the Lemma to the polygons $P_1 \ldots P_i$, $P_i \ldots P_j$, and $P_j \ldots P_1$, we conclude that there are $2^m - 1$, $2^k - 1$, $2^m - 1$ (where m and k are nonnegative integers) vertices in the interiors of the minor arcs $\widehat{P_1P_i}$, $\widehat{P_iP_j}$, $\widehat{P_jP_1}$, respectively. (In other words, $i = 2^m + 1$, $j = 2^k + i$.) Hence

$$n = 2^m - 1 + 2^k - 1 + 2^m - 1 + 3 = 2^{m+1} + 2^k,$$

where m and k are nonnegative integers. The above discussion can easily lead to a triangulation consisting of only isosceles triangles for $n = 2^{m+1} + 2^k$. ∎

Gregory Galperin suggested this problem.

5. If two of the a_i vanish, say a_2 and a_3, then r_1 must be zero and we are done. Assume at most one a_i vanishes. If any one a_i vanishes, say a_3, then $r_2/r_1 = -a_1/a_2$ is a nonnegative rational number. Write this number in lowest terms as p/q, and put $r = r_2/p = r_1/q$. We can then write $r_1 = qr$ and $r_2 = pr$. Performing the Euclidean algorithm on r_1 and r_2 will ultimately leave r and 0 on the blackboard. Thus we are done again.

Thus it suffices to consider the case where none of the a_i vanishes. We may also assume none of the r_i vanishes, as otherwise there is nothing to check. In this case we will show that we can perform an operation to obtain r_1', r_2', r_3' for which either one of r_1', r_2', r_3' vanishes, or there exist integers a_1', a_2', a_3', not all zero, with $a_1'r_1' + a_2'r_2' + a_3'r_3' = 0$ and

$$|a_1'| + |a_2'| + |a_3'| < |a_1| + |a_2| + |a_3|.$$

After finitely many steps we must arrive at a case where one of the a_i vanishes, in which case we finish as above.

If two of the r_i are equal, then we are immediately done by choosing them as x and y. Hence we may suppose $0 < r_1, r_2 < r_3$. Since we are free to negate all the a_i, we may assume $a_3 > 0$. Then either $a_1 < -a_3/2$ or $a_2 < -a_3/2$ (otherwise $a_1r_1 + a_2r_2 + a_3r_3 > (a_1 + a_3/2)r_1 + (a_2 + a_3/2)r_2 > 0$). Without loss of generality, we may assume $a_1 < -a_3/2$. Then choosing $x = r_1$ and $y = r_3$ gives

the triple $(r_1', r_2', r_3') = (r_1, r_2, r_3 - r_1)$ and $(a_1', a_2', a_3') = (a_1 + a_3, a_2, a_3)$. Since $a_1 < a_1 + a_3 < a_3/2 < -a_1$, we have $|a_1'| = |a_1 + a_3| < |a_1|$ and hence this operation has the desired effect. ■

Kiran Kedlaya suggested this problem.

6. Let n be the number of participants at the conference. We proceed by induction on n.

If $n = 1$, then we have one participant who can eat in either room; that gives us total of $2 = 2^1$ options.

Let $n \geq 2$. The case in which some participant, P, has no friends is trivial. In this case, P can eat in either of the two rooms, so the total number of ways to split n participants is twice as many as the number of ways to split $(n - 1)$ participants besides the participant P. By induction, the latter number is a power of two, 2^k, hence the number of ways to split n participants is 2^{k+1}. So we assume that every participant has at least one friend. We consider two different cases separately:

Case 1: *Some participant, Z, has an odd number of friends.*

Then the claim is that the number of possible seatings is unchanged after removing Z and reversing the relationship between X and Y in each pair (X, Y) of Z's friends.

Case 2: *Each participant has an even number of friends.*

In this case, each valid split of participants in two rooms gives us an even number of friends in either room. ■

Sam Vandervelde suggested this problem.

2008 Olympiad Results The top twelve students on the 2004 USAMO were (in alphabetical order):

David Benjamin	11	William Henry Harrison High School	West Lafayette	IN
TaoRan Chen	12	Bayside High School	Freshmeadows	NY
Paul Christiano	12	The Harker School	San Jose	CA
Sam Elder	12	Poudre High School	Fort Collins	CO
Shaunak Kishore	12	Unionville-Chaddsford High School	West Chester	PA
Delong Meng	11	Baton Rouge Magnet High School	Baton Rouge	LA
Evan O'Dorney	9	Venture High School	Berkeley	CA
Qinxuan Pan	11	Thomas S Wootton High School	Gaithersburg	MD
David Rolnick	11	Home School	Rupert	VT
Colin Sandon	12	Essex High School	Essex Junction	VT
Krishanu Sankar	12	Horace Mann High School	Hastings on Hudson	NY
Alex Zhai	12	University Laboratory High School	Champaign	IL

Colin Sandon and Evan O'Dorney were the winners of the Samuel Grietzer-Murray Klamkin Award, given to the top scorers on the USAMO. Colin Sandon and Evan O'Dorney tied for first place and were awarded college scholarships of $10,000 by the Akamai Foundation. Krishanu Roy Sankar and Qinxuan Pan tied for second place and were awarded scholarships of $7,000 by the Akamai Foundation. Delong Meng and Shaunak Kishore tied for throed place and were awarded scholarships of $5,000 by the Akamai Foundation. The Clay Mathematics Institute Award for a solution of outstanding elegance and carrying a $5,000 cash prize was presented to Evan O'Dorney for his solution on Problem 2.

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW
Washington, DC 20036