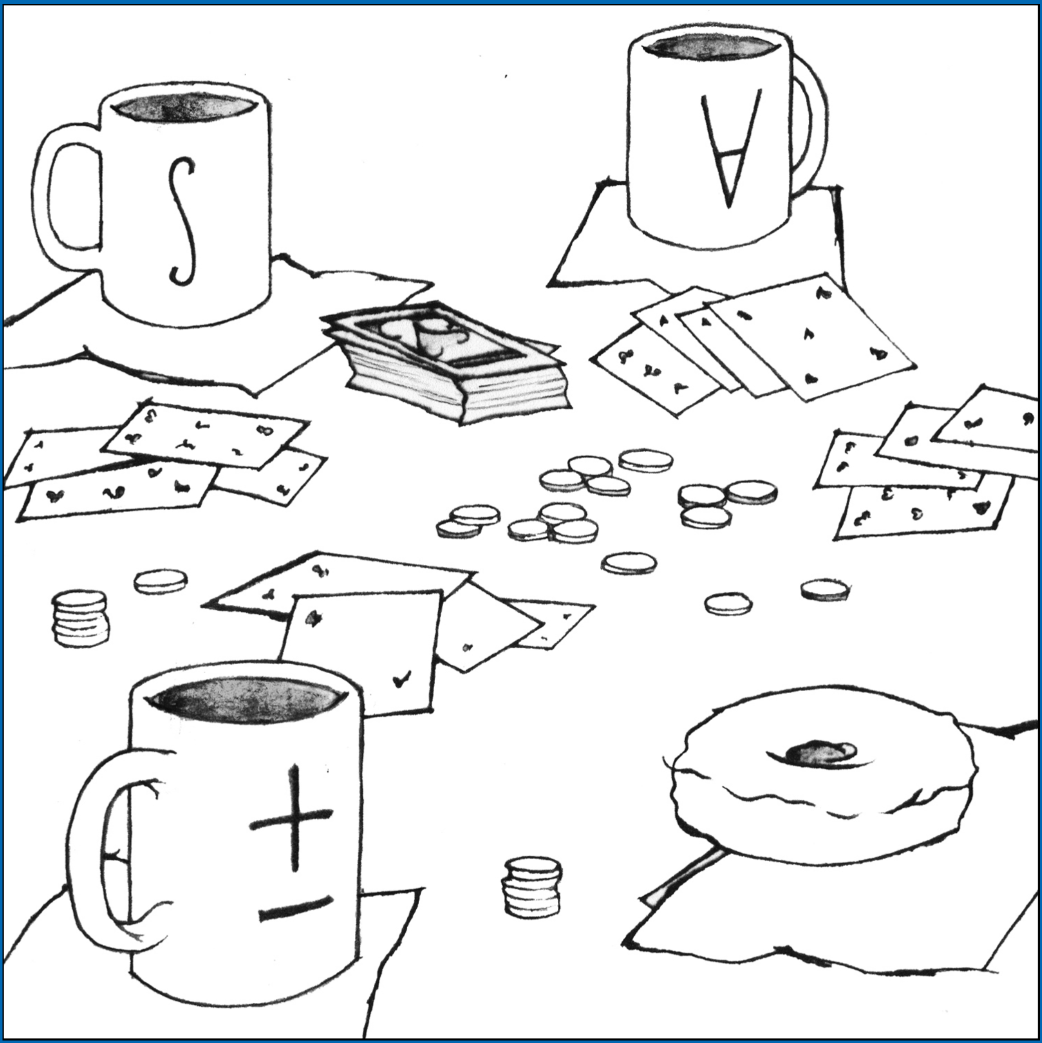
Vol. 84, No. 1, February 2011

MAA

MATHEMATICS MAGAZINE



Who are the players? (p. 3)

- The Inner Life of Markov Chains
- Crosscut Convex Quadrilaterals
- Zetas, Sequences, Rulers, and Butterflies

An Official Publication of the **Mathematical Association of America**

MAA
MATHEMATICAL ASSOCIATION OF AMERICA 1529 Eighteenth St., NW • Washington, DC 20036

CONTENTS

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. 83, at pages 73–74, and is available at the *Magazine*'s website 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 directly to the editor at mathmag@maa.org. A brief message containing contact information and 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 by Susan Stromquist

MATHEMATICS MAGAZINE (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D.C. 20036 and Lancaster, PA, 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

Vol. 84, No. 1, February 2011

MAA

MATHEMATICS MAGAZINE

LETTER FROM THE EDITOR

In this issue, Robert Mena and Will Murray introduce us to the inner workings of a Markov chain. Many of us know how to analyze Markov chains mechanically using large matrices. But the process in this article has a succession of stages, and the authors show us that by tracking it carefully through each stage, we get more detailed results and more valuable insights.

Rick Mabry's article is about quadrilaterals cut by straight lines, and he finds relationships among the areas formed. He gives visual proofs, and then tracks some of the relationships back to a problems journal from the 1940's. Will our own writings still attract researchers sixty years from now? Of course they will; we're mathematicians!

In the notes, Greg Markowsky connects two beautiful theorems of geometry, and Ernst Scheufens helps us understand the values of the zeta function. In between we are treated to prime-divisible sequences, well-chosen sums of consecutive integers, and a computer game that generalizes the classical secretary problem. Roger Alperin and Vladimir Drobot tell us how to construct a ruler, if we are being charged a fee for each mark.

The Putnam feature at page 74 required scrambling to meet the deadline. We thank the Putnam Committee for that task and for everything else they do.

There is a story behind the cover. What old joke does it illustrate? To find out, start by reading the first paragraph on page 3.

The first draft of the cover illustration had Styrofoam coffee cups, without handles. That's because the artist is a graduate student. Graduate students think all coffee cups are Styrofoam, or perhaps just paper—that is a common experience across all disciplines! I was able to explain why the coffee cups in the illustration needed handles. The artist did wonder, however, why the players in this evening poker game were drinking coffee, rather than, say, beer. That's easy, of course. Mathematicians can't turn beer into theorems!

The cover has a new color, too. Most editors choose to alternate between two colors, and I have chosen Crimson and Blue, the colors of the University of Kansas, my undergraduate college. Crimson is also a color of two other universities I attended: Pittsburg State University in Kansas (Crimson & Gold) and Harvard University (just Crimson). There is a historical connection. The Kansas schools are among many that have chosen Crimson to honor the inspiration they have received from Harvard.

Walter Stromquist, Editor

ARTICLES

Markov Chains for Collaboration

ROBERT MENA
rmena@csulb.edu

WILL MURRAY
wmurray@csulb.edu
California State University, Long Beach
Long Beach, CA 98040-1001

Introduction: Who wants to be a collaborator?

The math department at New Alarkania State University is comprised of Alan the analyst, Lorraine the logician, Stacy the statistician, and Tom the topologist. Each one is desperate for collaborators, so they start a Friday poker series. Each one is equally skilled, and they agree that the loser of each week's game (the first to run out of money) will renounce his or her former field and join the research team of the biggest winner.

In the first week, Stacy wins and Tom loses, so Tom gives up topology and joins Stacy to study statistics. The following week, Lorraine wins and Stacy loses, so Stacy becomes a logician. Next, Stacy wins and Lorraine loses, so no one has to switch. You have no doubt already guessed that eventually (with probability one) all of them will be working in the same field. (After the first week, for example, one field has already disappeared permanently, since as soon as Tom loses there are no more topologists.)

This is an example of a *Markov chain*, in which a system can be in a number of possible states, and at each time step there is a certain probability of moving to each of the other states (or of remaining in the same state). Kemeny and Snell ([**2**]) give an excellent background on Markov chains.

We will break our chain up into *stages*, numbered in reverse order according to how many fields are remaining. Thus, we start in Stage 4, meaning there are four fields left, but after one week we are certain to be in Stage 3. We will study three questions here:

1. *How long do we expect to stay in each stage?* The expected time in Stage 4 (or Stage n in the general case of n starters) is exactly one week, but after that it gets more complicated.
2. *When we first arrive at Stage $t-1$ from Stage t, what is the most likely configuration of the fields?* More precisely, what are the probabilities of arriving at different configurations of the players into $t-1$ teams? For example, with $n = 4$ starters, when we go down from three fields to two, are we more likely to have two teams of two players each, or a team of three and a lone wolf?
3. *How long does the game last?* In other words, what is the expected time until we reach the absorbing state in which everyone is on the same team? Of course, the answer here is just the sum of the answers from Question 1.

Math. Mag. **84** (2011) 3–15. doi:10.4169/math.mag.84.1.003.

We invite you to play with the small cases of $n = 3$, 4, or 5 starters, which are not too hard to work out from first principles. You will find that the answers to Question 3 are 4, 9, and 16 weeks respectively. It might not be obvious that this stunning pattern should continue to hold, but we will prove that with n starters, the expected time is indeed $(n-1)^2$ weeks. (Unfortunately, there appears to be no correspondingly congenial answer for the variance.)

The general answers to Questions 1 and 2 are not so obvious from analyzing small cases. For example, with $n = 5$ starters, the total expected time of 16 weeks breaks down into stages of $e_{54} = 1$, $e_{43} = \frac{5}{3}$, $e_{32} = \frac{10}{3}$, and $e_{21} = 10$ weeks. We will see that these come from binomial coefficients and that the answer to Question 2 comes from multinomial coefficients.

We organize the paper as follows: In the second section, we warm up by solving the case $n = 4$ from scratch, using no sophisticated machinery. Besides resolving the question for New Alarkania State, this will give us an informal preview of some of the notation and theorems coming later. Next, we introduce more formal notation and illustrate it with a larger example, $n = 6$. We then study the vectors of probabilities and discover multinomial coefficients as the answer to Question 2. With the probability vectors in hand, it is relatively quick to study the expected times and answer Questions 1 and 3. In the final section, we present a symmetric approach that answers Question 3 directly without reference to the answers to Questions 1 and 2.

$n = 4$: How long must New Alarkania wait?

In this section we will work out the case of four players from scratch using only basic probability; however, some of the notation and theory for later will become evident as we go along. As mentioned above, we organize the possible configurations into stages according to the number of teams left; thus we proceed in reverse order from Stage 4 (four individuals, [1111]) down to Stage 1 (a single team of four, [4]).

Starting at Stage 4 ([1111]), note that in the first week, one player must lose and join the winner's team. Therefore, the expected time to Stage 3 is exactly $e_{43} = 1$ week. The configuration at Stage 3 is necessarily [211], one team of two players and two individuals.

Now, from [211], the loser can be one of the players on the team of two, in which case the new configuration is still [211]. (If the winner is the other player on the team, then there is no change at all; if the winner is one of the two individuals, then the loser joins that individual, making a new team of two and leaving the loser's former teammate as an individual.) If the loser is one of the two individuals, however, we will go down to Stage 2. The new configuration depends on who the winner is, but we note first that since there is a $\frac{1}{2}$ chance of the loser being one of the two individuals, the expected waiting time is exactly $e_{32} = 2$ weeks.

When we do first get down to Stage 2, what configuration will we land in? We know that the loser in the previous week was one of the two individuals. There is a $\frac{2}{3}$ chance that the winner was a member of the team of two, in which case we land in [31]. There is a $\frac{1}{3}$ chance that the winner was the other individual, landing us in [22]. We thus have an answer for Question 2 at Stage 2: We say $L_2 := \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \end{pmatrix}$ is the *landing vector* at Stage 2, representing the probabilities that when we first arrive in Stage 2, we land in [31] or [22] respectively. (We had landing vectors at the previous stages as well, but because there was only one configuration in each stage, they were simply the trivial vectors $L_4 := (1)$, $L_3 := (1)$.)

Finally, we calculate the expected time e_{21} to go from Stage 2 to Stage 1. Here are the possible outcomes from configuration [31]:

Probability	Outcome	Explanation
$\frac{1}{2}$	Stay at [31].	Winner and loser are both from the team of three.
$\frac{1}{4}$	Move to [22].	Winner is the individual.
$\frac{1}{4}$	Move to [4].	Loser is the individual.

And here are the possiblities from [22]:

Probability	Outcome	Explanation
$\frac{2}{3}$	Move to [31].	Winner and loser are from different teams.
$\frac{1}{3}$	Stay at [22].	Winner and loser are on the same team.
0	Move to [4].	Not possible in one week.

We collect these probabilities in a matrix, denoted A_2, for later:

$$\begin{array}{c} \\ [31] \\ [22] \end{array} \begin{array}{c} \begin{array}{ccc} [31] & [22] & [4] \end{array} \\ \left(\begin{array}{cc|c} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{array}\right) \end{array} = \begin{array}{c} \\ [31] \\ [22] \end{array} \begin{array}{c} \begin{array}{ccc} [31] & [22] & [4] \end{array} \\ \left(\begin{array}{c|c} A_2 & \begin{array}{c} \frac{1}{4} \\ 0 \end{array} \end{array}\right) \end{array}$$

To find the expected time e_{21} to go from Stage 2 to Stage 1, let x_1 be the expected time to go from [31] to [4] and let x_2 be the expected time to go from [22] to [4] (necessarily via [31]). If we start at [31] and let one week go by, there is a $\frac{1}{2}$ chance that we will stay at [31], giving us a new expected time of x_1 plus the one week that just elapsed. There is a $\frac{1}{4}$ chance that we move to [22], giving us a new expected time of x_2 plus one. Finally, there is a $\frac{1}{4}$ chance that we move directly to [4], making the time exactly one week. We summarize this as an equation:

$$x_1 = \frac{1}{4}(x_1 + 1) + \frac{1}{2}(x_2 + 1) + \frac{1}{4}(1) = \frac{1}{4}x_1 + \frac{1}{2}x_2 + 1$$

Starting at [22] and letting one week elapse gives us a similar equation:

$$x_2 = \frac{2}{3}(x_1 + 1) + \frac{1}{3}(x_2 + 1) + 0(1) = \frac{2}{3}x_1 + \frac{1}{3}x_2 + 1$$

Combining these equations gives us a matrix equation that is easy to solve:

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= A_2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ (I - A_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= (I - A_2)^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{11}{2} \\ 7 \end{pmatrix} \end{aligned}$$

Recalling the landing vector of probabilities that we arrive at Stage 2 either in [31] or [22], the expected time to go to Stage 1 is then

$$e_{21} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} \frac{11}{2} \\ 7 \end{pmatrix} = 6 \text{ weeks.}$$

Finally, the total expected time to go from Stage 4 down to Stage 1 is the sum of the expected times at each stage, $e_{43} + e_{32} + e_{21} = 1 + 2 + 6 = 9$ weeks, or $(n-1)^2$ for $n = 4$.

Besides answering Questions 1–3 for New Alarkania, this small example already showcases several features that will be reflected in larger cases later:

- We depended heavily on *linearity of expectation* to break the total expected time into a sum of expected times $e_{t,t-1}$ to go from each Stage t to Stage $t-1$.
- Stage 2 (and for larger cases, almost all stages) consisted of multiple possible configurations, [31] and [22]. We described our arrival at Stage 2 in terms of a *landing vector* of probabilities $L_2 := \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \end{pmatrix}$ that we would first land in each configuration. These landing vectors are the answer to Question 2, but this one small example is not enough to see the general pattern.
- We can compute the expected time to go from Stage t to Stage $t-1$ as

$$e_{t,t-1} = L_t(I - A_t)^{-1}\mathbf{1},$$

  where L_t is the landing vector of probabilities for the configurations in Stage t, A_t is the matrix of internal transition probabilities between the various configurations in Stage t, and $\mathbf{1}$ is a column vector of ones of the appropriate length.
- In this small example, the expected times were all integers, $e_{43} = 1$, $e_{32} = 2$, and $e_{21} = 6$. That won't generalize, but they will follow a most interesting pattern. (We invite you to guess it now, with the reminder that the times for the case $n = 5$ are $e_{54} = 1$, $e_{43} = \frac{5}{3}$, $e_{32} = \frac{10}{3}$, and $e_{21} = 10$, giving a total time of $1 + \frac{5}{3} + \frac{10}{3} + 10 = 16 = (n-1)^2$ weeks.)

Keeping the lessons from $n = 4$ in mind, we now move on to address the general problem.

Notation and examples

Fix a value of n. We will consider the various partitions of n to be the states of the system. We will use both *partition notation*, where we list the parts as $n_1 + n_2 + \cdots + n_k$, which we will abbreviate as $n_1 n_2 \cdots n_k$, and *vector notation*, where we list the number of parts of each size as $(r_1 r_2 \cdots r_k)$, so $\sum i r_i = n$. (When using vector notation, we will always assume that the last entry is nonzero.)

Let $S(n,t)$ be the set of partitions of n into t parts, i.e., the set of all possible configurations at Stage t. Then the set of all partitions of n is $\cup_{t=1}^{n} S(n,t)$. We list the sets $S(n,t)$ in reverse order from $t = n$ to $t = 1$, and we assume that each $S(n,t)$ is given a consistent internal ordering.

For example, let $n = 6$. Then the states in partition notation are

$$\{[111111], [21111], [2211, 3111], [222, 321, 411], [33, 42, 51], [6]\},$$

and, respectively, in vector notation are

$$\{[(6)], [(41)], [(22), (301)], [(03), (111), (2001)],$$
$$[(002), (0101), (10001)], [(000001)]\}.$$

Let P be the probability transition matrix between the various possible states. Then P is block upper bidiagonal, where each diagonal block is A_t, the probability transition matrix from states in Stage t to each other, and each superdiagonal block is $A_{t,t-1}$, the probability transition matrix from states in Stage t to states in Stage $t-1$.

For $n = 6$, using the ordering above, we have the following matrix:

$$P = \left(\begin{array}{c|c|c|c|c|c}
A_6 & A_{65} & & & & \\ \hline
 & A_5 & A_{54} & & & \\ \hline
 & & A_4 & A_{43} & & \\ \hline
 & & & A_3 & A_{32} & \\ \hline
 & & & & A_2 & A_{21} \\ \hline
 & & & & & A_1
\end{array}\right)$$

$$= \frac{1}{30}\left(\begin{array}{c|c|cc|ccc|ccc|c}
0 & 30 & & & & & & & & & \\ \hline
 & 10 & 12 & 8 & & & & & & & \\ \hline
 & & 12 & 8 & 2 & 8 & 0 & & & & \\
 & & 9 & 6 & 0 & 6 & 9 & & & & \\ \hline
 & & & & 6 & 24 & 0 & 0 & 0 & 0 & \\
 & & & & 3 & 16 & 6 & 2 & 3 & 0 & \\
 & & & & 0 & 8 & 12 & 0 & 2 & 8 & \\ \hline
 & & & & & & & 12 & 18 & 0 & 0 \\
 & & & & & & & 8 & 14 & 8 & 0 \\
 & & & & & & & 0 & 5 & 20 & 5 \\ \hline
 & & & & & & & & & & 30
\end{array}\right)$$

For example, the middle rows of A_3 and A_{32} are obtained by noting that of the 30 possible choices for winner and loser from the partition 321 (in vector notation, (111)), 3 lead to the partition 222, 16 to 321, 6 to 411, 2 to 33, 3 to 42, and none to 51 (in vector notation, (03), (111), (2001), (002), (0101), (10001), respectively). We invite you to check the other values.

Probability vectors and multinomial coefficients

We define the *landing vectors* L_t recursively as follows. First, we set $L_n := (1)$ since we must start in Stage n in state (n). Now, for $n \geq t \geq 2$, assume that we start in one of the states in Stage t with probabilities given by the entries of L_t. We then define L_{t-1} to be the row vector whose jth entry is the probability that our first arrival in Stage $t-1$ from Stage t is in the jth state in Stage $t-1$.

Thus, in the example above with $n = 6$, we have $L_6 = (1)$, $L_5 = (1)$, and $L_4 = \left(\frac{3}{5} \quad \frac{2}{5}\right)$, because when we move from Stage 5, necessarily starting at (41) (in vector notation), to Stage 4, we have a $\frac{3}{5}$ chance of arriving in state (22) and a $\frac{2}{5}$ chance of arriving in state (301).

To calculate the L_t's, we define $P_{t,t-1}$ to be a matrix in which each row corresponds to a state in Stage t and each column to a state in Stage $t-1$. Entry (i, j) in $P_{t,t-1}$ is defined to be the probability that, given that we start in state i in Stage t, our first arrival in Stage $t-1$ is in state j. By a similar derivation to the one we used in the example with $n = 4$ above, we have

$$P_{t,t-1} = (I - A_t)^{-1} A_{t,t-1},$$

where I is the identity matrix of appropriate size. (This is also a standard result in the theory of Markov chains; see Theorem 3.3.7 in [**2**].)

We can now compute the L_t's recursively:

$$L_{t-1} = L_t P_{t,t-1} = L_t (I - A_t)^{-1} A_{t,t-1}.$$

For example, with $n = 6$, we have

$$\begin{aligned} L_5 &= L_6 (I - A_6)^{-1} A_{65} = (1)(1)(1) = (1) \\ L_4 &= L_5 (I - A_5)^{-1} A_{54} = (1)\left(\tfrac{3}{2}\right)\begin{pmatrix} \tfrac{2}{5} & \tfrac{4}{15} \end{pmatrix} = \begin{pmatrix} \tfrac{3}{5} & \tfrac{2}{5} \end{pmatrix} \\ L_3 &= L_4 (I - A_4)^{-1} A_{43} \\ &= \begin{pmatrix} \tfrac{3}{5} & \tfrac{2}{5} \end{pmatrix} \begin{pmatrix} 2 & \tfrac{2}{3} \\ \tfrac{3}{4} & \tfrac{3}{2} \end{pmatrix} \begin{pmatrix} \tfrac{1}{15} & \tfrac{4}{15} & 0 \\ 0 & \tfrac{1}{5} & \tfrac{3}{10} \end{pmatrix} \\ &= \begin{pmatrix} \tfrac{1}{10} & \tfrac{3}{5} & \tfrac{3}{10} \end{pmatrix} \end{aligned}$$

and so on.

We define $VS(n,t)$ to be the vector space whose basis is the set of partitions $S(n,t)$ in Stage t. Using vector notation for partitions, for

$$\mathbf{r} = \begin{pmatrix} r_1 & r_2 & \cdots & r_k \end{pmatrix} \in S(n,t),$$

we define the *multinomial coefficient*

$$m_{\mathbf{r}} := \binom{t}{r_1, r_2, \ldots, r_k} = \frac{t!}{r_1! r_2! \cdots r_k!}.$$

(Undergraduates will recall multinomial coefficients from combinatorial exercises about rearranging the letters of words like MISSISSIPPI; see Section 5.4 in [**1**] for details.)

Finally we define the vector $\mathbf{u}_t \in VS(n,t)$ by

$$\mathbf{u}_t := \sum_{\mathbf{r} \in S(n,t)} m_{\mathbf{r}} \mathbf{r}$$

and consider it as a row vector whose entries are the $m_{\mathbf{r}}$'s.

We can add the entries of a vector by multiplying it by $\mathbf{1}$, the column vector of appropriate size whose entries are all ones.

REMARK 1. *The sum of the coefficients of* $\mathbf{u}_t$ *is*

$$\mathbf{u}_t \mathbf{1} = \sum_{\mathbf{r} \in S(n,t)} m_{\mathbf{r}} = \binom{n-1}{t-1}.$$

Proof. One way to list the partitions of n into t parts is to make a line of n pebbles and then insert $t-1$ dividers into the $n-1$ spaces between the pebbles; there are $\binom{n-1}{t-1}$ ways to do this. However, most partitions will be counted multiple times in this list since the parts can appear in any order. In fact, the partition $\mathbf{r} = \begin{pmatrix} r_1 & r_2 & \cdots & r_k \end{pmatrix} \in S(n,t)$ will appear exactly

$$m_{\mathbf{r}} = \binom{t}{r_1, r_2, \ldots, r_k} = \frac{t!}{r_1! r_2! \cdots r_k!}$$

times, giving the desired result. ■

For example, with $n = 6$, we have the following:

$$\begin{aligned}
\mathbf{u}_6 &= 1(6) &&= (1); && \mathbf{u}_6\mathbf{1} = 1\\
\mathbf{u}_5 &= 5(41) &&= (5); && \mathbf{u}_5\mathbf{1} = 5\\
\mathbf{u}_4 &= 6(22) + 4(301) &&= (6 \quad 4); && \mathbf{u}_4\mathbf{1} = 10\\
\mathbf{u}_3 &= 1(03) + 6(111) + 3(2001) &&= (1 \quad 6 \quad 3); && \mathbf{u}_3\mathbf{1} = 10\\
\mathbf{u}_2 &= 1(002) + 2(0101) + 2(10001) &&= (1 \quad 2 \quad 2); && \mathbf{u}_2\mathbf{1} = 5\\
\mathbf{u}_1 &= 1(000001) &&= (1); && \mathbf{u}_1\mathbf{1} = 1
\end{aligned}$$

Note that we have $\mathbf{u}_t\mathbf{1} = \binom{5}{t-1}$, as predicted by Remark 1. Note also that the normalized version of $\mathbf{u}_3$ is

$$\frac{1}{\mathbf{u}_3\mathbf{1}}\mathbf{u}_3 = \frac{1}{10}\begin{pmatrix}1 & 6 & 3\end{pmatrix} = \begin{pmatrix}\frac{1}{10} & \frac{3}{5} & \frac{3}{10}\end{pmatrix} = L_3,$$

and the same is true for the other $\mathbf{u}_t$'s and L_t's. This elegant pattern for the landing vectors is the answer to Question 2, but we need several more theorems to justify it. The first two state that the $\mathbf{u}_t$ are eigenvectors for the probability transition matrices A_t, and they are also "chain eigenvectors" in the sense that $\mathbf{u}_t A_{t,t-1}$ is a scalar multiple of $\mathbf{u}_{t-1}$:

THEOREM 2.

$$\mathbf{u}_t A_t = \mathbf{u}_t d_t, \textit{ where } d_t := \frac{(n-t)(n+t-1)}{n(n-1)}.$$

We will discuss the proof of Theorem 2 below.

COROLLARY 3. *$\mathbf{u}_t$ is a left eigenvector for the matrix $(I - A_t)^{-1}$ with eigenvalue $\frac{1}{1-d_t}$.*

THEOREM 4.

$$\mathbf{u}_t A_{t,t-1} = \mathbf{u}_{t-1} h_t, \textit{ where } h_t := \frac{t(n-t+1)}{n(n-1)}.$$

Surprisingly, in our work later, we will only use the fact that $\mathbf{u}_t A_{t,t-1}$ is a multiple of $\mathbf{u}_{t-1}$; the actual value of h_t is immaterial. We will explain this after Theorem 8 below.

The proof of Theorem 2 (respectively, Theorem 4) depends on some careful combinatorial bookkeeping. We will suppress the computational details of the proofs, partly because of the tedium involved and partly because we have an independent way to answer Question 3 that we will present in full detail later. Instead, we will just give a sketch here and then illustrate with a numerical example.

The main idea of both proofs is to track which states $\mathbf{s} \in S(n, t)$ (respectively, $\mathbf{s} \in S(n, t-1)$) can be reached directly from which states $\mathbf{r} \in S(n, t)$, which we denote by $\mathbf{r} \to \mathbf{s}$. For $\mathbf{r} \to \mathbf{s}$, we define $\delta(\mathbf{r}, \mathbf{s})$ to be the number of possible winner-loser pairs in state $\mathbf{r}$ that will take us to state $\mathbf{s}$, that is, the numerator of the corresponding entry in A_t (respectively $A_{t,t-1}$), where the denominator is $n(n-1)$.

The key step in the proof of Theorem 2 is then to switch from summing over $\mathbf{r}$ to summing over $\mathbf{s}$:

$$
\begin{aligned}
\mathbf{u}_t A_t &= \sum_{\mathbf{r}\in S(n,t)} m_{\mathbf{r}} \mathbf{r} A_t && \text{by definition of } \mathbf{u}_t \\
&= \frac{1}{n(n-1)} \sum_{\mathbf{r}\in S(n,t)} \sum_{\{\mathbf{s}:\mathbf{r}\to\mathbf{s}\}} m_{\mathbf{r}} \delta(\mathbf{r},\mathbf{s})\mathbf{s} && \text{by the action of } A_t \\
&= \frac{1}{n(n-1)} \sum_{\mathbf{s}\in S(n,t)} \sum_{\{\mathbf{r}:\mathbf{r}\to\mathbf{s}\}} m_{\mathbf{r}} \delta(\mathbf{r},\mathbf{s})\mathbf{s} && \text{switching the summation} \\
&= \frac{1}{n(n-1)} \sum_{\mathbf{s}\in S(n,t)} (n-t)(n+t-1) m_{\mathbf{s}} \mathbf{s} && \text{(see below)} \\
&= \mathbf{u}_t \frac{(n-t)(n+t-1)}{n(n-1)} && \text{by definition of } \mathbf{u}_t
\end{aligned}
$$

The work is in justifying the second to last equality above that

$$\sum_{\{\mathbf{r}:\mathbf{r}\to\mathbf{s}\}} m_{\mathbf{r}} \delta(\mathbf{r},\mathbf{s}) = t(n-t+1) m_{\mathbf{s}}.$$

This requires several pages of unenlightening calculation. The proof of Theorem 4 is similar, and similarly tedious. We have spared you the full details, and instead we will illustrate with a larger concrete example. Let $n = 10$ and $t = 4$; then in partition notation we have

$$
\begin{aligned}
S(10,4) &= \{3331, 3322, 4321, 4411, 4222, 5311, 5221, 6211, 7111\} \\
S(10,3) &= \{433, 442, 541, 532, 631, 622, 721, 811\}
\end{aligned}
$$

and in vector notation we have

$$
\begin{aligned}
S(10,4) = \{&(103), (022), (1111), (2002), (0301), (20101), (12001), \\
&(210001), (3000001)\} \\
S(10,3) = \{&(0021), (0102), (10011), (01101), (101001), (020001), \\
&(1100001), (20000001)\}.
\end{aligned}
$$

Then

$$
\begin{aligned}
\mathbf{u}_4 &= \begin{pmatrix} 4 & 6 & 24 & 6 & 4 & 12 & 12 & 12 & 4 \end{pmatrix}, \\
\mathbf{u}_3 &= \begin{pmatrix} 3 & 3 & 6 & 6 & 6 & 3 & 6 & 3 \end{pmatrix}
\end{aligned}
$$

with $\mathbf{u}_4\mathbf{1} = 84 = \binom{9}{3}$ and $\mathbf{u}_3\mathbf{1} = 36 = \binom{9}{2}$, as predicted by Remark 1.

The corresponding blocks of the transition matrix are

$$
\begin{pmatrix} A_4 \mid A_{43} \end{pmatrix} = \frac{1}{90}
\left(\begin{array}{ccccccccc|cccccccc}
18 & 9 & 54 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
8 & 40 & 24 & 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
8 & 4 & 40 & 6 & 3 & 8 & 12 & 0 & 0 & 2 & 3 & 0 & 4 & 0 & 0 & 0 & 0 \\
0 & 0 & 16 & 24 & 0 & 32 & 0 & 0 & 0 & 0 & 2 & 16 & 0 & 0 & 0 & 0 & 0 \\
0 & 24 & 24 & 0 & 18 & 0 & 24 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 10 & 15 & 0 & 26 & 6 & 15 & 0 & 0 & 0 & 6 & 2 & 10 & 0 & 0 & 0 \\
0 & 0 & 20 & 0 & 5 & 8 & 28 & 20 & 0 & 0 & 0 & 0 & 4 & 0 & 5 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 12 & 12 & 36 & 12 & 0 & 0 & 0 & 0 & 4 & 2 & 12 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 21 & 42 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 21
\end{array}\right).
$$

Note that $\mathbf{u}_4 A_4 = \mathbf{u}_4 \frac{13}{15} = \mathbf{u}_4 d_4$ and $\mathbf{u}_4 A_{43} = \mathbf{u}_3 \frac{14}{45} = \mathbf{u}_3 h_4$, as predicted by Theorem 2 and Theorem 4.

We can now justify our answer to Question 2:

THEOREM 5. *For all t, the landing vector L_t is the normalized $\mathbf{u}_t$, that is,*

$$L_t = \frac{1}{\mathbf{u}_t\mathbf{1}}\mathbf{u}_t.$$

Proof. First, we note that $L_n = (1) = \mathbf{u}_n$. Proceeding downwards by induction, we assume the theorem for L_t and show it for L_{t-1}:

$$\begin{aligned}
L_{t-1} &= L_t P_{t,t-1} && \text{by construction of } L_{t-1}\\
&= L_t(I - A_t)^{-1}A_{t,t-1} && \text{by construction of } P_{t,t-1}\\
&= \frac{1}{\mathbf{u}_t\mathbf{1}}\mathbf{u}_t(I - A_t)^{-1}A_{t,t-1} && \text{by the induction hypothesis}\\
&= \frac{1}{\mathbf{u}_t\mathbf{1}}\frac{1}{1-d_t}\mathbf{u}_t A_{t,t-1} && \text{by Corollary 3}\\
&= \frac{1}{\mathbf{u}_t\mathbf{1}}\frac{1}{1-d_t}\mathbf{u}_{t-1}h_t && \text{by Theorem 4}
\end{aligned}$$

This shows that L_{t-1} is a scalar multiple of $\mathbf{u}_{t-1}$. But since we know that L_{t-1} is a probability vector, i.e., that its entries sum to one, we must have that

$$L_{t-1} = \frac{1}{\mathbf{u}_{t-1}\mathbf{1}}\mathbf{u}_{t-1},$$

as desired. ■

REMARK 6. *The proof of Theorem 5 gives an alternate way to find $\mathbf{u}_t\mathbf{1}$.*

Proof. We can find a relationship between $\mathbf{u}_{t-1}\mathbf{1}$ and $\mathbf{u}_t\mathbf{1}$:

$$\begin{aligned}
L_{t-1}\mathbf{1} &= \frac{h_t}{1-d_t}\frac{1}{\mathbf{u}_t\mathbf{1}}\mathbf{u}_{t-1}\mathbf{1} && \text{from the proof above}\\
1 &= \frac{h_t}{1-d_t}\frac{1}{\mathbf{u}_t\mathbf{1}}\mathbf{u}_{t-1}\mathbf{1} && \text{since } L_{t-1} \text{ is a probability vector}\\
\mathbf{u}_{t-1}\mathbf{1} &= \frac{1-d_t}{h_t}\mathbf{u}_t\mathbf{1} && \text{by cross multiplication}\\
&= \frac{t-1}{n-t+1}\mathbf{u}_t\mathbf{1} && \text{by definition of } d_t \text{ and } h_t
\end{aligned}$$

This gives us the recursive system

$$\begin{aligned}
\mathbf{u}_n\mathbf{1} &= 1\\
\mathbf{u}_{n-1}\mathbf{1} &= \frac{n-1}{1}\mathbf{u}_n\mathbf{1} = \frac{n-1}{1}\\
\mathbf{u}_{n-2}\mathbf{1} &= \frac{n-2}{2}\mathbf{u}_{n-1}\mathbf{1} = \frac{n-2}{2}\frac{n-1}{1}\\
&\vdots\\
\mathbf{u}_t\mathbf{1} &= \frac{t}{n-t}\cdots\frac{n-2}{2}\frac{n-1}{1} = \binom{n-1}{t-1},
\end{aligned}$$

confirming our result from Remark 1. ■

Expected times

We are now ready to answer Questions 1 and 3. Recall that L_t is the row vector whose jth entry is the probability that our first arrival in Stage $t-1$ from Stage t is in the jth state in Stage $t-1$. We define $e_{t,t-1}$ to be the expected time from our first arrival in Stage t to our first arrival in Stage $t-1$. We have an immediate answer for Question 1.

THEOREM 7.

$$e_{t,t-1} = \frac{n(n-1)}{t(t-1)} = \frac{\binom{n}{2}}{\binom{t}{2}}$$

Proof. When we worked out the case for $n = 4$ we derived a formula for $e_{t,t-1}$ that clearly generalizes to larger cases. (This is a standard result in the theory of Markov chains; see Theorem 3.3.5 in [**2**].) We proceed from that formula:

$$\begin{aligned}
e_{t,t-1} &= L_t(I - A_t)^{-1}\mathbf{1} \\
&= L_t \frac{1}{1-d_t}\mathbf{1} && \text{by Corollary 3 and Theorem 5} \\
&= \frac{1}{1-d_t} L_t \mathbf{1} && \text{since } \frac{1}{1-d_t} \text{ is a scalar} \\
&= \frac{1}{1-d_t} && \text{since } L_t \text{ is a probability vector} \\
&= \frac{n(n-1)}{t(t-1)} && \text{by definition of } d_t
\end{aligned}$$ ■

We now just add the times at each stage to answer Question 3:

THEOREM 8. *The expected time to the final state is* $(n-1)^2$.

Proof. We use a partial fraction expansion:

$$\begin{aligned}
\sum_{t=2}^{n} e_{t,t-1} &= \sum_{t=2}^{n} \frac{n(n-1)}{t(t-1)} && \text{by Theorem 7} \\
&= n(n-1)\sum_{t=2}^{n}\left(\frac{1}{t-1} - \frac{1}{t}\right), && \text{a telescoping series} \\
&= n(n-1)\left(1 - \frac{1}{n}\right) \\
&= (n-1)^2
\end{aligned}$$ ■

One slightly surprising element of the proofs above is that we never used the formula for the "chain eigenvalue" h_t from Theorem 4. (We did use the value of h_t in the proof of Remark 6, but Remark 6 was not used to prove anything else.) This is less surprising when we realize that the value of h_t can be derived from the value of d_t by the following method, which is independent of the formula in Theorem 4. Note that the row vector $L_t \begin{pmatrix} A_t & A_{t,t-1} \end{pmatrix}$ gives the complete set of probabilities of landing in the various states in Stage t and Stage $t-1$ one step after landing in Stage t. Accordingly, the entries in this row vector add to one. But we can calculate this vector:

$$
\begin{aligned}
L_t A_t &= L_t d_t && \text{by Theorems 2 and 5} \\
L_t A_{t,t-1} &= \mathbf{u}_t \frac{1}{\mathbf{u}_t \mathbf{1}} A_{t,t-1} && \text{by Theorem 5} \\
&= \mathbf{u}_{t-1} \frac{h_t}{\mathbf{u}_t \mathbf{1}} && \text{by Theorem 4} \\
&= L_{t-1} \frac{(\mathbf{u}_{t-1}\mathbf{1}) h_t}{\mathbf{u}_t \mathbf{1}} && \text{by Theorem 5}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
L_t \begin{pmatrix} A_t & A_{t,t-1} \end{pmatrix} \mathbf{1} &= 1 && \text{by the discussion above} \\
\begin{pmatrix} L_t d_t & L_{t-1} \frac{(\mathbf{u}_{t-1}\mathbf{1}) h_t}{\mathbf{u}_t \mathbf{1}} \end{pmatrix} \mathbf{1} &= 1 && \text{by the calculations immediately above} \\
d_t + \frac{(\mathbf{u}_{t-1}\mathbf{1}) h_t}{\mathbf{u}_t \mathbf{1}} &= 1 && \text{since } L_t \text{ and } L_{t-1} \text{ are probability vectors} \\
d_t + \frac{(t-1) h_t}{n-t+1} &= 1 && \text{by Remark 1.}
\end{aligned}
$$

Thus, d_t and h_t are dependent on each other, and if we use a particular value of one, then we are also implicitly using the corresponding value of the other. And note that the value of d_t did indeed play a key role in the proof of Theorem 7 above.

A symmetric approach

Although we think the answers to Questions 1 and 2 are interesting in their own right, we can derive the answer to Question 3 independently without going through the calculations above. In particular, this method does not rely on the proofs of Theorems 2 and 4.

We start with n players, each of whom initially represents a different field. We arbitrarily choose one field to focus on, say, statistics. At any point in the game, we define a set of random variables $x_0, \ldots, x_n$, where x_i represents the number of *future wins by statisticians*, given that there are i statisticians currently remaining. (Note that it does not matter what the configuration of the other $n-i$ players into teams is.) We have easy boundary values: $x_0 = 0$, since if statistics has been wiped out as a field, then there can be no future converts to statistics; and $x_n = 0$, since if everyone is now a statistician then the game is over.

We now set up a system of equations for the other x_i, $1 \le i \le n-1$. In each round, there are $n(n-1)$ choices for the winner and loser. With i statisticians currently, there are four possibilities for how the winner and loser can be arranged with respect to the statisticians:

1. Both winner and loser are statisticians. There are $i(i-1)$ ways this can happen. The number of wins by statisticians has gone up by one, and the new expectation at the following round is again x_i, since we again have i statisticians.
2. Only the winner is a statistician. There are $i(n-i)$ ways this can happen. The number of wins by statisticians has gone up by one, and the new expectation at the following round is x_{i+1}, since we then have $i+1$ statisticians.

3. Only the loser is a statistician. There are $i(n-i)$ ways this can happen. The number of wins by statisticians is unchanged, and the new expectation at the following round is x_{i-1}.
4. Neither the winner nor the loser is a statistician. There are $(n-i)(n-i-1)$ ways this can happen. The number of wins by statisticians is unchanged, and the new expectation at the following round is again x_i.

This gives us the following equation:

$$x_i = \frac{i(i-1)}{n(n-1)}(1+x_i) + \frac{i(n-i)}{n(n-1)}(1+x_{i+1}) + \frac{i(n-i)}{n(n-1)}x_{i-1} + \frac{(n-i)(n-i-1)}{n(n-1)}x_i$$

Mercifully, this simplifies rather dramatically:

$$2x_i - (x_{i-1} + x_{i+1}) = \frac{n-1}{n-i}$$

This gives us a linear system for the x_i's:

$$\begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{n-1}{n-2} \\ \frac{n-1}{n-3} \\ \vdots \\ n-1 \end{pmatrix}$$

We denote the $(n-1)\times(n-1)$ matrix on the left by M_n. It is an amusing exercise to compute M_n^{-1}; for example, with $n=6$ we have

$$M_6 = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}, \quad M_6^{-1} = \frac{1}{6}\begin{pmatrix} 5 & 4 & 3 & 2 & 1 \\ 4 & 8 & 6 & 4 & 2 \\ 3 & 6 & 9 & 6 & 3 \\ 2 & 4 & 6 & 8 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}.$$

The pattern in the right-hand matrix is that the (i,j)-entry is $i(n-j)$ for entries above the main diagonal and $j(n-i)$ for entries below. In other words,

$$\left(M_n^{-1}\right)_{i,j} = \frac{1}{n}\min\{i,j\}\,[n-\max\{i,j\}]\,.$$

To answer Question 3, we need to know the expected number of future wins by statisticians at the very start of the game. We start with one statistician, so we solve our system for x_1 using the first row of M_n^{-1}:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{pmatrix} = M_n^{-1} \begin{pmatrix} 1 \\ \frac{n-1}{n-2} \\ \vdots \\ n-1 \end{pmatrix}$$

$$\begin{aligned} x_1 &= \frac{1}{n}\begin{pmatrix} n-1 & n-2 & \cdots & 1 \end{pmatrix}\begin{pmatrix} 1 & \frac{n-1}{n-2} & \cdots & n-1 \end{pmatrix}^T \\ &= \frac{1}{n}\left[(n-1)+(n-1)+\cdots+(n-1)\right] \\ &= \frac{(n-1)^2}{n} \end{aligned}$$

We have just computed the expected number of total wins by *statisticians*. By symmetry, every other field expects the same number of wins, so the total number of rounds of the game (again, exploiting linearity of expectation) is $n\frac{(n-1)^2}{n} = (n-1)^2$. This confirms our answer to Question 3 from the small games and the derivation in the previous section.

Finally, we address the temptation to hope that a Markov chain with such a nice expectation might also have an interesting variance. Following Theorem 3.3.5 in [**2**], we can compute the variance of the time to absorption via the matrix $N := (I - A)^{-1}$, where A is the submatrix of P obtained by deleting the final row and column, which correspond to the absorbing state. We then define the column vector $\tau := N\mathbf{1}$ (the expected time to absorption from each state), and let τ_{sq} be the column vector whose entries are the squares of those in τ. Then [**2**] tells us that the variance of the time to absorption is the first entry of the vector

$$\tau_2 := (2N - I)\tau - \tau_{\text{sq}}.$$

For $n = 2, 3, 4$, the variances turn out to be 0, 6, and 32, raising the hope that an interesting sequence of integers might ensue. Sadly, for $n = 5$ and $n = 6$, the variances are $\frac{890}{9}$ and $\frac{469}{2}$, respectively. We challenge you to discover, prove, and interpret the general pattern!

Acknowledgment We thank John Brevik for suggesting this problem and Kent Merryfield and Peter Ralph for useful conversations.

REFERENCES

1. Richard A. Brualdi, *Introductory Combinatorics*, 5th ed., Prentice Hall, Upper Saddle River, NJ, 2009.
2. John G. Kemeny and J. Laurie Snell, *Finite Markov Chains*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1976.

Summary Consider a system of n players in which each initially starts on a different team. At each time step, we select an individual winner and an individual loser randomly and the loser joins the winner's team. The resulting Markov chain and stochastic matrix clearly have one absorbing state, in which all players are on the same team, but the combinatorics along the way are surprisingly elegant. The expected number of time steps until each team is eliminated is a ratio of binomial coefficients. When a team is eliminated, the probabilities that the players are configured in various partitions of n into t teams are given by multinomial coefficients. The expected value of the time to absorbtion is $(n-1)^2$ steps. The results depend on elementary combinatorics, linear algebra, and the theory of Markov chains.

ROBERT MENA joined the faculty at Long Beach State in 1988 after 15 years in the faculty at the University of Wyoming, and graduate school at the University of Houston. He has also spent time at Caltech and Ohio State. He is an enthusiastic solver of quote and other sorts of acrostics as well as ken ken puzzles. Even after 40 years of teaching, he still purports to enjoy the company of his students and the thrill of teaching mathematics.

WILL MURRAY did graduate work in algebra and is also interested in probability, combinatorics, and analysis. He has taught at Berkeley, Long Beach State, the Royal University of Phnom Penh, and the University of Médéa in Algeria. An avid juggler and traveler, he has performed on five continents and written and lectured about the mathematics of juggling. At home, he enjoys playing with his pets, one of whom starred on *America's Got Talent*.

Crosscut Convex Quadrilaterals

RICK MABRY
Louisiana State University in Shreveport
Shreveport, LA 71115
Richard.Mabry@LSUS.edu

In all that follows, $\mathcal{A}$ is an arbitrary convex plane quadrilateral; let's call such a polygon a *quad*. We let the vertices of $\mathcal{A}$ be given cyclically by the four-tuple (A_0, A_1, A_2, A_3) and name a midpoint M_i on each segment (A_i, A_{i+1}), taking all subscripts modulo 4, as in FIGURE 1. We then *crosscut* quad $\mathcal{A}$ by drawing *medians* (A_i, M_{i+2}). These medians intersect one another at the vertices of a new "inner quad" $\mathcal{B}$ in the interior of the "outer quad" $\mathcal{A}$.

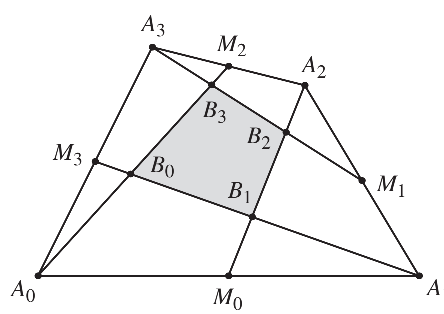


Figure 1 Crosscut quad $\mathcal{A}$ with shaded inner quad $\mathcal{B}$

Our results are inspired by the well known, pretty result ([**1**, p. 49], [**8**, p. 22]) that if $\mathcal{A}$ is a square, then

$$|\mathcal{A}| = 5\,|\mathcal{B}|\,, \tag{1}$$

where $|\cdot|$ denotes area. (See FIGURE 2. In this case it is clear that $\mathcal{B}$ is also a square.)

It follows from familiar facts about shear transformations that (1) remains true when $\mathcal{A}$ is a parallelogram. That (1) does not hold in general is easily seen by letting one

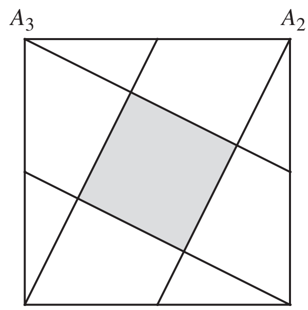


Figure 2 The classic crosscut square

Math. Mag. **84** (2011) 16–25. doi:10.4169/math.mag.84.1.016. © Mathematical Association of America

vertex of $\mathcal{A}$ approach another, in which case the shape of $\mathcal{A}$ approaches that of a triangle. In such a limiting case we get

$$|\mathcal{A}| = 6\,|\mathcal{B}|\,, \tag{2}$$

a fact that we leave as an easy exercise for the reader. (See FIGURE 3, where our four medians have coalesced into two medians of a triangle, these meeting at the triangle's centroid.) We will continue to refer to such a figure as a quad, albeit a *degenerate* one, and this will be the only type of degeneracy we need to consider—two vertices of a quad merging to form a nondegenerate triangle. Otherwise, a (nondegenerate) quad shall be convex with four interior angles strictly between zero and 180 degrees.

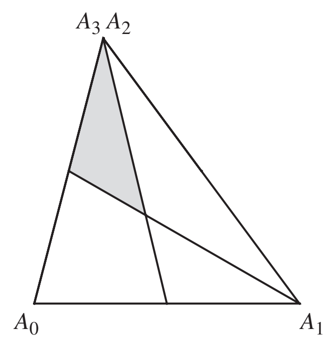


Figure 3 A degenerate case of two coincident vertices

It will be shown presently that the general case lies between (1) and (2). Actually, we prove a bit more:

THEOREM 1. *For an arbitrary outer quad $\mathcal{A}$, the following properties hold.*

(a) *The inner quadrilateral $\mathcal{B}$ is a quad and*

$$5\,|\mathcal{B}| \le |\mathcal{A}| \le 6\,|\mathcal{B}|\,. \tag{3}$$

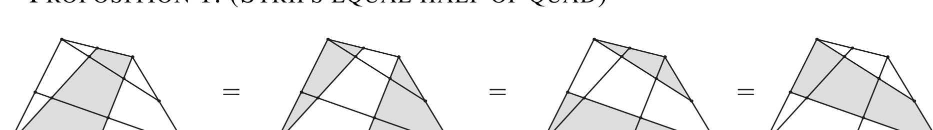
(b) $|\mathcal{A}| = 5\,|\mathcal{B}|$ *if and only if $\mathcal{B}$ is a trapezoid.*

(c) $|\mathcal{A}| = 6\,|\mathcal{B}|$ *if and only if $\mathcal{A}$ is a degenerate quad with two coincident vertices.*

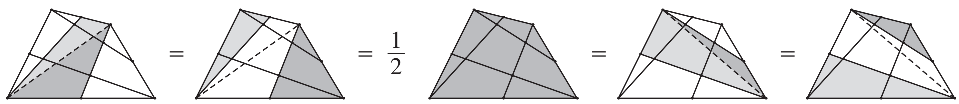
Theorem 1 has gathered dust since 2000 (or earlier—see the Epilogue) while the author occasionally tried, to no avail, to find a "Proof Without Words" (PWW) of (3) or some other *visual proof*, as is done in [**1**] and [**8**] for the case of a square. Later in this note a fairly visual proof will be given that $\mathcal{B}$ being a trapezoid implies that $|\mathcal{A}| = 5\,|\mathcal{B}|$, but that is far from the entire theorem. In part, the purpose of this note is to open this challenge to a wider audience.

In the meantime, we do have a few other visual propositions to offer. In the following three results and their proofs (WW), the values being added, subtracted, and equated are the areas of the shaded regions of an arbitrary (fixed) quad $\mathcal{A}$.

PROPOSITION 1. (STRIPS EQUAL HALF OF QUAD)

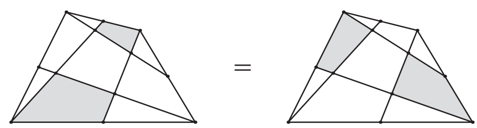


PWW of Proposition 1.

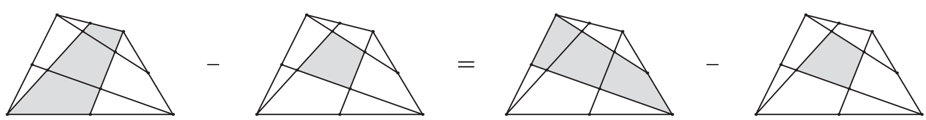


■

PROPOSITION 2. (PAIRS OF OPPOSING FLAPS ARE EQUAL)

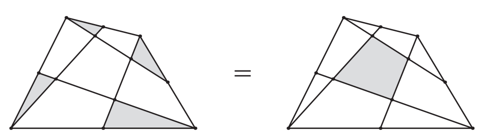


PWW of Proposition 2.

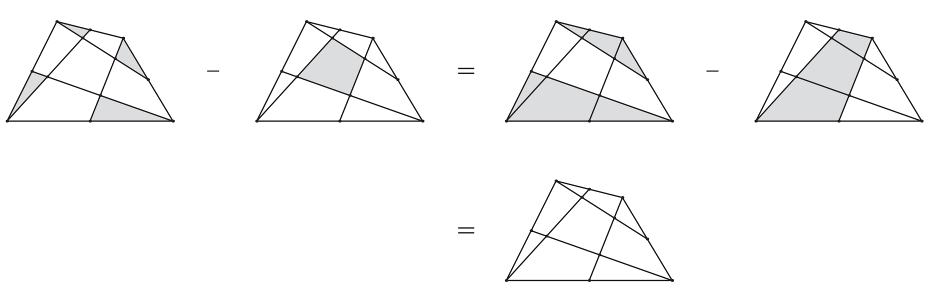


■

PROPOSITION 3. (CORNERS EQUAL INNER QUAD)



PWW of Proposition 3.



■

The greatest visual appeal in this note might be concentrated in Proposition 3. It is a challenge (left to the readers and unfulfilled by the author) to prove some of the other facts in more visual ways, especially Theorem 1. Meanwhile we apply some easy vector-based methods that have some sneaky appeal of their own. Complex variables can be used to the same effect.

Notation, convention, and basic calculation First let's set some notation. Given a sequence $P_0, P_1, \ldots, P_n$ of points in $\mathbb{R}^2$, we let $(P_0, P_1, \ldots, P_n)$ denote either the ordered tuple of points or the polygon formed by taking the points in order. We shall assume that when $(P_0, P_1, \ldots, P_n)$ is given, *the sequence is oriented positively* (counterclockwise) in the plane, with this one exception: If the tuple is a pair, then we consider it *directed* so that we may use it as a vector as well as a line segment, the context making clear which. So, for instance, we write $\mathcal{A} = (A_0, A_1, A_2, A_3)$, and likewise we take $\mathcal{B} = (B_0, B_1, B_2, B_3)$ for our inner quad, where $B_i = (A_i, M_{i+2}) \cap (A_{i+1}, M_{i+3})$. We let $|(P_0, P_1, \ldots, P_n)|$ denote the area of the polygon $(P_0, P_1, \ldots, P_n)$, but $|(P, Q)|$

will denote the length of a segment or vector (P, Q). Likewise, $|\mathbf{w}|$ will denote the length of any vector $\mathbf{w}$. As usual, a point P is identified with the vector (O, P) via the usual canonical identification once an origin O is selected.

Areas are calculated using the magnitude of the cross product. It will be convenient to abuse notation slightly and, for $\mathbf{x} = \langle x_1, x_2\rangle$ and $\mathbf{y} = \langle y_1, y_2\rangle$, set $\mathbf{x} \times \mathbf{y} = x_1 y_2 - x_2 y_1$.

The area of any triangle is then $|(\mathbf{x}, \mathbf{y}, \mathbf{z})| = \frac{1}{2}(\mathbf{y} - \mathbf{x}) \times (\mathbf{z} - \mathbf{x})$, which is positive because of our orientation convention mentioned earlier.

For arbitrary P_0, P_1, P_2, P_3, the polygon formed from the sequence is a nondegenerate convex quadrilateral (and therefore also *simple*, that is, with no self-crossings) if and only if $(P_i, P_{i+1}) \times (P_{i+1}, P_{i+2})$ has constant positive or constant negative sign for each i (mod 4). (An analogous statement cannot be made for convex n-gons with $n \geq 5$—consider the pentagram, or, if the fewest crossings are desired, have a look at a “foxagon” like $(\langle 0, 0\rangle, \langle 3, 9\rangle, \langle -1, 5\rangle, \langle 1, 5\rangle, \langle -3, 9\rangle)$.) We will appeal to cross products to verify the convexity of a certain octagon that arises in our figures.

Diagonals rule Any two fixed, independent vectors $\mathbf{u}$ and $\mathbf{v}$ in the plane correspond to the diagonals of infinitely many different quadrilaterals, as is illustrated in FIGURE 4, where the diagonals of the quadrilaterals shown generate identical vectors. Clearly, such quadrilaterals need not be quads—shown in that figure are a nonconvex and a nonsimple quadrilateral (neither are quads) and two quads, one being a parallelogram. To ensure convexity, it is necessary and sufficient that the diagonals intersect each other. We accomplish this in the following way. Let the origin O be the intersection of the diagonals, and for scalars s and t, set

$$(A_0, A_1, A_2, A_3) = ((1 - s)\mathbf{u}, (1 - t)\mathbf{v}, -s\,\mathbf{u}, -t\,\mathbf{v}), \tag{4}$$

as in FIGURE 5. When s and t in $[0, 1]$ the quadrilateral will be convex; it will be a quad when s and t are in $(0, 1)$. The diagonals (A_0, A_2) and (A_1, A_3) have lengths $|\mathbf{u}|$ and $|\mathbf{v}|$, respectively.

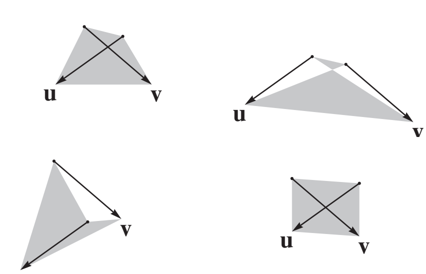


Figure 4 Two linearly independent vectors **u** and **v** form (nonuniquely) the diagonals of a quadrilateral

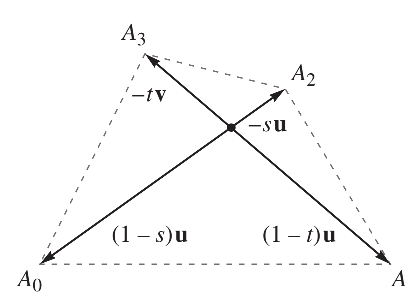


Figure 5 Splitting **u**, **v** with s, t to specify a unique quad

It is a simple matter to compute the area of any simple polygon by triangulation (partitioning the polygon into triangles). For a simple, positively oriented quadrilateral $\mathcal{Q} = (Q_0, Q_1, Q_2, Q_3)$ one may also use the familiar fact that $|\mathcal{Q}| = \frac{1}{2}(Q_0, Q_2) \times (Q_1, Q_3)$. It is then clear that in the case of our main quad, $|\mathcal{A}| = \frac{1}{2}\mathbf{u} \times \mathbf{v}$.

We exploit the fact that all of the vertices of the polygons that appear in the context of crosscut quads are linear combinations of $\mathbf{u}$ and $\mathbf{v}$, with rational functions of s, t

as coefficients. Therefore, since $\mathbf{u} \times \mathbf{u} = \mathbf{v} \times \mathbf{v} = 0$, all of our polygonal areas take the form $F(s, t)\, \mathbf{u} \times \mathbf{v}$ for some rational function F of s, t. All our results derive from $F(s, t)$, as we could scale our figures to have $\mathbf{u} \times \mathbf{v} = 1$.

Proof of Theorem 1. First note that $\mathcal{B}$ is indeed convex, being the intersection of two (clearly) convex sets. To compute $|\mathcal{B}|$, we first find expressions for the points B_i. First, we have that $B_0 = (A_0, M_2) \cap (A_1, M_3)$, where $M_i = \frac{1}{2}(A_i + A_{i+1})$. There are then scalars $q, r \in (0, 1)$ for which

$$B_0 = A_0 + q(M_2 - A_0) = A_1 + r(M_3 - A_1).$$

This implies that

$$\frac{q}{2}\big((A_2 - A_0) + (A_3 - A_0)\big) = A_1 - A_0 + \frac{r}{2}\big((A_3 - A_1) + (A_0 - A_1)\big),$$

hence, by (4),

$$\frac{q}{2}\big(-\mathbf{u} - t\,\mathbf{v} - (1 - s)\mathbf{u}\big) = (1 - t)\mathbf{v} - (1 - s)\mathbf{u} + \frac{r}{2}\big(-\mathbf{v} + (1 - s)\mathbf{u} - (1 - t)\mathbf{v}\big).$$

The linear independence of $\mathbf{u}$ and $\mathbf{v}$ allows us to separately equate their coefficients to solve for q and r, obtaining

$$q = 2(1 - s)/(4 - 2s - t) \quad \text{and} \quad r = 2(2 - s - t)/(4 - 2s - t).$$

(The most diligent of readers will pause to verify that both q and r lie in (0, 1) when s and t do.) We can use q (say) to calculate

$$B_0 = A_0 + q(M_2 - A_0) = \frac{(1 - s)(2 - s - t)}{4 - 2s - t}\,\mathbf{u} - \frac{(1 - s)t}{4 - 2s - t}\,\mathbf{v}.$$

Once B_0 is obtained, the symmetry of our construction in FIGURE 5 can be used to compute the remaining B_i—our convention for indexing points ensures that $P_{i+1} = \langle -y(t, 1 - s),\ x(t, 1 - s)\rangle$ when $P_i = \langle x(s, t),\ y(s, t)\rangle$ is given.

We can now calculate the area of $\mathcal{B}$. We could use the formula for the area of a quadrilateral already mentioned, but instead we'll employ the result of Proposition 3, as we'll make use of the result later. Denoting by $\mathcal{C}_i$ the corner triangle (A_i, B_i, M_{i-1}), we easily compute

$$|\mathcal{C}_0| = |(A_0, B_0, M_3)| = \frac{(1 - s)t}{4(4 - 2s - t)}(\mathbf{u} \times \mathbf{v}). \tag{5}$$

Again, symmetry can be used to compute the remaining three corner areas from the first; generally, if $|\mathcal{R}_i| = g(s, t)$, then $|\mathcal{R}_{i+1}| = g(t, 1 - s)$. (The fact that $(-\mathbf{v}) \times \mathbf{u} = \mathbf{u} \times \mathbf{v}$ is used for this.)

Now we break out the algebra software (if we haven't already) and find, using $|\mathcal{B}| = |\mathcal{C}_0| + |\mathcal{C}_1| + |\mathcal{C}_2| + |\mathcal{C}_3|$, that

$$6\,|\mathcal{B}| - |\mathcal{A}| = \frac{5\left(2 - 3s + s^2 + 4s\,t - 2t^2\right)\left(4s + t - 2s^2 - 4s\,t + t^2\right)}{2(2 - s + 2t)(3 + s - 2t)(4 - 2s - t)(1 + 2s + t)}(\mathbf{u} \times \mathbf{v}). \tag{6}$$

It is now a routine exercise to show that for $s, t \in (0, 1)$, each of the factors above is positive. This shows that $6\,|\mathcal{B}| \ge |\mathcal{A}|$. Similarly, we can compute

$$|\mathcal{A}| - 5\,|\mathcal{B}| = \frac{(1 - 3s + t)^2(2 - s - 3t)^2}{2(2 - s + 2t)(3 + s - 2t)(4 - 2s - t)(1 + 2s + t)}(\mathbf{u} \times \mathbf{v}), \tag{7}$$

which is clearly nonnegative on $[0, 1] \times [0, 1]$. Part (a) of the theorem is now established.

Part (b) of the theorem is fairly easy in view of the fact that (7) shows that $|\mathcal{A}| - 5\,|\mathcal{B}| = 0$ if and only if $1 - 3s + t = 0$ or $2 - s - 3t = 0$. Consider the easily verified fact that

$$(B_0, B_1) \times (B_0, B_3) \\ = -\frac{(1-3s+t)\left(2s+t-s^2-st+t^2\right)\left(2+s-2t-s^2-st+t^2\right)}{(2-s+2t)(3+s-2t)(4-2s-t)(1+2s+t)}(\mathbf{u} \times \mathbf{v}).$$

None of the factors, other than $1 - 3s + t$, is ever zero for $s, t \in (0, 1)$; and it follows that for a crosscut quad $\mathcal{A}$, the segments (B_0, B_1) and (B_2, B_3) are parallel if and only if $1 - 3s + t = 0$. Similarly (after all, there is a symmetry at work here), one finds that (B_1, B_2) is parallel to (B_0, B_4) if and only if $2 - s - 3t = 0$. This establishes part (b). (We can also see from this that $\mathcal{B}$ is a parallelogram if and only if $1 - 3s + t = 0 = 2 - s - 3t$, which happens if and only if $s = t = 1/2$, which is in turn true if and only if $\mathcal{A}$ is a parallelogram.)

Finally, for part (c), it was probably noticed earlier that for all $s, t \in [0, 1]$, all the factors in the denominator of (6) are ≥ 1 and that the numerator is zero if and only if $\langle s, t\rangle$ is one of $\langle 0, 0\rangle$, $\langle 1, 0\rangle$, $\langle 0, 1\rangle$, or $\langle 1, 1\rangle$. It is clear from our construction that each of these four cases is equivalent to the (degenerate) situation of the merging of two vertices of $\mathcal{A}$. (For example, $A_2 = A_3$ when $\langle s, t\rangle = \langle 0, 0\rangle$, as in FIGURE 3.) ■

Diagonal triangles There are probably many amusing relationships lurking among the various pieces of our crosscut quad. As an example (found, as with the others, by messing around with *Geometer's Sketchpad*), our next theorem contains a cute result, whose visual proof, too, we abandon to the reader.

First a few easy preliminaries. The diagonals of a quad partition the quad into four triangles. With O being the intersection of the diagonals of $\mathcal{A}$ (the origin mentioned earlier), we define the *diagonal triangles* $\mathcal{D}_i = (A_i, O, A_{i-1})$. If the diagonals of $\mathcal{A}$ are added to the picture, as in FIGURE 6, it is clear that the points B_i must always lie within the respective $\mathcal{D}_i$. Incidentally, it is a simple exercise to show that the centroids D_i of the $\mathcal{D}_i$ form a parallelogram whose area is $(2/9)\,|\mathcal{A}|$. (Engel [**4**, prob. 69, sec. 12.3.1] shows that this result holds when O is *any* point in the quad.) Also known is that the products of the areas of opposing diagonal triangles $\mathcal{D}_i$ are equal. That is,

$$|\mathcal{D}_1|\,|\mathcal{D}_3| = |\mathcal{D}_0|\,|\mathcal{D}_2|\,, \tag{8}$$

which is easily checked by noting $|\mathcal{D}_0| = (1 - s)\,t/2$ and its cyclic counterparts. (A PWW lurks in FIGURE 5. In fact, (8) holds when O is any point on a diagonal of a quadrilateral $A_0A_1A_2A_3$, regardless of whether the quadrilateral is convex or simple. Cross products are superfluous.)

Corner triangles Another odd relationship involves ratios of areas of corner triangles to diagonal triangles. It may be unrelated to (8), but it has a similar flavor.

THEOREM 2. *For an arbitrary quad $\mathcal{A}$,*

$$\frac{|\mathcal{D}_0|}{|\mathcal{C}_0|} + \frac{|\mathcal{D}_2|}{|\mathcal{C}_2|} = \frac{|\mathcal{D}_1|}{|\mathcal{C}_1|} + \frac{|\mathcal{D}_3|}{|\mathcal{C}_3|} = 10.$$

Proof. Notice that the areas of the $\mathcal{D}_i$ are the numerators of the formulas we have for the areas of respective triangles $\mathcal{C}_i$ (see (5)). These cancel in each of the ratios

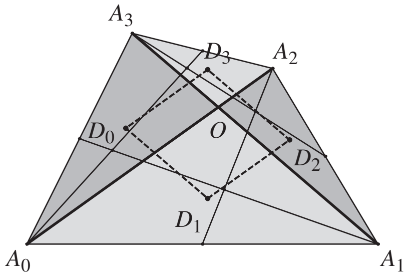


Figure 6 The $\mathcal{D}_i$, individually shaded, and the parallelogram of their centroids

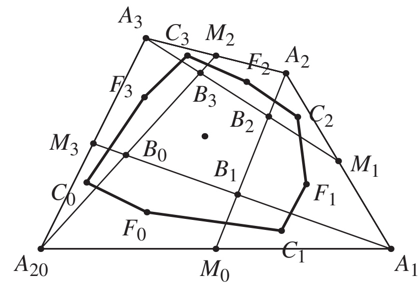


Figure 7 A convex octagon of centroids

$|\mathcal{D}_i| \,/\, |\mathcal{C}_i|$. Thus,

$$\frac{|\mathcal{D}_0|}{|\mathcal{C}_0|} + \frac{|\mathcal{D}_2|}{|\mathcal{C}_2|} = 2(4 - 2s - t) + 2(1 + 2s + t) = 10.$$

It is unnecessary to write a similar formula for the second sum, as it follows by symmetry. ■

A convex octagon In FIGURE 7, the centroids of the nine partitioned regions of our crosscut quad are shown. The outer eight form an octagon: Let $\mathcal{F}_i = (A_i, M_i, B_{i+1}, B_i)$, $i = 0, 1, 2, 3$, denote the flaps and let C_i and F_i denote the centroids of the corners $\mathcal{C}_i$ and the flaps $\mathcal{F}_i$, respectively.

THEOREM 3. *The octagon* $\mathcal{O} = (C_0, F_0, C_1, F_1, C_2, F_2, C_3, F_3)$ *is convex.*

All we offer toward a proof of the convexity is the suggestion already given concerning cross products. (The author's colleague Zsolt Lengvárszky has a nice visual proof.) It suffices to show that $(C_0, F_0) \times (F_0, C_1) > 0$ and $(F_0, C_1) \times (C_1, F_1) > 0$ for every choice of $s, t \in (0, 1)$. It turns out that, using our machinations here, the latter of these inequalities (which would appear from FIGURE 7 to be the more sensitive of the two) is fairly easy, as one gets a rational function of s, t with nice, positive polynomial factors. The former also factors into positive polynomial factors, although more work is involved and one of the factors is of degree five in s, t. We claim very tight bounds on the octagon's relative area:

$$1.888\cdots = \frac{270}{143} \le \frac{|\mathcal{A}|}{|\mathcal{O}|} \le \frac{216}{113} = 1.911\ldots,$$

where the lower bound occurs when $\mathcal{A}$ is a parallelogram, the upper when $\mathcal{A}$ becomes degenerate. (We provide no proof for this conjecture, but the author will attempt one if offered a sufficient cash incentive.)

A look at FIGURE 6 suggests that the centroid of each of the diagonal triangles D_i is inside the corresponding $\mathcal{F}_i$. This too is easily shown using cross products; it suffices to note that $(A_0, M_2) \times (A_0, D_0) = st/6 > 0$ and that D_0 lies on (O, M_3). The reader is challenged to give a nice visual proof of this fact, but it cannot be denied that using cross products is very quick indeed.

A visual proof for the trapezoid case Here now is the visually oriented proof, mentioned earlier, that $|\mathcal{A}| = 5\,|\mathcal{B}|$ when $\mathcal{B}$ is a trapezoid. (This is one direction of Theorem 1(b).) In FIGURE 8, triangle (M_1, A_2, A_3) is rotated 180° about M_1 forming (M_1, A_1, A_3'), and (M_3, A_0, A_1) is likewise rotated about M_3 to form (M_3, A_3, A_1').

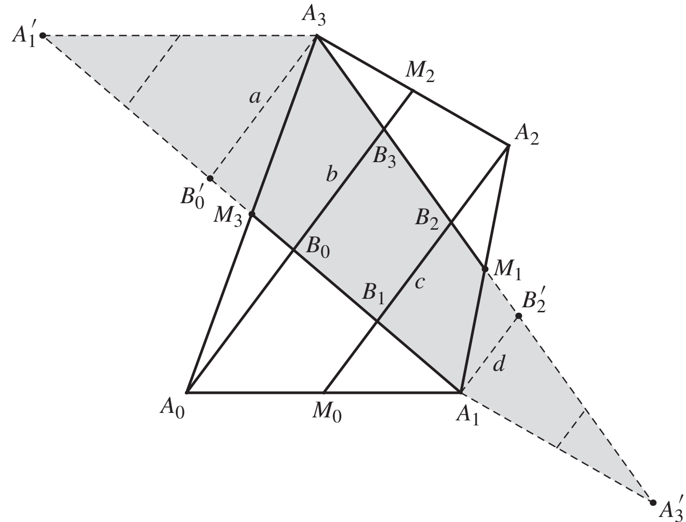


Figure 8 $\mathcal{B}$ is a trapezoid with $(B_1, B_2) \parallel (B_0, B_3)$

(For a point P, denote its rotated counterpart by P'.) Now let h (not annotated) be the perpendicular distance between segments (B_1, B_2) and (B_0, B_3), and let $a = |(B_0', A_3)|$, $b = |(B_0, B_3)|$, $c = |(B_1, B_2)|$, and $d = |(A_1, B_2')|$. Then it is evident that $a + d = b + c$, and we therefore have

$$\begin{aligned}|\mathcal{A}| &= \left|(B_0', A_3, A_1')\right| + \left|(B_0', A_1, B_2', A_3)\right| + \left|(A_1, A_3', B_2')\right| \\ &= \frac{1}{2}a(2h) + \frac{a+d}{2}(3h) + \frac{1}{2}d(2h) \\ &= 5\left(\frac{b+c}{2}\,h\right) = 5\,|\mathcal{B}|\,.\end{aligned}$$

If a condition on $\mathcal{A}$ itself is desired in order that $\mathcal{B}$ be a trapezoid, perhaps the following will satisfy. For fixed A_0, A_1, A_3, we will have $(M_0, A_2) \parallel (M_2, A_0)$ if and only if A_2 lies on the line joining the midpoint M_0 of (A_0, A_1) and the point that lies one-third the way from A_1 to A_3. This will force $(B_1, B_2) \parallel (B_0, B_3)$, as in FIGURE 8. Cyclically permute vertices for the remaining possibilities.

One can think of immediate variations and generalizations to the problems explored in this note. Note that there is a certain *chirality* or handedness in our choice of cross-cutting; FIGURE 9 gives a version with alternate medians. We leave it as an exercise to prove that the areas of the inner quads of the two variations are equal (for a fixed outer quad) if and only if either the outer quad is a trapezoid or one of the diagonals of the outer quad bisects the other. Is there some visual proof of that?

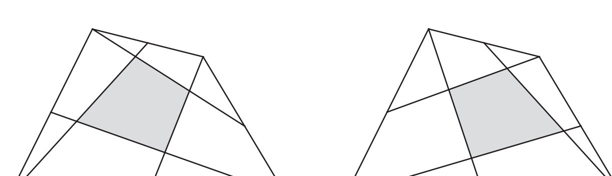


Figure 9 The original and alternate crosscut quads

To generalize, one can use other *cevians* in place of our medians, by letting $M_i = (1 - r)A_i + r\,A_{i+1}$ for some fixed ratio r other than $r = 1/2$, as in FIGURE 10. It is then not difficult to establish a generalization to Theorem 1, in which the minimum and maximum for $|\mathcal{A}|\,/\,|\mathcal{B}|$ are replaced by $(r^2 - 2r + 2)/r^2$ and $(r^2 - r + 1)/r^3$, respectively. One can make more cuts to form an n-crosscut, skewed chessboard, as in FIGURE 11 (where $n = 4$). Obvious analogs of Theorems 1 and 2, and Propositions 2 and 3, hold for such multi-crosscuttings. For more generality, try $m \times n$-crosscut, skewed chessboards. An article by Hoehn [**6**] suggests further problems. As for a multi-crosscut generalization of Theorem 3, well, let's just say that what might seem an obvious generalization is not. (Not what? Not true or not obvious? The intrepid reader should venture forth.)

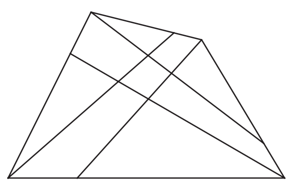


Figure 10 Crosscutting cevians with $r = 1/4$

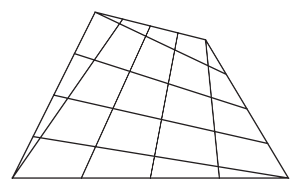


Figure 11 4-crosscut skewed chessboard

Epilogue It turns out that a form of Theorem 1 has appeared earlier. Our diligent referee found a reference to it [**3**, p. 132 (item 15.19)], which in turn names a problems column as a source [**9**]. For the same geometric configuration as ours, but using our notation, the statement in [**3**] gives the inequality $5\,|\mathcal{B}| \le |\mathcal{A}| < 6\,|\mathcal{B}|$ (so it doesn't count the degenerate case). However, [**3**] also states that $|\mathcal{A}| = 5\,|\mathcal{B}|$ "only for a parallelogram." This is incorrect, as we have shown that this equality holds if and only if $\mathcal{B}$ is a trapezoid, in which case neither $\mathcal{A}$ nor $\mathcal{B}$ need be a parallelogram. (It isn't clear in [**3**] whether $\mathcal{A}$ or $\mathcal{B}$ is intended as a parallelogram, but they turn out to be equivalent conditions.) Thus it was necessary to track down the original problem in [**9**] for comparison.

The problems column in question was in *Gazeta Matematică*, which is known to every Romanian mathematician, and is near and dear to most. It is one of the journals of *Societatea de Ştiinţe Matematice din România* (the Romanian Mathematical Society) [**2**]. The problem was posed by the eminent Romanian mathematician, Tiberiu Popoviciu (1906–1975), whose contributions to mathematics are too numerous to mention here. He is immortalized also by the Tiberiu Popoviciu Institute of Numerical Analysis, which he founded in 1957 (a short biography appears on the institute's website [**7**]).

Locating the problem, printed in 1943 (surely a difficult year), was not easy, as no library on the WorldCat® network has the volume. Fortunately, the entire collection is available in electronic form [**5**]. The generous help of Eugen Ionascu (Columbus State University) was enlisted, first to find someone who has access to the electronic format, and second for a translation into English (the text is Romanian). The translation revealed that, in *Gazeta*, Popoviciu gives the inequality as $5\,|\mathcal{B}| \le |\mathcal{A}| \le 6\,|\mathcal{B}|$ and challenges the reader to prove the inequality for all convex quadrilaterals $\mathcal{A}$ and to determine when equality holds. That seems to be the last mention of the problem. If there is a follow-up in later issues of *Gazeta Matematică*, it is hiding well. (It would be nice to know the solution intended by Popoviciu, which is likely more elegant than ours.)

Hearing of the search for the 1943 *Gazeta*, Aurel Stan (at The Ohio State University at Marion) had the following reaction. "*Gazeta Matematică* is one of the dearest things to my heart, although for many years I have not opened it, and I feel that I have betrayed it. It is one of the oldest journals in the world dedicated to challenging mathematical problems for middle and high school students. It has appeared without interruption since 1895." The Hungarian journal *Középiskolai Matematikai Lapok* (Mathematical Journal for High Schools) has a similar mission and has been published since 1894 except for a few years during WWII. Professor Stan continues: "Even during the two world wars the Romanian officers who had subscriptions had it [*Gazeta*] delivered to them in the military camps. It is probably the main reason why today so many foreign-born mathematicians in the United States are from Romania. We all grew up with it. Each month I waited for the newest issue." Professor Ionascu concurred with these sentiments, and described getting hooked on the journal in the seventh grade. Hungarian colleagues say similar things about *Matematikai Lapok*. Professor Stan mentions that even during the time of the Communist regime in Romania, there was a high level of mathematics and respect for mathematicians, adding, "We owe a big part of it to *Gazeta Matematică*."

Acknowledgments Certainly we all owe thanks to those earliest problem-solving journals and societies, and we can make partial payment on that debt by getting our students involved with their present-day counterparts. On that note, the author thanks his colleagues and the students in the Senior Seminar course at LSUS for their time spent listening to much of the above. Special thanks go to Roger B. Nelsen at Lewis & Clark College, who visited LSUS for a few short but enjoyable days during the spring semester of 2008, and whose many articles and books inspire so much visual thinking in mathematics. Thanks, too, to Aurel Stan and to Eugen Ionascu and his far-flung colleagues for help locating Popoviciu's problem.

REFERENCES

1. C. Alsina and R. B. Nelsen, *Math made visual: creating images for understanding mathematics*, Mathematical Association of America, Washington, DC, 2006.
2. M. Berinde and V. Berinde, The Romanian Mathematical Society, *European Mathematical Society Newsletter* **40** (June 2001) 20–22.
3. O. Bottema, R. Ž. Djordjevic, R. R. Janic, D. S. Mitrinovic, and P. M. Vasic, *Geometric Inequalities*, Wolters-Noordhoff Publishing, Groningen, 1969.
4. A. Engel, *Problem-Solving Strategies*, Springer, New York, 1998.
5. Gazeta Matematică (Ediţie Electronică), http://www.gazetamatematica.net.
6. L. Hoehn, Area relations on the skewed chessboard, *College Math. J.* **34** (2003) 232–236. doi:10.2307/3595809
7. Institutul de Calcul "Tiberiu Popoviciu," http://www.ictp.acad.ro.
8. R. B. Nelsen, *Proof Without Words II: More Exercises in Visual Thinking*, Mathematical Association of America, Washington, DC, 2000.
9. T. Popoviciu, Problem 5897, *Gaz. Mat.* **49** (1943) 322.

Summary A convex quadrilateral ABCD is "crosscut" by joining vertices A, B, C, and D to the midpoints of segments CD, DA, AB, and BC, respectively. Some relations among the areas of the resulting pieces are explored, both visually and analytically.

RICK MABRY'S professional and pleasurable pursuits (usually one and the same) seem mainly to involve subdividing time and/or space. As a drummer he cuts up time, and asks, Is it good when the pieces are reassembled and there is some left over? As a mathematician, he cuts up lines and planes and tries to be more careful, though he sometimes does cut corners (as in the present article) using *Mathematica* and Geometer's Sketchpad.

NOTES

From Fourier Series to Rapidly Convergent Series for Zeta(3)

ERNST E. SCHEUFENS
DTU Mathematics
Technical University of Denmark
DK-2800 Kgs. Lyngby, Denmark
e.e.scheufens@mat.dtu.dk

In this note we present a method for computing values of $\zeta(s)$ for integers $s \geq 2$. The method involves Fourier series, and it succeeds in giving exact values for $\zeta(s)$ when s is even. When s is odd the method does not give values in closed form, but we show how it can be extended to provide series expressions for $\zeta(3)$ (and other $\zeta(s)$) that converge much more rapidly than the series that defines the function.

The Riemann zeta function $\zeta(s)$ is for $s > 1$ defined by the infinite series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{1}$$

Euler studied this series for integers $s \geq 2$. He was the first to find the exact value of $\zeta(2)$, which he did in 1734 by generalizing the factorization of polynomials to transcendental functions and using the Taylor expansion for $\sin x$. Later, in a 1744 publication, he gave values of $\zeta(s)$ for even values of s up to $s = 26$. The exact values of $\zeta(s)$ for the first six even values of s are given in the following table.

s	2	4	6	8	10	12
$\zeta(s)$	$\frac{\pi^2}{6}$	$\frac{\pi^4}{90}$	$\frac{\pi^6}{945}$	$\frac{\pi^8}{9450}$	$\frac{\pi^{10}}{93555}$	$\frac{691\pi^{12}}{638512875}$

Generally when $s = 2m$, $m \in \mathbb{N} = \{1, 2, \ldots\}$ the values of the zeta function are related to the Bernoulli numbers by the formula

$$\zeta(2m) = (-1)^{m+1} \frac{(2\pi)^{2m}}{2(2m)!} B_{2m} \tag{2}$$

which was also first proved by Euler. The Bernoulli numbers can be defined recursively by

$$B_0 = 1, \quad \text{and} \quad \sum_{k=0}^{n-1} \binom{n}{k} B_k = 0 \text{ for } n \geq 2. \tag{3}$$

Notice that (2) is valid for $m = 0$ too, since $\zeta(0) = -\frac{1}{2}$, if we extend the definition of $\zeta(s)$ by analytic continuation.

Math. Mag. **84** (2011) 26–32. doi:10.4169/math.mag.84.1.026. © Mathematical Association of America

After his triumph in finding exact values of $\zeta(n)$ for even n, Euler tried to develop a technique for finding an exact value for $\zeta(3)$, but the best he could do was to evaluate the related series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)^3} = \frac{\pi^3}{32}. \tag{4}$$

At a later point he conjectured that

$$\zeta(3) = \alpha(\ln 2)^2 + \beta\frac{\pi^2}{6}\ln 2$$

for some rational numbers α and β.

Little progress was made for a long period after Euler, but in 1978 the French mathematician Roger Apéry proved that $\zeta(3)$ is an irrational number, a great increase in our knowledge about $\zeta(3)$. Therefore $\zeta(3)$ is now known as Apéry's constant. It is still a famous unsolved problem in mathematics to say almost anything more about Apéry's constant, or about $\zeta(2m+1)$ for any $m \in \mathbb{N}$. The best result would be to find exact values involving π, $\ln 2$, and other well-known mathematical constants. Lacking such, there is value in finding rapidly convergent series for $\zeta(3)$ and for $\zeta(s)$ for other odd values of s.

In the following we need a relation between $\zeta(s)$ and $\eta(s)$, the eta function or the alternating zeta function, defined for $s > 0$ by the series

$$\eta(s) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s}\ . \tag{5}$$

From (1) we get $(1 - 2^{1-s})\zeta(s) = \eta(s)$, so for $s > 1$ we have the relation

$$\zeta(s) = (1 - 2^{1-s})^{-1}\eta(s). \tag{6}$$

In the next section we will confirm the known exact values of $\zeta(2)$ and $\zeta(4)$, and generally show how we can determine $\zeta(2m)$, $m \in \mathbb{N}$ from Fourier series for even periodic functions. Applying these Fourier series with a special value such as $x = \pi$, we can determine $\zeta(2)$ and $\zeta(4)$ very simply. Using Fourier series for odd periodic functions does not help in finding exact values of $\zeta(3), \zeta(5), \ldots,$ in the same simple way, but in the following section we shall indicate how to find rapidly convergent series representations for these values.

Euler's work is described by Ayoub [**2**] and Dunham [**5**], and Apéry's result by van der Poorten [**7**]. A good reference for Bernoulli numbers is Apostol's article in this MAGAZINE [**1**] and a good reference for Fourier series is Tolstov's book [**10**]. The series representations we derive below are known, and are described, for example, by Srivistava [**8**]. As we note below, some elements of the method go back to Euler.

Evaluation of $\zeta(2m)$ from Fourier Series

For $m \in \mathbb{N}$, let f be the even 2π-periodic function given by $f(x) = x^{2m}$, $x \in [-\pi, \pi]$. Since f is continuous and piecewise differentiable, we have

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(nx), \quad x \in \mathbb{R}, \tag{7}$$

where

$$a_0 = \frac{2}{\pi} \int_0^{\pi} x^{2m}\, dx = \frac{2\pi^{2m}}{2m+1} \tag{8}$$

and

$$a_n = \frac{2}{\pi} \int_0^{\pi} x^{2m} \cos(nx) dx, \quad n \in \mathbb{N}. \tag{9}$$

If $m = 1$ we get from (7), (8), and (9) that $a_0 = 2\pi^2/3$ and $a_n = 4(-1)^n/n^2$, $n \in \mathbb{N}$, so

$$x^2 = \frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{4(-1)^n}{n^2} \cos(nx), \quad x \in [-\pi, \pi]. \tag{10}$$

Letting $x = \pi$ in (10) gives

$$\pi^2 = \frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{4}{n^2},$$

and thus

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{4}\left(\pi^2 - \frac{\pi^2}{3}\right) = \frac{\pi^2}{6}. \tag{11}$$

Alternatively letting $x = 0$ in (10) we can evaluate $\zeta(2)$ via (5) and (6). Several other proofs for $\zeta(2) = \frac{\pi^2}{6}$ are known. A good reference is Chapman [**3**], where fourteen proofs are given with references. The proof leading to (11) is most similar to Proof 5 in [**3**]. Euler's original proof from 1734 is given as Proof 7 in [**3**].

If $m = 2$ we get from (7), (8), and (9) that $a_0 = 2\pi^4/5$ and $a_n = 8(-1)^n(\pi^2/n^2 - 6/n^4)$, $n \in \mathbb{N}$, so

$$x^4 = \frac{\pi^4}{5} + \sum_{n=1}^{\infty} 8(-1)^n \left(\frac{\pi^2}{n^2} - \frac{6}{n^4}\right) \cos(nx), \quad x \in [-\pi, \pi]. \tag{12}$$

Letting $x = \pi$ in (12) gives

$$\pi^4 = \frac{\pi^4}{5} + \sum_{n=1}^{\infty} 8\left(\frac{\pi^2}{n^2} - \frac{6}{n^4}\right)$$

and by using (11)

$$\zeta(4) = \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{1}{48}\left(-\pi^4 + \frac{\pi^4}{5} + 8\pi^2\zeta(2)\right) = \frac{\pi^4}{90}. \tag{13}$$

We have now seen how we can evaluate $\zeta(2)$ and $\zeta(4)$ from Fourier series for a periodic version of $f(x) = x^{2m}$, $-\pi \le x \le \pi$ for $m = 1$ and $m = 2$. And of course we can continue with $m = 3, 4, \ldots$ and evaluate $\zeta(6), \zeta(8), \ldots$.

Information about $\zeta(3)$ from Fourier Series

In the light of how we have evaluated $\zeta(2)$ and $\zeta(4)$ from Fourier series in the previous section, it is an obvious idea to consider a periodic version of $f(x) = x^3$, $-\pi < x \le$

π when we are looking for information on $\zeta(3)$. But the periodic version of x^3 is not continuous, so we would not have uniform convergence of the Fourier series. So instead we consider the odd 2π-periodic function f given by

$$f(x) = \pi^2 x - x^3, \quad -\pi \le x \le \pi, \tag{14}$$

which is continuous and piecewise differentiable. Thus we have

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx), \quad x \in \mathbb{R}, \tag{15}$$

where

$$b_n = \frac{2}{\pi}\int_0^{\pi} (\pi^2 x - x^3)\sin(nx)\,dx = 12\frac{(-1)^{n-1}}{n^3}, \quad n \in \mathbb{N},$$

so that

$$f(x) = 12\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^3}\sin(nx), \quad x \in \mathbb{R}, \tag{16}$$

and now the Fourier series on the right converges uniformly for $x \in \mathbb{R}$. Setting $x = 0$ or $x = \pi$ in (16) as we did in the last sections just gives $0 = 0$. Setting $x = \pi/2$ in (16) gives

$$\frac{3\pi^3}{8} = 12\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^3}\sin\left(\frac{n\pi}{2}\right) = 12\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)^3},$$

or

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)^3} = \frac{\pi^3}{32},$$

which is Euler's result, equation (4). That is the best we can do by setting x to a particular value in (16).

It is the sine function which gives us problems, so now we choose a method where the sine function disappears using the well-known result

$$\int_0^{\infty} \frac{\sin(nx)}{x}dx = \frac{\pi}{2}, \quad n \in \mathbb{N}. \tag{17}$$

The series for $f(x)/x$ obtained by dividing (16) by x converges uniformly. Therefore

$$\int_0^{\infty} \frac{f(x)}{x}dx = 12\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^3}\int_0^{\infty} \frac{\sin(nx)}{x}dx = 12\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^3}\frac{\pi}{2} = 6\pi\,\eta(3).$$

Now using (6) with $s = 3$ gives

$$\zeta(3) = \frac{4}{3}\eta(3) = \frac{2}{9\pi}\int_0^{\infty} \frac{f(x)}{x}dx. \tag{18}$$

Because f is 2π-periodic

$$\int_0^{\infty} \frac{f(x)}{x}dx = \int_0^{\pi} (\pi^2 - x^2)\,dx + \sum_{k=1}^{\infty} \int_{(2k-1)\pi}^{(2k+1)\pi} \frac{\pi^2(x-2k\pi) - (x-2k\pi)^3}{x}dx,$$

where

$$\int_0^{\pi} (\pi^2 - x^2)\, dx = \frac{2\pi^3}{3},$$

and

$$\int_{(2k-1)\pi}^{(2k+1)\pi} \frac{\pi^2(x - 2k\pi) - (x - 2k\pi)^3}{x} dx$$
$$= 2\pi^3 \left[(4k^3 - k) \ln \frac{2k+1}{2k-1} - 4k^2 + \frac{2}{3} \right], \quad k \in \mathbb{N}.$$

Then from (18) we find

$$\zeta(3) = \frac{4\pi^2}{27} + \frac{4\pi^2}{9} \sum_{k=1}^{\infty} \left[(4k^3 - k) \ln \frac{2k+1}{2k-1} - 4k^2 + \frac{2}{3} \right]. \tag{19}$$

It is not obvious from equation (19) itself that the series on the right is convergent. It follows from the derivation of the equation. However, we shall give another proof that leads ultimately to another series for $\zeta(3)$. First we rewrite (19) in a form with a more rapidly convergent series. Setting

$$a_k = (4k^3 - k) \ln \frac{2k+1}{2k-1} - 4k^2 + \frac{2}{3} = (4k^3 - k) \ln \frac{1 + \frac{1}{2k}}{1 - \frac{1}{2k}} - 4k^2 + \frac{2}{3},$$

and using that

$$\ln \frac{1+x}{1-x} = \sum_{n=0}^{\infty} \frac{2}{2n+1} x^{2n+1}, \quad -1 < x < 1,$$

we get

$$\begin{aligned}
\sum_{k=1}^{\infty} a_k &= \sum_{k=1}^{\infty} \left[(4k^2 - 1) \sum_{n=1}^{\infty} \frac{1}{2n+1} \left(\frac{1}{2k} \right)^{2n} - \frac{1}{3} \right] \\
&= \sum_{k=1}^{\infty} \left[\left(4k^2 \sum_{n=1}^{\infty} \frac{1}{2n+1} \left(\frac{1}{4k^2} \right)^n - \frac{1}{3} \right) - \sum_{n=1}^{\infty} \frac{1}{2n+1} \left(\frac{1}{4k^2} \right)^n \right] \\
&= \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \left(\frac{1}{2n+3} - \frac{1}{2n+1} \right) \left(\frac{1}{4k^2} \right)^n,
\end{aligned} \tag{20}$$

from which it follows, that $a_k < 0$, and

$$|a_k| = \sum_{n=1}^{\infty} \left(\frac{1}{2n+1} - \frac{1}{2n+3} \right) \left(\frac{1}{4k^2} \right)^n < \frac{2}{15} \sum_{n=1}^{\infty} \left(\frac{1}{4k^2} \right)^n = \frac{2}{15} \cdot \frac{1}{4k^2 - 1}.$$

We know that

$$\sum_{k=1}^{\infty} \frac{1}{4k^2 - 1} = \frac{1}{2} \left(\sum_{k=1}^{\infty} \left(\frac{1}{2k-1} - \frac{1}{2k+1} \right) \right) = \frac{1}{2},$$

and thus the comparison test shows that $\sum_{k=1}^{\infty}|a_k|$ is convergent. We have now shown that the series on the right in (19) is (absolutely) convergent. Inserting the result from (20) into (19) we get

$$\begin{aligned}\zeta(3) &= \frac{4\pi^2}{27} + \frac{4\pi^2}{9}\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\left(\frac{1}{2n+3} - \frac{1}{2n+1}\right)\left(\frac{1}{4k^2}\right)^n \\ &= \frac{4\pi^2}{27} - \frac{8\pi^2}{9}\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{(2n+1)(2n+3)}\left(\frac{1}{4k^2}\right)^n.\end{aligned}$$

The series on the last line has positive terms, so we can change the order of summation to get

$$\begin{aligned}\zeta(3) &= \frac{4\pi^2}{27} - \frac{8\pi^2}{9}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{1}{(2n+1)(2n+3)}\left(\frac{1}{4k^2}\right)^n \\ &= \frac{4\pi^2}{27} - \frac{8\pi^2}{9}\sum_{n=1}^{\infty}\frac{1}{(2n+1)(2n+3)4^n}\sum_{k=1}^{\infty}\frac{1}{k^{2n}},\end{aligned}$$

and thus we can evaluate $\zeta(3)$ from

$$\zeta(3) = \frac{4\pi^2}{27} - \frac{8\pi^2}{9}\sum_{n=1}^{\infty}\frac{\zeta(2n)}{(2n+1)(2n+3)4^n}, \tag{21}$$

where the terms depend on $\zeta(2n)$, which, as we have seen, are given by equation (2). Using $\zeta(0) = -\frac{1}{2}$, we can express (21) more compactly as

$$\zeta(3) = -\frac{8\pi^2}{9}\sum_{n=0}^{\infty}\frac{\zeta(2n)}{(2n+1)(2n+3)4^n}. \tag{22}$$

If $s > 1$ it follows from (6) that we have $\zeta(s) < (1 - 2^{1-s})^{-1}$ since $\eta(s) < 1$. So using the Nth partial sum of the infinite series in (21) or (22), we can compute $\zeta(3)$ with an error R_N satisfying

$$\begin{aligned}|R_N| &< \frac{8\pi^2}{9}\frac{\zeta(2N+2)}{(2N+3)(2N+5)}\sum_{n=N+1}^{\infty}\frac{1}{4^n} \\ &< \frac{8\pi^2}{27}\frac{1}{(2N+3)(2N+5)(4^N - \frac{1}{2})}.\end{aligned}$$

Using the 25th partial sum in (22) we have an error bound $|R_{25}| < 9 \cdot 10^{-19}$, and approximately the value $\zeta(3) = 1.20205690315959429$. For comparison using the Nth partial sum of $\sum_{n=1}^{\infty} 1/n^3$, we can compute $\zeta(3)$ with an error R_N satisfying

$$R_N = \sum_{n=N+1}^{\infty}\frac{1}{n^3} < \int_N^{\infty}\frac{1}{x^3}dx = \frac{1}{2N^2},$$

from which we get the bound $R_N < 8 \cdot 10^{-4}$ if $N = 25$.

We can write the series representation (22) in another form, because

$$\sum_{n=0}^{\infty}\frac{\zeta(2n)}{(2n+1)4^n} = -\frac{1}{2}\ln 2,$$

see [**9**, p. 837]. Inserting in (22) gives the alternative form

$$\zeta(3) = \frac{2\pi^2}{9}\left(\ln 2 + 2\sum_{n=0}^{\infty}\frac{\zeta(2n)}{(2n+3)4^n}\right), \tag{23}$$

cf. [**8**]. Notice that the term $(2\pi^2/9)\ln 2$ has the form $\beta(\pi^2/6)\ln 2$ as Euler conjectured could be a term in an exact value for $\zeta(3)$, cf. [**5**].

Conclusion Starting with the continuous and piecewise differentiable 2π-periodic versions of x^2 and x^4, $x \in [-\pi, \pi]$ we have seen how we can evaluate $\zeta(2)$ and $\zeta(4)$ simply by setting $x = \pi$ in the Fourier series of the periodic versions of x^2 and x^4, equations (11) and (13). In the same manner we can evaluate $\zeta(6), \zeta(8), \ldots$ by setting $m = 3, 4, \ldots$ in the periodic version of $f(x) = x^{2m}$, $x \in [-\pi, \pi]$. Starting with the continuous, piecewise differentiable 2π-periodic version of $f(x) = \pi^2 x - x^3$, $x \in [-\pi, \pi]$ we have illustrated how, by integraton of $f(x)/x$, we can find a series with logarithmic terms for $\zeta(3)$, equation (19), and then use power series for logarithmic functions to get some rapidly convergent series for $\zeta(3)$, equations (21) and (22). The series representations (23) for $\zeta(3)$ is well-known [**8**, p. 585]. A series representation analogous to (22) was contained in a 1772 paper by Euler [**2**, pp. 1084–1085] and later rediscovered by others [**8**, p. 571]. Several other known series representations of $\zeta(3)$, and more generally several other known series representations of $\zeta(2n+1)$, $n \in \mathbb{N}$, can also be found in [**8**]. There are also some interesting integral representations for $\zeta(2n+1)$, $n = 1, 2, 3, \ldots$ which can be found from Fourier series for odd functions [**4**] and [**6**].

Acknowledgment Many thanks to Preben K. Alsholm for his valuable assistance.

REFERENCES

1. Tom M. Apostol, A primer on Bernoulli numbers and polynomials, *Math. Mag.* **81** (2008) 178–190.
2. Raymond Ayoub, Euler and the zeta function, *Amer. Math. Monthly* **81** (1974) 1067–1086.
3. Robin Chapman, Evaluating $\zeta(2)$, http://pictor.math.uqam.ca/~plouffe/OEIS/archive_in_pdf/zeta2.pdf (2003).
4. D. Cvijović and J. Klinowski, Integral representations of the Riemann zeta function for odd-integer arguments, *J. Comput. Appl. Math.* **142** (2002) 435–439. doi:10.1016/S0377-0427(02)00358-8
5. William Dunham, *Euler: The Master of Us All*, Mathematical Association of America, Washington, DC, 1999.
6. J. Duoandikoetxea, A sequence of polynomials related to the evaluation of the Riemann zeta function, *Math. Mag.* **80** (2007) 38–45.
7. A. van der Poorten, A proof that Euler missed... Apéry's proof of the irrationality of $\zeta(3)$, *Math. Intelligencer* **1** (1979) 195–203. doi:10.1007/BF03028234
8. H. M. Srivastava, Some families of rapidly convergent series representations for the zeta functions, *Taiwanese J. Math.* **4** (2000) 569–599; available at http://www.math.nthu.edu.tw/~tjm/abstract/0012/tjm0012_3.pdf.
9. H. M. Srivastava, M. L. Glasser, and V. S. Adamchik, Some definite integrals associated with the Riemann zeta function, *Journal for Analysis and Its Applications* **19** (2000) 831–846.
10. G. P. Tolstov, *Fourier Series*, Dover, New York, 1976.

Summary Exact values of the Riemann zeta function $\zeta(s)$ for even values of s can be determined from Fourier series for periodic versions of even power functions, but there is not an analogous method for determining exact values of $\zeta(s)$ for odd values of s. After giving a brief historical overview of $\zeta(s)$ for integer values of s greater than one, and showing how we can determine $\zeta(2)$ and $\zeta(4)$ from Fourier series, we consider the Fourier series for a continuous and piecewise differentiable odd periodic function from which we can find a series with logarithmic terms for $\zeta(3)$. Using power series for logarithmic functions on this series, a rapidly convergent series for $\zeta(3)$ is obtained. Using a partial sum of this series we can compute $\zeta(3)$ with an error which is much smaller than the error obtained by using a similar partial sum of the infinite series defining $\zeta(3)$.

Three Approaches to a Sequence Problem

GREGORY MINTON
Massachusetts Institute of Technology
Cambridge, MA 02139
gminton@math.mit.edu

In mathematics we often find that a single problem may be solved by a variety of methods, with each yielding new insights and perspective. In this note we solve a particular problem with three different methods, finding that each method suggests the same generalization (although for varying reasons!). These methods are not new; our goal in collecting them here is to highlight the connections between different techniques.

The focus of our study has appeared (in slightly different forms) in the problem sections of the *American Mathematical Monthly* in 1908 [**3**] and the *College Math Journal* in 1999 [**4**].

DEFINITION. *An integer sequence* $\{x_n\}$ *is* ***prime-divisible*** *if* $p \mid x_p$ *for every prime* p.

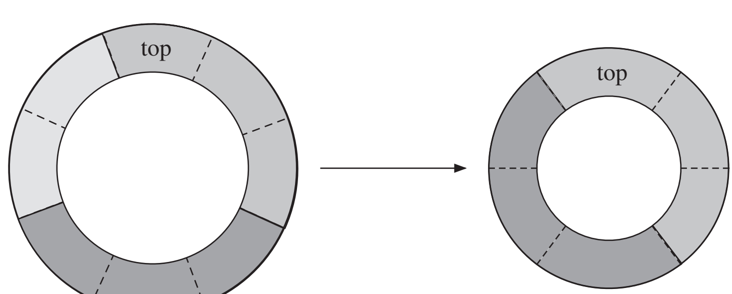
PROBLEM. *Prove that the sequence defined by* $x_1 = 0$, $x_2 = 2$, $x_3 = 3$, *and*

$$x_n = x_{n-2} + x_{n-3} \quad \text{for } n \geq 4$$

is prime-divisible.

Combinatorial argument Our first method relies on the combinatorial maxim of "telling a story" about what the sequence is counting. In the same spirit as Benjamin and Quinn [**1**], we use tilings to interpret the sequence. Consider a circular strip consisting of n equal cells. We wish to tile this strip with pieces that cover two cells (called dominos) and pieces that cover three cells (called triominos). Unlike many other problems of arrangements on a circle, we consider tilings that are related by a rotation to be distinct; in other words, a tiling has a fixed orientation.

The question, quite naturally, is how many tilings are possible for a given n. Denote this count by t_n. We show that, in fact, $t_n = x_n$. It is clear that $t_1 = 0$, $t_2 = 2$ (there are two possible rotational "phases" for the domino), and similarly $t_3 = 3$. Now for $n \geq 4$, pick a particular cell of the circular strip as the "top." For any given tiling, locate the piece covering the top; remove the piece directly beside it (in the counterclockwise direction), and paste the strip back together. What results is an oriented tiling of a smaller strip. One example is shown below.



Math. Mag. **84** (2011) 33–37. doi:10.4169/math.mag.84.1.033. © Mathematical Association of America

If the piece removed was a domino, then what results is a tiling of a $(n-2)$-length strip; if the piece removed was a triomino, then what results is a tiling of a $(n-3)$-length strip. In fact, we get a bijection between the set of tilings of length n and the set of tilings of length $n-2$ or $n-3$. Thus $t_n = t_{n-2} + t_{n-3}$; since $\{t_n\}$ obeys the same initial conditions and recurrence relation as $\{x_n\}$, the two are equal for all n. We have indeed combinatorially represented the sequence in our problem.

Now consider tilings of a prime length p. We prove by contradiction that all of the p rotations of a given tiling are distinct, showing that the total number of tilings must be divisible by p. Suppose there is a tiling such that a rotation by k cells $(0 < k < p)$ leaves the tiling invariant. Since p is prime, k and p are relatively prime; thus there exist integers x and y such that

$$kx + py = 1.$$

Interpreting this physically, if we perform x rotations of the tiling by k cells (under which it is invariant) and y rotations by p cells (again, under which it is invariant), then the result is a rotation by one cell. This shows the tiling is invariant under rotation by a single cell, which is impossible as there are no length 1 pieces. This contradiction completes the proof.

If we examine this argument with an eye for generalization, we see that the absence of any length 1 piece plays the pivotal role. Therefore we do not expect to be able to handle recurrence relations where x_n depends directly on x_{n-1}. We can, however, handle a more general recurrence relation by using a technique found in Benjamin and Quinn [**1**]: we *color* the pieces. Suppose the scenario is the same as above, except there are γ different colors of dominos to use and δ different colors of triominos. By the same proof, the number of ways of tiling a p-length strip will again be divisible by p, and this number is the pth term of the sequence

$$x_1 = 0, x_2 = 2\gamma, x_3 = 3\delta, x_n = \gamma x_{n-2} + \delta x_{n-3} \quad \text{for all } n \geq 4.$$

Thus for any positive integers γ, δ, the sequence above is prime-divisible.

Generating functions Our second method to solve this problem uses the generating function for the series x_n. The generating function for a sequence $\{a_1, a_2, \ldots\}$ is the formally defined series $f(t) = a_1 t^1 + a_2 t^2 + \cdots$. Thus, in our case we are interested in the function

$$f(x) = 0t^1 + 2t^2 + 3t^3 + 2t^4 + 5t^5 + \cdots.$$

The coefficients x_n grow exponentially in n; the ratio test thus shows that for small enough t, $f(t)$ is a well-defined, smooth function. Consider the power series $(t^2 + t^3) f(t)$. Thanks to the recurrence relation, this must match $f(t)$ in all terms of order at least 4. We find that

$$(t^2 + t^3) f(t) = f(t) - 2t^2 - 3t^3,$$

and so

$$f(t) = \frac{2t^2 + 3t^3}{1 - t^2 - t^3} = (-t)\frac{d}{dt}\left[\log(1 - t^2 - t^3)\right].$$

The pth term of the sequence is given from the generating function by $f^{(p)}(0)/p!$. If we take p derivatives of the above product with the Leibniz rule, notice that only one term survives after setting $t = 0$. This term corresponds to applying exactly one derivative to $(-t)$. Thus,

$$\begin{aligned} x_p &= \frac{1}{p!} \cdot \left(p \cdot (-1) \cdot \frac{d^p}{dt^p} \left[\log(1 - t^2 - t^3) \right] \right) \Big|_{t=0} \\ &= (-p) \cdot (\text{coefficient of } t^p \text{ in } \log(1 - t^2 - t^3)). \end{aligned}$$

Now using the Taylor series for $\log(1 - t)$, we find that

$$-\log(1 - t^2 - t^3) = (t^2 + t^3) + \frac{(t^2 + t^3)^2}{2} + \frac{(t^2 + t^3)^3}{3} + \cdots .$$

For each $n \geq p$, $(t^2 + t^3)^n$ contributes no t^p term. Thus, the coefficient c of t^p is a sum of fractions, all of whose denominators are less than p. For prime p, this implies that the denominator of c in reduced form is relatively prime with p, so $x_p = -pc$ is indeed divisible by p. This completes the second proof.

As before, we look to generalize this proof technique to other sequences. If the recurrence relation had x_{n-1} dependence, we would have introduced a linear term into the $t^2 + t^3$ factor. Thus we could no longer have argued away the pth term in the log series, which was a key step in the proof. For this reason we do not expect to handle recurrences with x_{n-1} dependence. For the most general remaining recurrence, $x_n = \gamma x_{n-2} + \delta x_{n-3}$, this proof technique still works as long as the numerator of the resulting generating function $f(t)$ is (a constant times) t times the derivative of the denominator. Since the denominator is $1 - \gamma t^2 - \delta t^3$ and the numerator is $x_1 t + x_2 t^2 + x_3 t^3$, prime-divisibility holds for multiples of the initial conditions $x_1 = 0$, $x_2 = -2\gamma$, $x_3 = -3\delta$. Except for a superficial negative sign, this is the same form as the generalization we found in the previous section!

Field-theoretic solution For our third and final solution method, we use some field theoretic ideas. This proof requires slightly more background than the others; we refer the reader to any introductory algebra text, such as Dummit and Foote [2].

Consider the characteristic polynomial

$$f(t) = t^3 - t - 1$$

for the sequence $\{x_n\}$. This has three distinct roots r_1, r_2, and r_3 in the complex plane; thus there exist complex constants A, B, and C such that

$$x_n = Ar_1^n + Br_2^n + Cr_3^n$$

for all n. We show that $A = B = C = 1$.

It suffices to verify equality for $n = 0$, $n = 1$, and $n = 2$, since those initial conditions determine the sequence. (While the problem did not specify a zeroth term of $\{x_n\}$, it is defined uniquely by extending the recurrence relation. Solving $x_3 = x_1 + x_0$, we see that $x_0 = 3$.) Obviously $r_1^0 + r_2^0 + r_3^0 = 3 = x_0$. Further, since

$$f(t) = t^3 - t - 1 = (t - r_1)(t - r_2)(t - r_3),$$

by equating quadratic coefficients we have $r_1^1 + r_2^1 + r_3^1 = 0 = x_1$. Equating linear coefficients, $r_1 r_2 + r_1 r_3 + r_2 r_3 = -1$; thus

$$r_1^2 + r_2^2 + r_3^2 = (r_1 + r_2 + r_3)^2 - 2(r_1 r_2 + r_1 r_3 + r_2 r_3) = 2 = x_2.$$

This shows that, indeed, $x_n = r_1^n + r_2^n + r_3^n$ for all n.

Now fix a prime p. By the Multinomial Theorem for the power of a three-term sum, a generalization of the Binomial Theorem, we have

$$(r_1 + r_2 + r_3)^p = \sum_{\substack{a,b,c\geq 0 \\ a+b+c=p}} \frac{p!}{a!\,b!\,c!} r_1^a r_2^b r_3^c.$$

The coefficient $p!/(a!\,b!\,c!)$ is always an integer and, unless (a, b, c) is some permutation of $(p, 0, 0)$, it is divisible by p. Thus we may write $(r_1 + r_2 + r_3)^p = r_1^p + r_2^p + r_3^p + p \cdot z$, where

$$z = \sum_{\substack{0\leq a,b,c<p \\ a+b+c=p}} \frac{(p-1)!}{a!\,b!\,c!} r_1^a r_2^b r_3^c$$

is some integer-linear combination of products of r_1, r_2, and r_3. Substituting $x_n = r_1^n + r_2^n + r_3^n$,

$$\begin{aligned} x_1^p &= x_p + p \cdot z \\ -x_p &= p \cdot z. \end{aligned}$$

Now z is an algebraic integer; since $p \cdot z$ is an integer, it follows that z is in fact an integer. Thus $p \mid x_p$, completing the proof.

We again attempt to generalize this proof. It was crucial that f have three distinct roots r_1, r_2, r_3, and that the sum $r_1 + r_2 + r_3$ vanish. This restricts us to recurrences of the form $x_n = \gamma x_{n-2} + \delta x_{n-3}$. Now the proof will proceed for any multiple of $x_n = r_1^n + r_2^n + r_3^n$; this gives the initial conditions $x_1 = 0$, $x_2 = 2\gamma$, $x_3 = 3\delta$. We have yet again found the same generalization.

Conclusions Our three methods offer different interpretations of the problem. Depending on one's point of view, the initial conditions that made our sequence prime-divisible arose from

- counting base case tilings of a circular strip;
- matching the numerator of the generating function with the derivative of the denominator; or
- using sums of powers of roots of the characteristic polynomial.

Also notably, the three methods used the condition of primality in slightly different ways.

It is especially compelling that each of these interpretations led to the same natural generalization of the problem, giving initial conditions for $x_n = \gamma x_{n-2} + \delta x_{n-3}$ to be prime-divisible. The fact that this generalization arose thrice suggests that it is actually the "correct" one. In fact, with suitable restrictions on the characteristic polynomial, it does cover all prime-divisible third-order recurrences.

If those restrictions are relaxed, though, there are more families of prime-divisible sequences. We prove this, giving a catalogue of such sequences, in an upcoming companion paper. In the mean time, we invite the reader to use any of these three approaches (or another!) to discover these additional recurrence families for yourself.

Acknowledgment The author is funded by a Hertz Foundation Graduate Fellowship. He would also like to acknowledge fellow Harvey Mudd College alumni Fred Johnson and Theodore Spaide, from whom he first learned of the combinatorial proof for this problem.

REFERENCES

1. A. Benjamin and J. Quinn, *Proofs that Really Count: The Art of Combinatorial Proof*, Mathematical Association of America, Washington, DC, 2003.
2. D. Dummit and R. Foote, *Abstract Algebra*, 3rd ed., John Wiley, Hoboken, NJ, 2004.
3. E. Escott, Solution to Problem 151, *Amer. Math. Monthly* **15** (1908) 187. doi:10.2307/2969581
4. R. Tudoran, "A Well-Known Sequence," Solution to Problem 653, *College Math. J.* **31** (2000) 223–224. doi:10.2307/2687495

Summary In this note we examine a well-studied problem concerning the terms of a certain linear recurrence modulo prime numbers. We present three solutions to this problem and examine the similarities and differences between them. In particular, despite using primality in different ways, all three proofs yield the same generalization of the original problem.

Isoperimetric Sets of Integers

STEVEN J. MILLER
Williams College
Williamstown, MA 01267
Steven.J.Miller@williams.edu

FRANK MORGAN
Williams College
Williamstown, MA 01267
Frank.Morgan@williams.edu

EDWARD NEWKIRK
Brown University
Providence, RI 02912
enewkirk@math.brown.edu

LORI PEDERSEN
Williams College
Williamstown, MA 01267
lap811@aol.com

DEIVIDAS SEFERIS
Williams College
Williamstown, MA 01267
deividas.seferis@gmail.com

The celebrated isoperimetric theorem says that the circle provides the least-perimeter way to enclose a given area. In this note we discuss a generalization which arose at a departmental research seminar [**1**] and which moves the isoperimetric problem from geometry to number theory and combinatorics. Instead of Euclidean space, let's take the set $\mathbb{N}_0$ of nonnegative integers:

$$\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}.$$

For any subset S of $\mathbb{N}_0$, we define volume and perimeter as follows:

$$\begin{aligned} \mathrm{vol}(S) &:= \text{sum of elements of } S \\ \mathrm{per}(S) &:= \text{sum of elements of } S \text{ whose predecessor and successor} \\ &\quad\ \ \text{are not both in } S. \end{aligned}$$

Math. Mag. **84** (2011) 37–42. doi:10.4169/math.mag.84.1.037.

For example, for $S = \{0, 1, 4, 5, 6, 7\}$, $\mathrm{vol}(S) = 23$ and $\mathrm{per}(S) = 0 + 1 + 4 + 7 = 12$. This definition is a natural generalization of the original problem; the perimeter comes from elements that are on the "boundary" of our set, as these elements belong to S but have a neighbor that does not.

We can now state the problem we want to consider:

ISOPERIMETRIC SET PROBLEM. *Among all sets $S \subset \mathbb{N}_0$ whose volume is n, find the set S with smallest perimeter.*

We write $P(n)$ for the smallest perimeter consistent with volume n. For example, among all sets of nonnegative integers that sum to $n = 9$, the smallest possible perimeter turns out to be 6, with $S = \{2, 3, 4\}$. For $n = 19$, the smallest perimeter is 10, with $S = \{0, 1, 3, 4, 5, 6\}$. For $n \leq 135$, we find $P(n)$ by an exhaustive computer search; FIGURE 1 shows a plot.

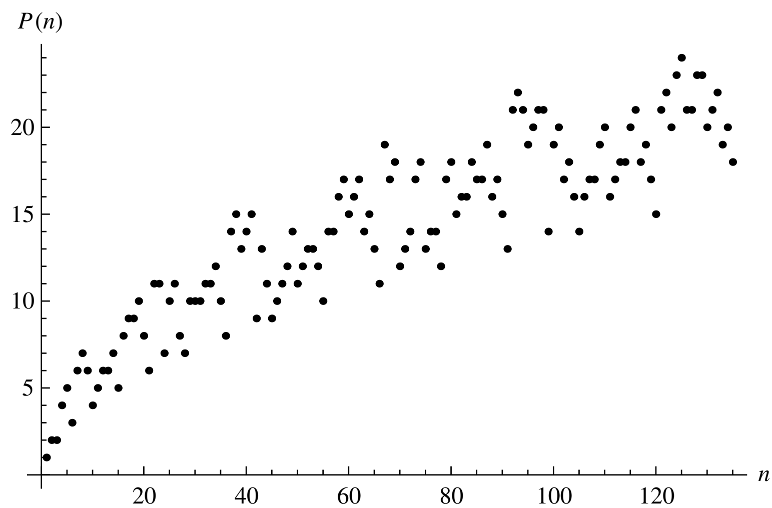


Figure 1 Plot of minimum perimeter $P(n)$ versus prescribed volume n.

How large is a typical $P(n)$ relative to n? Does $P(n)$ grow linearly with n, or is it significantly slower? Unfortunately, our plot above has only 135 data points, and it is very easy to be misled as to the limiting behavior from such a small data set. (Similar limitations of computations arise when counting prime numbers. The famous Prime Number Theorem states that the fraction of numbers from 1 to n which are prime is about $1/\log n$. This tends to zero, albeit very slowly. Of the numbers from 1 to 10, 40% are prime (2, 3, 5, and 7), and of the numbers from 1 to 20, still 40% are prime, without any indication that the percentage is approaching 0.)

As our problem generalizes the classical isoperimetric problem, perhaps the solution to that problem can suggest what the true behavior of $P(n)$ should be. Although our problem lies on the number line, each point is weighted by its value, in some sense adding a dimension. So perhaps the two-dimensional classical isoperimetric problem would be a good guide. We know that the optimal solution in that case is a circle. As the area of a circle of radius r is $A = \pi r^2$ and the perimeter is $P = 2\pi r$, simple algebra yields that the least perimeter for a given area is $P = 2\sqrt{\pi}\, A^{1/2}$. Given this, it is not unreasonable to conjecture for large n that $P(n)$ should approximately equal a constant times $\sqrt{n}$.

It turns out that the classical problem *does* provide the right intuition, as the following proposition shows. It says that $P(n)$ is asymptotic to $\sqrt{2}\, n^{1/2}$; that is, the ratio of

$P(n)$ to $\sqrt{2}\,n^{1/2}$ approaches 1 as n tends to infinity. The two bounds in the proposition are graphed in FIGURE 2.

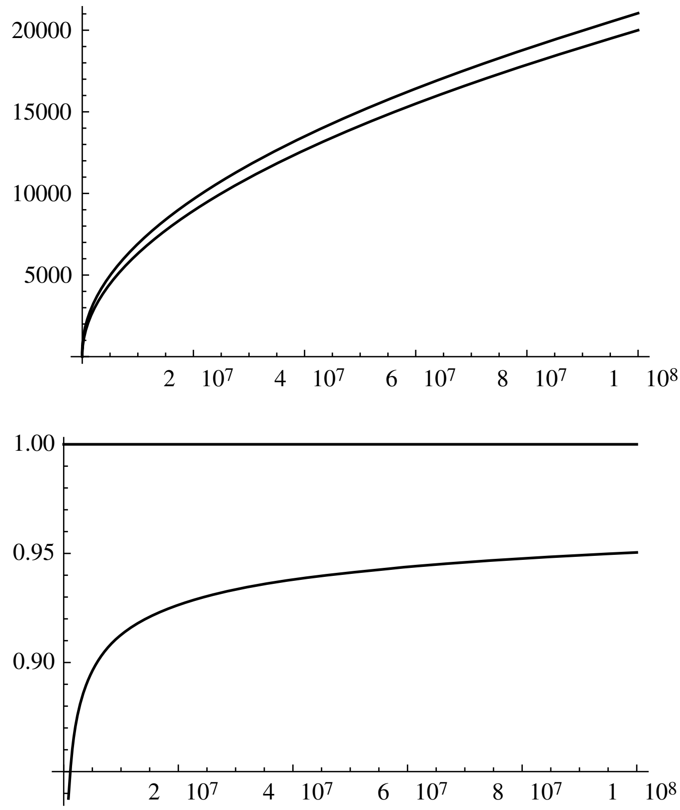


Figure 2 *Top:* Plot of the upper bound $\sqrt{2}\,n^{1/2} + (2n^{1/4} + 8)\log_2\log_2 n + 58$ and the lower bound $\sqrt{2}\,n^{1/2} - 1/2$. As n approaches infinity, our two bounds asymptotically approach $\sqrt{2}\,n^{1/2}$. *Bottom:* Plot of the ratio of the upper and lower bound. The two bounds have the same growth rate in the limit, but the convergence is very slow.

As we will see, the proof of the lower bound is trivial. The proof of the upper bound is a tricky induction argument. The hardest part is finding the right induction hypothesis, a process which, for us, involved much trial and error.

PROPOSITION. $P(n) \sim \sqrt{2}\,n^{1/2}$. *Indeed, for* $n \geq 2$,

$$\sqrt{2}\,n^{1/2} - 1/2 < P(n) < \sqrt{2}\,n^{1/2} + (2n^{1/4} + 8)\log_2\log_2 n + 58.$$

Proof. The perimeter of S is at least as large as the largest element of S, which we denote by m. For the lower bound, if $\mathrm{vol}(S) = n$ then m must satisfy

$$n \leq 0 + 1 + \cdots + m = \frac{m(m+1)}{2}. \tag{1}$$

By the quadratic formula, equality holds if m equals $f(n)$, where

$$f(n) := \left(2n + \frac{1}{4}\right)^{1/2} - \frac{1}{2},$$

and the inequality in (1) is equivalent to

$$m \geq f(n) = \left(2n + \frac{1}{4}\right)^{1/2} - \frac{1}{2} > \sqrt{2}\, n^{1/2} - 1/2.$$

For the upper bound, we use the following greedy algorithm. Given n, we will construct a set S with volume n and reasonably small perimeter. Choose the largest term m_1 in S to be as small as possible. For the remaining volume, continue taking consecutive numbers as long as possible, through $m_2 \leq m_1$. Choose the next term $m_3 < m_2$ as small as possible. Continue. For example, for $n = 19$, this algorithm yields $S = \{6, 5, 4, 3, 1, 0\}$, which turns out to be optimal. For $n = 11$ it yields $\{5, 4, 2\}$ with perimeter 11, which is worse than the optimal $\{5, 3, 2, 1, 0\}$ with perimeter 8.

First note that $m_1 < f(n) + 1$, because for any integer $m \geq f(n)$, (1) holds. Similarly for any odd k, $m_k < f(m_{k-1} - 1) + 1$ because by choice of m_{k-1}, the remaining volume is less than $m_{k-1} - 1$.

For even k, we stopped at m_k because the next (smaller) integer $m_k - 1$ exceeded the remaining volume. The sequence of chosen integers stopping at m_k began at m_{k-1}, which by its choice exceeds the sum of all successive unused integers, including $m_k - 1$; or excluding $m_k - 1$, the sum of all the integers starting with $p = m_k - 2$:

$$\frac{p(p+1)}{2} < m_{k-1}.$$

Hence, $m_k - 2 < f(m_{k-1})$. In summary, for all $k \geq 2$,

$$m_k < f(m_{k-1}) + 2. \tag{2}$$

We now consider a simpler new function $g(n)$ which is closely related to $f(n)$ and avoids the complications of the $+2$. Letting

$$g(m) := \sqrt{2}\, m^{1/2} + 2,$$

we find

$$\begin{aligned} f(m) + 2 &= \left(2m + \frac{1}{4}\right)^{1/2} + \frac{3}{2} \\ &\leq \sqrt{2}\, m^{1/2} + \frac{1}{2} + \frac{3}{2} \\ &= \sqrt{2}\, m^{1/2} + 2 = g(m). \end{aligned}$$

We now consider *compositional powers* $g^k(n)$, where for example $g^3(n) = g(g(g(n)))$. By (2), for $k \geq 2$, the compositional power $g^k(n)$ satisfies

$$m_k < g^k(n), \tag{3}$$

without the pesky $+2$ of (2). Now all we need is an upper bound on $g^k(n)$. We'll prove by induction that for $k \geq 1$

$$g^k(n) < 2^{1-1/2^k} n^{1/2^k} + 8. \tag{4}$$

The base case $k = 1$ is immediate. The induction step takes just a little algebra:

$$\begin{aligned} g^{k+1}(n) &< 2^{1/2}\left(2^{1-1/2^k} n^{1/2^k} + 8\right)^{1/2} + 2 \\ &< 2^{1/2}\left(2^{1/2-1/2^{k+1}} n^{1/2^{k+1}} + 8^{1/2}\right) + 4 \\ &= 2^{1-1/2^{k+1}} n^{1/2^{k+1}} + 8. \end{aligned}$$

Now we can translate our upper bound for $g^k(n)$ into an upper bound for $P(n)$. Indeed, by (3) and (4) it follows that

$$\begin{aligned} P(n) &\le m_1 + m_2 + m_3 + \cdots \\ &< (2^{1/2}n^{1/2} + 2) + (2^{3/4}n^{1/4} + 8) + (2^{7/8}n^{1/8} + 8) + \cdots \\ &< 2^{1/2}n^{1/2} + (2n^{1/4} + 8)(\log_2 \log_2 n - 1) + 68, \end{aligned}$$

because there are at most $\log_2 \log_2 n$ terms with $n^{1/2^k} \ge 2$ and the rest by (4) are less than 12, with sum at most 66. Since $2n^{1/4} + 8 \ge 10$,

$$P(n) < \sqrt{2}\, n^{1/2} + (2n^{1/4} + 8) \log_2 \log_2 n + 58,$$

as desired. ■

Volume and perimeter within other sets It is interesting to let another ordered set X, other than the nonnegative integers, play the role of $\mathbb{N}_0$. For example, for the set of harmonic numbers $X = \{1/1, 1/2, 1/3, \ldots\}$ one may attain arbitrary positive volume (via infinitely many terms). We know almost nothing about the minimum perimeter.

Roger Bolton proposed including negative numbers, say all the integers in order:

$$X = \mathbb{Z} := \{\ldots, -1, 0, 1, 2, \ldots\}. \tag{5}$$

In this case the perimeter should be defined as a sum of the absolute values. Now certain types of holes in the previous solution can be filled to reduce the perimeter. First, if the smallest term in the previous solution is $k > 1$, then adding the terms $-(k-1), \ldots, 0, \ldots, k-1$ will leave the volume unaffected and reduce the perimeter by 1. For instance, the minimizer for volume 9 can be improved from $\{2, 3, 4\}$ with perimeter 6 to $\{-1, 0, 1, 2, 3, 4\}$ with perimeter 5. Second, if a nonnegative sequence includes a passage like $\ldots, a-1, a+1, \ldots,$ then including $\{-a, a\}$ will leave the volume untouched while reducing the perimeter by a. For instance, the minimizer for volume 19 can be improved from $\{0, 1, 3, 4, 5, 6\}$ with perimeter 10 to $\{-2, 0, 1, 2, 3, 4, 5, 6\}$ with perimeter 8.

Often in number theory a related problem allowing differences as well as sums is significantly easier to solve than the original allowing just sums. (One famous example is Waring's Problem and the Easier Waring's Problem. Waring's Problem states that for each k there is an $s(k)$ such that all positive integers are a sum of at most $s(k)$ kth powers. The Easier Waring's Problem allows differences as well as sums. For example, 7 is not the sum of three squares, but $7 = 3^2 - 1^2 - 1^2$. It can be proved in half a page [**2**, p. 102].)

Similarly, we can study the Easier Isoperimetric Sequence Problem, where we allow pluses and minuses when summing the elements of the subset S to obtain the prescribed volume n. In this case it is significantly easier to analyze the fluctuations about $\sqrt{2}\, n^{1/2}$. Let $\mathrm{EP}(n)$ denote the minimum perimeter (with all contributions to perimeter still positive) for our related problem, which is obviously at most $P(n)$. The lower bound is still $\sqrt{2}\, n^{1/2} - 1/2$, but now we can remove the $\log_2 \log_2 n$ factor in the upper bound, as it is easy to show that $\mathrm{EP}(n) < \sqrt{2}\, n^{1/2} + 4$. Indeed, choose the smallest k such that $0 + 1 + 2 + \cdots + k = k(k+1)/2 \ge n$; then $k < \sqrt{2}\, n^{1/2} + 1$. To obtain a sum of exactly n, take a minus sign on one term, changing the sum by an even integer to n or $n+1$, and in the latter case drop the 1, adding 2 to the perimeter. In the exceptional case when $k(k+1)/2 = n + 3$, drop the 1 and the 2, adding 3 to the perimeter. In any case, $\mathrm{EP}(n) < \sqrt{2}\, n^{1/2} + 4$, and thus the fluctuations about $\sqrt{2}\, n^{1/2}$ cannot be larger than 5.

We end with some questions for further research. What does the minimum perimeter function $P(n)$ say about the number theoretical properties of a set X? What are some interesting examples? What is the true scale of fluctuations of $P(n)$ about $\sqrt{2}\,n^{1/2}$ when X is the nonnegative integers? Are the fluctuations frequently as large as $\log_2 \log_2 n$? In other words, what can we say about $P(n) - \sqrt{2}\,n^{1/2}$? Is there some $h(n)$ so that $(P(n) - \sqrt{2}\,n^{1/2})/h(n)$ has a nice limiting distribution as $n \to \infty$?

Acknowledgment We thank our colleagues for comments. Miller and Morgan acknowledge partial support from the National Science Foundation.

REFERENCES

1. Frank Morgan, Isoperimetric sequences, http://blogs.williams.edu/Morgan/, 2008.
2. Melvyn B. Nathanson, *Additive Number Theory: The Classical Bases*, Graduate Texts in Mathematics **164**, Springer-Verlag, New York, 1996.

Summary The celebrated isoperimetric theorem says that the circle provides the least-perimeter way to enclose a given area. In this note we discuss a generalization which moves the isoperimetric problem from the world of geometry to number theory and combinatorics. We show that the classical isoperimetric relationship between perimeter P and area A, namely $P = cA^{1/2}$, holds asymptotically in the set of nonnegative integers.

Choosing Rarity: An Exercise in Stopping Times

EWA KUBICKA
GRZEGORZ KUBICKI
University of Louisville
Louisville, Kentucky 40292
gkubicki@louisville.edu

Let's play a computer game. You will see on your computer screen thirteen envelopes in one of two colors: red or blue. The envelopes will appear one by one in random order. Each envelope will be visible for 10 seconds and then will disappear forever. You know that there will be five envelopes of one color and eight of the other, and each of the five envelopes contains \$100, while the eight envelopes of the other color are empty. You can select only one envelope, by clicking on it while it appears on the screen. Once you click on an envelope, it opens to reveal either \$100 or the taunt "You lose!" and the game is over. Of course, you don't know which color is the winning, or "rare," color until you make a choice. How should you proceed to maximize your chances of getting a prize? Does your probability of winning decrease if the number of prizes is smaller?

We describe the optimal strategy for the general case with m envelopes of one color and n of the other, and give the probability of winning. In fact, you will see that, playing with five winning envelopes out of thirteen, the probability of success is 0.902 if you play optimally. However, with only three lucky envelopes, the optimal strategy allows you to win with a probability of 0.965.

Math. Mag. **84** (2011) 42–48. doi:10.4169/math.mag.84.1.042.

The situation we have just described is an example of an optimal stopping time problem and is a variation of the famous secretary problem. In the classic secretary problem, we have n candidates for a secretary position. These candidates are linearly ordered, or ranked, from the best (ranked 1) to the worst (ranked n). They are coming for interviews in some random order and after t interviews, $t \leq n$, we can determine the relative ranking of only those t candidates. We would like to stop the interview process, but can only offer a job to the most recently interviewed candidate. We want to maximize the probability of hiring the best secretary.

The classic secretary problem appeared in the late 1950s and its solution was presented in a paper by Lindley [**6**] in 1961. You can read about the rich history of the problem and its many generalizations in an article by Thomas Ferguson [**2**].

The optimal strategy tells us to interview and reject $t-1$ candidates, where t is the smallest number satisfying the inequality

$$\frac{1}{t} + \frac{1}{t+1} + \cdots + \frac{1}{n-1} \leq 1$$

and then stop on the first one that is better than all candidates seen so far. This means that we have to wait throughout the initial n/e (roughly) interviews before making a stopping decision. The probability of success is approximately $1/e$.

For example, for six candidates ($n = 6$) and the permutation $\pi = (4, 6, 2, 1, 3, 5)$, the full history of interviews is given in FIGURE 1. At each time the candidates with high values are placed below the candidates with low values. In this sample permutation π the fourth best candidate appears first, followed by the sixth best (the worst) candidate, and so on. The optimal strategy tells us to stop at time $t = 3$ since $\frac{1}{3} + \frac{1}{4} + \frac{1}{5} = \frac{47}{60}$ and $\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} > 1$. Unfortunately, the third candidate, who is best so far, is not the best one, but the second best.

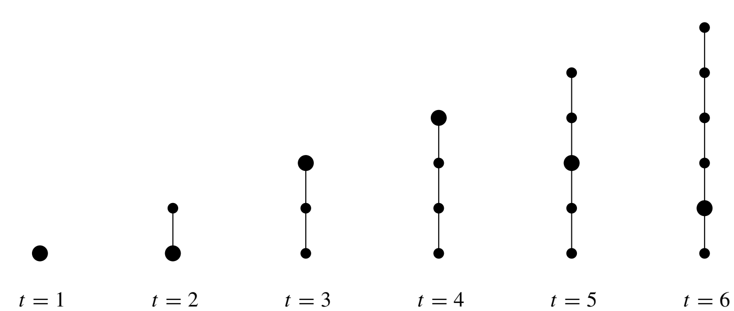


Figure 1 History of interviews for the permutation $\pi = (4, 6, 2, 1, 3, 5)$, meaning that the fourth best candidate appears first, and so on. Each large bullet represents the latest candidate at time t.

Game details and states In our computer game, we have two types of envelopes, the empty ones forming the set M and the winning envelopes forming the set N. We assume that $|M| = m > |N| = n$ and call the game $G(m, n)$. Since the elements of M, as well as N, are indistinguishable, what we observe is a binary sequence whose elements are the names of the colors, say a and b.

Therefore, for any pair of positive integers m and n, with $m > n$, the sample space $U_{m,n}$ is the set of all sequences of length $m + n$ over the binary alphabet $\{a, b\}$ having either n as and m bs or n bs and m as. Every sequence in $U_{m,n}$ is equally probable.

During the game, for a given sequence $\bar{\mathbf{u}}$ (unknown to us), longer and longer initial segments of $\bar{\mathbf{u}}$ are revealed. We develop a vocabulary of *states* to identify features of the initial segment of $\bar{\mathbf{u}}$ at the current time k. Before we define these formally, we give an example: If $\bar{\mathbf{u}} = (aaabbbabbaabaa)$, then at time $k = 5$ we see the initial segment $aaabb$ of $\bar{\mathbf{u}}$; we write $\bar{\mathbf{u}} \in S(3, 2)$ to mean that the *current character* b occurs twice and the other character a occurs three times. At time $k = 6$, we see the initial segment $aaabba$ of $\bar{\mathbf{u}}$ and write $\bar{\mathbf{u}} \in S(2, 4)$. Now the two bs cause us to write 2 in the first slot, because b is no longer the current character.

For $s \geq 0, t \geq 1$, $\min(s, t) \leq n$, and $\max(s, t) \leq m$, define the state $S(s, t)$ as follows: For any sequence $\bar{\mathbf{u}}$ from $U_{m,n}$, $\bar{\mathbf{u}} \in S(s, t)$ if among $s + t$ initial characters of $\bar{\mathbf{u}}$ there are exactly t copies of the character occurring in the position $s + t$ and s copies of the opposite character. Let $\mathfrak{S}_{m,n}$ be the set of all states $S(s, t)$ with $s \geq 0, t \geq 1$, $\min(s, t) \leq n$, and $\max(s, t) \leq m$.

As an illustration, consider the game $G(8, 5)$ and assume that a randomly selected sequence is $\bar{\mathbf{u}} = (aaabbabbaabaa)$. In TABLE 1 we mark consecutive states for this sequence as the game progresses.

TABLE 1: Consecutive states for the sequence ($aaabbabbaabaa$)

s \ t	1	2	3	4	5	6	7	8
0	1	2	3					
1								
2				6				
3	4	5						
4			7	8				
5				9				
6				10	11			
7					12			
8					13			

Notice that many sequences could lead to the same state; for example, all sequences whose initial segments of length five are $aaabb$, $babba$, or $baaab$ belong to the state $S(3, 2)$.

Suppose that after time $s + t$ we see an $\{a, b\}$-sequence $\bar{\mathbf{u}}$ belonging to the state $S(s, t)$ and we stop at that time. If $\bar{\mathbf{u}}(s + t)$ is the $(s + t)$ element of that sequence (a or b), then the conditional probability of winning given that we are currently in the state $S(s, t)$ turns out to be

$$P[\bar{\mathbf{u}}(s+t) \in N \mid \bar{\mathbf{u}} \in S(s,t)] = \frac{\binom{n}{t}\binom{m}{s}}{\binom{n}{s}\binom{m}{t} + \binom{n}{t}\binom{m}{s}}. \tag{1}$$

To see why, notice that the expression in the numerator counts all the winning sequences in the state $S(s, t)$, while the denominator counts all the sequences in the state $S(s, t)$, both winning and losing. Recall again that when evaluating the conditional probability evaluated in (1), we do not know the whole sequence $\bar{\mathbf{u}}$, but only its initial segment of length $s + t$. We do not know which character a or b is the winning character. This will be determined by the remaining elements of $\bar{\mathbf{u}}$. Let's notice that this probability does not depend on the permutation of the elements of the initial segment, but only on the number of occurrences of the current character.

Intuitively, a stopping time should tell us when to stop the game, based only on the portion of the sequence through the kth entry. More formally, a *stopping time* for the set of sequences $U_{m,n}$ is a function τ on $U_{m,n}$ whose range is $\{1, 2, 3, \ldots, m+n\}$, such that if for some sequence $\bar{\mathbf{u}}$, $\tau(\bar{\mathbf{u}}) = k$, then $\tau(\bar{\mathbf{v}}) = k$ for any other sequence $\bar{\mathbf{v}}$ whose initial segments until time k agree with the initial segments of $\bar{\mathbf{u}}$.

In our case, however, because of the observation about probabilities we made earlier, the value $\tau(\bar{\mathbf{u}}) = k$ depends only on the number t of occurrences of the character $\bar{\mathbf{u}}(k)$ through time k. In other words, the stopping time depends only on the current state $S(k-t, t)$ and not the whole history of states determined by shorter initial segments of $\bar{\mathbf{u}}$. Therefore, a stopping time τ can be defined by specifying a subset T of the set of states $\mathfrak{S}_{m,n}$. We should stop the first time we enter a state in T. If, for a given $\bar{\mathbf{u}}$, this state is $S(s, t)$, then $\tau(\bar{\mathbf{u}}) = k = s + t$.

Optimal stopping time We are ready to formulate the theorem describing an optimal stopping time for the game $G(m, n)$. This optimal stopping time tells us to stop at the first moment when we see at least n envelopes in each color. It is obvious that seeing more than n envelopes of a single color guarantees that we will win by picking the other color. It is also clear that if we have seen n envelopes of one color and we are looking at the nth envelope of the other color, then we must stop. We could still win with the probability $1/2$, but if we wait, we will certainly lose.

THEOREM 1. *An optimal stopping time τ_0 is the stopping time determined by the subset $T_0 = \{S(s, n) \mid s \geq n\}$. For this stopping time τ_0, the probability of winning equals*

$$P[\bar{\mathbf{u}}(\tau_0) \in N] = 1 - \frac{n(n+1)(n+2)\cdots(2n-1)}{(m+1)(m+2)(m+3)\cdots(m+n)}. \tag{2}$$

Proof. We establish four facts:

(a) $P[\bar{\mathbf{u}}(\tau_0) \in N \mid \bar{\mathbf{u}} \in S(s, n)] = \begin{cases} 1, & \text{if } s > n; \\ \frac{1}{2}, & \text{if } s = n. \end{cases}$

(b) It is never beneficial to stop in a state $S(s, t)$ with $t > s$ and $s < n$.

(c) The probability of winning is as in (2).

(d) If $t \leq s$ and τ is a stopping time that tells us to stop in a state $S(s, t)$ with $t < n$, then $P[\bar{\mathbf{u}}(\tau) \in N \mid \bar{\mathbf{u}} \in S(s, n)] \leq P[\bar{\mathbf{u}}(\tau_0) \in N \mid \bar{\mathbf{u}} \in S(s, n)]$.

To justify (a), let us notice that putting $s = t = n$ in (1) gives $P[\bar{\mathbf{u}}(2n) \in N \mid \bar{\mathbf{u}} \in S(n, n)] = \frac{1}{2}$; if $s > n$, then the current character is certainly the winning character.

To justify (b), suppose that we stop at the state $S(s, t)$ with $t > s$ and $s < n$. Then

$$P[\bar{\mathbf{u}}(s+t) \in N \mid \bar{\mathbf{u}} \in S(s, t)] = \frac{\binom{n}{t}\binom{m}{s}}{\binom{n}{s}\binom{m}{t} + \binom{n}{t}\binom{m}{s}} = \frac{1}{1 + \frac{(m-s)!\,(n-t)!}{(m-t)!\,(n-s)!}} < \frac{1}{2},$$

since

$$\frac{(m-s)!\,(n-t)!}{(m-t)!\,(n-s)!} = \frac{(m-t+1)(m-t+2)\cdots(m-s)}{(n-t+1)(n-t+2)\cdots(n-s)} > 1.$$

However, if we use τ_0 and do not stop at $S(s, t)$ we will reach a state in T_0 which gives us, after stopping there, the probability of winning of at least $\frac{1}{2}$.

To justify (c), we notice that the only sequences for which the stopping time τ_0 leads to a loss are those where all n elements from the winning set N arrive before the nth element of the larger set M. Equivalently, all elements of N arrive in the first $2n-1$ positions. Therefore,

$$P[\bar{\mathbf{u}}(\tau_0) \in N] = 1 - \frac{\binom{2n-1}{n}}{\binom{m+n}{n}} = 1 - \frac{n(n+1)(n+2)\cdots(2n-1)}{(m+1)(m+2)(m+3)\cdots(m+n)}.$$

To justify (d), we will prove that the probabilities of failure when using τ_0 and τ satisfy the reverse inequality. We have

$$P[\bar{\mathbf{u}}(\tau) \in M \mid \bar{\mathbf{u}} \in S(s,t)] = \frac{\binom{n}{s}\binom{m}{t}}{\binom{n}{s}\binom{m}{t} + \binom{n}{t}\binom{m}{s}}.$$

On the other hand, if we do not stop at $S(s,t)$ but wait and stop later according to the strategy τ_0, then

$$P[\bar{\mathbf{u}}(\tau_0) \in M \mid \bar{\mathbf{u}} \in S(s,t)]$$
$$= \frac{\binom{n}{s}\binom{m}{t}}{\binom{n}{s}\binom{m}{t} + \binom{n}{t}\binom{m}{s}} \cdot \frac{\binom{2n-s-t-1}{n-s}}{\binom{m+n-s-t}{n-s}} + \frac{\binom{n}{t}\binom{m}{s}}{\binom{n}{s}\binom{m}{t} + \binom{n}{t}\binom{m}{s}} \cdot \frac{\binom{2n-s-t-1}{n-t}}{\binom{m+n-s-t}{n-t}}.$$

The first term gives the probability that among $s+t$ elements leading to the state $S(s,t)$, we had t elements from M, s from N, and the remaining $n-s$ elements of N will all come before the time $2n-s-t$. Similarly, the second term gives the probability that t elements leading to the state $S(s,t)$ were from N, s were from M, and the remaining $n-t$ elements of N will all come before the time $2n-s-t$. Assume first that $t \le s < n$. Then

$$\binom{m+n-s-t}{n-s} > \binom{2n-s-t}{n-s} \quad \text{or} \quad \frac{(m+n-s-t)!}{(m-t)!\,(n-s)!} > \frac{(2n-s-t)!}{(n-s)!\,(n-t)!}.$$

Equivalently,

$$\frac{(m+n-s-t)!}{(m-t)!\,(n-s)!} > \frac{(2n-s-t-1)!\,(n-t+n-s)}{(n-s)!\,(n-t)!}$$

or

$$\frac{(m+n-s-t)!}{(m-t)!\,(n-s)!} > \frac{(2n-s-t-1)!}{(n-t-1)!\,(n-s)!} + \frac{(2n-s-t-1)!}{(n-s-1)!\,(n-t)!},$$

which implies, after dividing both sides by $(m+n-s-t)!$ and multiplying by $\frac{m!\,n!}{t!\,s!}$, the inequality

$$\binom{n}{s}\binom{m}{t} > \binom{n}{s}\binom{m}{t} \cdot \frac{\binom{2n-s-t-1}{n-s}}{\binom{m+n-s-t}{n-s}} + \binom{n}{t}\binom{m}{s} \cdot \frac{\binom{2n-s-t-1}{n-t}}{\binom{m+n-s-t}{n-t}}.$$

Dividing both sides by $\binom{n}{s}\binom{m}{t} + \binom{n}{t}\binom{m}{s}$ gives

$$P[\bar{\mathbf{u}}(\tau) \in M \mid \bar{\mathbf{u}} \in S(s,t)] > P[\bar{\mathbf{u}}(\tau_0) \in M \mid \bar{\mathbf{u}} \in S(s,t)], \quad \text{as desired.}$$

If $t < n$ but $s = n$, then easy calculations show that

$$P[\bar{\mathbf{u}}(\tau) \in M \mid \bar{\mathbf{u}} \in S(s,t)] = P[\bar{\mathbf{u}}(\tau_0) \in M \mid \bar{\mathbf{u}} \in S(s,t)]. \qquad \blacksquare$$

To illustrate the theorem let us evaluate relevant probabilities for the game $G(8,5)$. According to the last theorem, using the stopping time τ_0, we win with the probability $P[\bar{\mathbf{u}}(\tau_0) \in N] = 1 - \frac{5\cdot6\cdot7\cdot8\cdot9}{9\cdot10\cdot11\cdot12\cdot13} = \frac{129}{143} \simeq 0.902$. For example, for the sequence $\bar{\mathbf{u}} = (aaabbabbaabaa)$ we have $\tau_0(\bar{\mathbf{u}}) = 11$ and we win on this sequence. However, for the sequence $\bar{\mathbf{u}} = (aaabbabbbaaaa)$ we have $\tau_0(\bar{\mathbf{u}}) = 10$ and we lose on this sequence. For a random $\bar{\mathbf{u}}$ in which only ten initial characters are known, for example $\bar{\mathbf{u}} = (aaabbabbba\_\_\_)$, $\tau_0(\bar{\mathbf{u}}) = 10$ and $P[\bar{\mathbf{u}}(\tau_0) \in N \mid \bar{\mathbf{u}} \in S(5,5)] = \frac{1}{2}$. Notice that when we see only seven characters of $\bar{\mathbf{u}}$, namely $\bar{\mathbf{u}} = (aaabbab\_\_\_\_\_\_)$, and we stop at the state $S(4,3)$, not according to the optimal strategy τ_0, then $P[\bar{\mathbf{u}}(7) \in N \mid \bar{\mathbf{u}} \in S(4,3)] = 1 - \frac{56}{196} = \frac{5}{7} \simeq 0.714$. Although this probability is higher than 0.5, it is still possible from the state $S(4,3)$ to reach other states, not $S(5,5)$, where τ_0 assures us of winning with probability 1.

The optimal stopping time τ_0 described in the theorem is not unique. For example, the stopping time τ_1 associated with the set of states $T_1 = \{S(s,t) \mid s \geq n\}$ is optimal as well. In fact, any stopping time τ associated with a set of states T such that $T_0 \subseteq T \subseteq T_1$ is optimal. Not only does such τ have the same probability of winning as τ_0, but τ and τ_0 are either both winning or both losing identically on every sequence in the sample space. They all fail on sequences in which all n winning characters appear before the nth losing character. For the game $G(8,5)$, the set T_0 consists of the four last states of column 5 in TABLE 1, but the set T_1 consists of twenty-three states in the last four rows of TABLE 1.

For $m + n = 13$, the probability of winning while using τ_0 changes with n. TABLE 2 gives these probabilities for the full range of n, $1 \leq n \leq 6$. The highest probability of winning occurs for $n = 3$.

TABLE 2: Probabilities of winning for all games with $m + n = 13$

Game	Probability of winning
$G(7,6)$	$\frac{19}{26} \simeq 0.731$
$G(8,5)$	$\frac{129}{143} \simeq 0.902$
$G(9,4)$	$\frac{136}{143} \simeq 0.951$
$G(10,3)$	$\frac{138}{143} \simeq 0.965$
$G(11,2)$	$\frac{25}{26} \simeq 0.962$
$G(12,1)$	$\frac{12}{13} \simeq 0.923$

Exercises for the reader

1. Show that the probabilities of winning while using τ_0 for the games $G(n+1,n)$ approach $3/4$ as $n \to \infty$.
2. Show that the probabilities of winning while using τ_0 for the games $G(m,1)$ approach 1 as $m \to \infty$.
3. Prove that for a fixed value of $m + n$ the sequence $a_n = P[\bar{\mathbf{u}}(\tau_0) \in N]$, where $|N| = n$, is unimodal for $1 \leq n < m$, more precisely

$$a_n \leq a_{n+1} \quad \text{for } 1 \leq n \leq \frac{m-2}{4}$$

and the sequence becomes decreasing for larger values of n.

Final comments More comprehensive studies of the secretary problem appear in a book by Berezovsky and Gnedin [**1**] or in a paper by Ferguson [**2**].

Other variants of the secretary problem were studied later for partial orders (complete binary tree [**7**], general partial order [**3, 8**]), for graphs and digraphs [**5**], and threshold stopping times [**4**].

The game described in this paper could be generalized to allow p different types of elements, with $p > 2$. If the objective is to "choose rarity" by stopping on an element in the smallest set, an optimal strategy seems to be analogous to the one described in the paper, but finding probabilities of winning is more challenging.

REFERENCES

1. B. A. Berezovsky and A. V. Gnedin, *The Best Choice Problem*, Nauka, Moscow, 1984.
2. T. Ferguson, Who solved the secretary problem? *Statist. Sci.* **4** (1989) 282–296. doi:10.1214/ss/1177012493
3. N. Georgiou, M. Kuchta, M. Morayne, and J. Niemiec, On a universal best choice algorithm for partially ordered sets, *Random Structures Algorithms* **32**(3) (2008) 263–273. doi:10.1002/rsa.20192
4. A. V. Gnedin, Multicriteria extensions of the best choice problem: Sequential selection without linear order, pp. 153–172 in F. T. Bruss, T. S. Ferguson, and S. M. Samuels (eds.), *Strategies for Sequential Search and Selection in Real Time*, American Mathematical Society, Providence, RI, 1992.
5. G. Kubicki and M. Morayne, Graph-theoretic generalization of the secretary problem; the directed path case, *SIAM J. Discrete Math.* **19**(3) (2005) 622–632. doi:10.1137/S0895480104440596
6. D. V. Lindley, Dynamic programming and decision theory, *Appl. Statist.* **10** (1961) 39–51. doi:10.2307/2985407
7. M. Morayne, Partial-order analogue of the secretary problem: The binary tree case, *Discrete Math.* **184** (1998) 165–181. doi:10.1016/S0012-365X(97)00091-5
8. J. Preater, The best-choice problem for partially ordered sets, *Oper. Res. Lett.* **25** (1999) 187–190. doi:10.1016/S0167-6377(99)00053-X

Summary You enter an online game, hoping to win a prize. You know that thirteen envelopes of two colors will appear one by one on your screen, and you know that there will be five of one color and eight of the other. Each of the five envelopes, those with the "rare" color, contains \$100, while the others are empty. However, you do not know which color is which, and you can only select one envelope; once you click, the game is over. How should you play to maximize the probability of winning? We provide an optimal strategy for this game and a generalization.

Golomb Rulers

ROGER C. ALPERIN
San Jose State University
San Jose, CA 95192
alperin@math.sjsu.edu

VLADIMIR DROBOT
San Jose State University
San Jose, CA 95192
drobot@pacbell.net

The Math Factor podcast posed the problem of finding the smallest number of inch marks on a 12 inch ruler so that one could still measure any integer length from 1 to 12 as a distance between marks. One needs only four additional marks besides 0 and 12; for example 1, 4, 7, 10 works, as you can check. This entertaining problem led to others during the next few minutes (you can listen at http://mathfactor.uark.edu/2005/10) and inspired us to look for generalizations. After several false starts

Math. Mag. **84** (2011) 48–55. doi:10.4169/math.mag.84.1.048. © Mathematical Association of America

and numerous literature searches we uncovered the theory of *Golomb* and *minimal spanning rulers*, a generalization to the natural numbers and relations of an unsolved conjecture of Erdős and Turan.

We analyze our first problem, which led us to Golomb rulers. A property of the ruler of length 6 with marks at 0, 1, 4, 6 is that each of the lengths 1, 2, 3, 4, 5, and 6 can be measured and it can be done in only one way. *Can one choose marks on a ruler of length 12 so that each length from 1 to 12 can be measured in only one way?*

Golomb rulers are sets of integers (marks) with the property that if a distance can be measured using these marks then it can be done in a unique way.

DEFINITION 1. A set $\mathcal{G}$ of integers

$$a_1 < a_2 < \cdots < a_{p-1} < a_p$$

is called a *Golomb ruler* if for every two distinct pairs of these integers, say $a_i < a_j$ and $a_m < a_n$, we have $a_j - a_i \neq a_n - a_m$.

The *size* of $\mathcal{G}$ is defined to be p (the number of marks in $\mathcal{G}$) and is denoted $\#\mathcal{G}$. The *length* of $\mathcal{G}$ is defined to be $a_p - a_1$ (the largest distance that can be measured using the marks from $\mathcal{G}$).

It is clear that we can translate these sets: if $\mathcal{G} = \{a_1, a_2, \ldots, a_p\}$ is a Golomb ruler then so is $\{a_1 + b, a_2 + b, \ldots, a_p + b\}$. This makes the choice of a_1 immaterial, so it will usually be taken to be 0. It is also clear that if $\mathcal{G} = \{a_1, a_2, \ldots, a_p\}$ is a Golomb ruler, then so is the reflection of $\mathcal{G}$ around the midpoint $(a_1 + a_p)/2$. For example, $\{0, 1, 4, 6\}$ is a Golomb ruler, as is $\{0, 2, 5, 6\}$ obtained by reflecting the first ruler around the point 3. To simplify the statements of some of the theorems, a set $\{a_1\}$ consisting of a single point is considered to be a Golomb ruler.

Golomb rulers have numerous applications. The best known is an application to radio astronomy. Radio telescopes (antennas) are placed in a linear array. For each pair of these antennas, the received signals are subtracted from each other and an inference can be then made as to the location of the source. These inferences can be made much more accurate if all the distances between the antennas are multiples of the same common length, and many such pairs with distinct distances between them are available and can be utilized. The problem maximizing the number of distinct distances between the pairs, while minimizing the number of the antennas and the length of the array, was first considered by Solomon W. Golomb [**1, 2, 8, 10**].

Other applications include assignments of channels in radio communications, X-ray crystallography, and self-orthogonal codes. Rankin [**12**] gives more information about these applications. There is also a wealth of information in various writings by Martin Gardner [**5, 6, 7**].

The Golomb ruler $\{0, 1, 4, 6\}$ has the additional property that every integral distance between 1 and 6 can be measured. We call such a ruler *perfect*.

DEFINITION 2. A Golomb ruler $\mathcal{G}$ of length N is called *perfect* if every integer d, $1 \leq d \leq N$, can be expressed as $d = a - a'$, for some $a, a' \in \mathcal{G}$.

Since $\mathcal{G}$ is a Golomb ruler, the representation of each d is unique. Unfortunately, there are very few perfect Golomb rulers.

THEOREM 1. (GOLOMB) *Together with their translations and reflections around the midpoint, the only perfect Golomb rulers are* $\{0\}$, $\{0, 1\}$, $\{0, 1, 3\}$, *and* $\{0, 1, 4, 6\}$.

This theorem was proved by Golomb, but apparently he never published it. There are several places where the proof appears [**4, 12**], but they are not very easily accessible, so we present here a slight modification of the original argument.

Proof. If $\mathcal{G}$ is a perfect Golomb ruler of size p and length N, then, since there are N distances to be measured and the number of distinct pairs of these points is $\binom{p}{2}$, we must have $N = \frac{1}{2}p(p-1)$, so N is a triangular number. The triangular numbers below 10 are 0, 1, 3, 6 corresponding to the rulers listed in the theorem.

Assume then that $\mathcal{G}$ is a perfect Golomb ruler of length $N > 9$ and seek a contradiction. Without loss of generality we may assume that a_1, the smallest number in $\mathcal{G}$, is equal to 0 and so the largest number is $a_p = N$. By hypothesis, every number $1 \le d \le N$ is uniquely realizable as a difference of two marks in $\mathcal{G}$. Since N is realizable, 0 and N must belong to $\mathcal{G}$. Since $N-1$ is realizable, either 1 or $N-1$ belongs to $\mathcal{G}$. By reflecting $\mathcal{G}$ around $N/2$, we may assume that $1 \in \mathcal{G}$. Next, since $N > 3$, $N-2$ must be realized. Since $N-2 > 1$, $\mathcal{G}$ must contain another point.

The possible pairs realizing $N-2$ are $\{2, N\}$, $\{1, N-1\}$, $\{0, N-2\}$. The first two produce duplications: $1-0 = 2-1$ and $1-0 = N-(N-1)$. The third is the only possibility, so $\mathcal{G}$ contains $N-2$ as well as $0, 1, N$. The realized distances are $1, 2, N-3, N-2, N-1$, and N.

Since $N-4 \notin \{1, 2\}$ we need one of the pairs $\{0, N-4\}$, $\{1, N-3\}$, $\{2, N-2\}$, $\{3, N-1\}$, $\{4, N\}$ to realize $N-4$. All but the last case produce distances already realized: $(N-2)-(N-4) = N-(N-2)$, $1-0 = (N-2)-(N-3)$, $2-0 = N-(N-2)$, $1-0 = N-(N-1)$.

The last case passes inspection, so $\mathcal{G}$ contains $0, 1, 4, N-2$, and N. The distances that can be realized by $\mathcal{G}$ are $1, 2, 3, 4, N-6, N-4, N-3, N-2, N-1$, and N.

Finally, consider the distance $N-5$. Since $N-5 \notin \{1, 2, 3, 4\}$ and $N > 9$ this distance has not been realized. The possible pairs for realizing the distance $N-5$ are $\{0, N-5\}$, $\{1, N-4\}$, $\{2, N-3\}$, $\{3, N-2\}$, $\{4, N-1\}$, $\{5, N\}$. The reader may easily check that each of these cases leads to a duplication. This contradiction shows that $N < 9$ and the constructions above give the perfect rulers asserted by the theorem. ∎

Since perfect Golomb rulers essentially do not exist, we seek "almost perfect" rulers. Roughly speaking, given a length N, we try to place as many points as possible in the interval $[0, N]$ so that the resulting set forms a Golomb ruler. Alternatively, given the size p of the ruler (the number of marks), we try to construct a Golomb ruler of shortest possible length N with p points. Such rulers are called optimal.

DEFINITION 3. For every positive integer p, let $G(p)$ be the shortest possible length of a Golomb ruler with p marks.

A Golomb ruler with p marks is called *optimal* if its length is $G(p)$. Dimitromanolakis [**4**] discusses optimal Golomb rulers in detail. For example, $G(6) = 17$, and there are 4 optimal rulers of size 6 and length 17: $\{0, 1, 4, 10, 12, 17\}$, $\{0, 1, 4, 10, 15, 17\}$, $\{0, 1, 8, 11, 13, 17\}$, and $\{0, 1, 8, 12, 14, 17\}$.

Computer searches give the largest known value of $G(p)$. The current record is $G(26) = 492$ and the corresponding optimal Golomb ruler has marks

0 1 33 83 104 110 124 163 185 200 203 249 251 258

314 318 343 356 386 430 440 456 464 475 487 492.

The search took several years and this ruler is not even known to be unique [**13, 14**]. *Wikipedia* is a good place to look for the latest record values of $G(p)$.

Given a Golomb ruler with p marks, there are $\binom{p}{2} \sim \frac{1}{2}p^2$ distinct distances one can measure with this ruler. Thus, one expects $G(p)$ to be roughly at least $\frac{1}{2}p^2$. It is a conjecture, with strong empirical evidence, that $G(p) < p^2$; but it is only a conjecture.

Golomb rulers also have a close connection with additive number theory. It is completely outside the scope of this paper to discuss this connection in any depth, and we only state some facts and invite the reader to investigate further.

DEFINITION 4. A subset $\mathcal{B}$ of integers contained in $[1, N]$ is called a B_2 *basis*, if for any two distinct pairs of integers from $\mathcal{B}$, say a, a' and b, b' we have

$$a + a' \neq b + b'.$$

There is an old conjecture of Erdős and Turan which states that a B_2 basis with $\lfloor\sqrt{N}\rfloor$ elements can be constructed in $[1, N]$ for any N. This is very closely related to the conjecture that $G(p) < p^2$. Halberstam and Roth [**9**] give a comprehensive discussion of additive number theory and the connection with Golomb rulers.

Perfect rulers on $\mathbb{N}$ In this section we study *infinite rulers*. These are sets $\mathcal{G}$ of nonnegative integers, such that any positive integer d is realized as a distance between some two elements of $\mathcal{G}$. If we require, in addition, the representation to be unique, we may speak of infinite perfect Golomb rulers.

DEFINITION 5. A subset $\mathcal{G}$ of the set $\mathbb{N}$ of natural numbers is called an *infinite perfect Golomb ruler* if

(1) for every positive integer d, there are elements $a, a' \in \mathcal{G}$ so that $d = a - a'$, and
(2) for every such d this representation is unique.

It is not entirely clear that such things exist, but in fact they do and also they can be made arbitrarily *thin* (sparse) depending on the choice of a function φ.

THEOREM 2. *Let* $\varphi : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ *be strictly increasing with* $\varphi(x) \to \infty$ *as* $x \to \infty$. *There is an infinite perfect Golomb ruler* $\mathcal{G} \in \mathbb{N}$ *such that for* $x > x_0 = x_0(\mathcal{G}, \varphi)$

$$\#\{k \mid k \in \mathcal{G},\ k \leq x\} \leq \varphi(x). \tag{1}$$

Proof. We preview the steps: First we choose a rapidly increasing sequence γ_k, $k = 1, 2, \ldots$, and then construct $\mathcal{G}$ by successively adding the points $\{\gamma_k, \gamma_k + k\}$. If a duplication would occur as a result of this addition, then we do not add the pair. Various things have to be proved, for example, that skipping a pair does not result in some integer d not being realized as the difference of two elements of $\mathcal{G}$, etc. The details follow.

Choose a strictly increasing function $\psi(x)$ such that

$$x < \frac{1}{2}\varphi(\psi(x)) \tag{2}$$

and define a sequence $\{\gamma_k\}$ by

$$\begin{aligned} &\gamma_1 = 0 \\ &\gamma_{k+1} > \psi(k+1) + 2(\gamma_k + k) + 1, \text{ for } k \geq 1. \end{aligned} \tag{3}$$

Define $\mathcal{A}_1 = \{\gamma_1, \gamma_1 + 1\}$ which of course equals $\{0, 1\}$, set

$$\mathcal{A}_{k+1} = \begin{cases} \mathcal{A}_k \cup \{\gamma_{k+1}, \gamma_{k+1} + (k+1)\} \\ \quad \text{if this set } \mathcal{A}_k \cup \{\gamma_{k+1}, \gamma_{k+1} + (k+1)\} \text{ has no duplicate distances} \\ \mathcal{A}_k \text{ otherwise,} \end{cases}$$

and let

$$\mathcal{G} = \bigcup_{k=1}^{\infty} \mathcal{A}_k.$$

First we show that the set $\mathcal{G}$ satisfies the density condition (1) in the statement of the Theorem. Let $x > 1$ be given and let k_0 be the largest integer such that $\gamma_{k_0} \le x$. Then

$$\#\{k \mid k \in \mathcal{G}, k \le x\} \le 2k_0$$

because the elements of $\mathcal{G}$ come in pairs: γ_p and $\gamma_p + p$. Now

$$k_0 < \frac{1}{2}\varphi\big(\psi(k_0)\big) < \frac{1}{2}\varphi(\gamma_{k_0}) \le \frac{1}{2}\varphi(x).$$

The first inequality follows from (2) and the second and third follow from (3) and the fact that φ is monotonically increasing. Thus the density claim (1) of the Theorem is true.

By construction, there is no duplication of distances in $\mathcal{G}$. After all, there is no duplication of distances in any of the sets $\mathcal{A}_k$.

It remains to show that every distance d is realized as a difference of two elements of $\mathcal{G}$. It suffices to analyze the pairs not included by our process: When a duplication occurs by inclusion of $\{\gamma_p, \gamma_p + p\}$, then we claim that

$$p = a - a' \text{ where } a, a' \in \mathcal{A}_{p-1} \tag{4}$$

that is, p is already realized as a distance in the set $\mathcal{A}_{p-1}$. This would occur, for example, in the following situation: Let a and a' be two points in a set $\mathcal{A}_q$, for some q such that $a - a' > q$. Then, further along in the process, adding the pair $\{\gamma_{a-a'}, \gamma_{a-a'} + (a - a')\}$ would surely create a duplication. The claim is that this is essentially the only way it could happen. Now, if this claim is true, then either every distance d occurs in $\mathcal{G}$ through the addition of the pair $\{\gamma_d, \gamma_d + d\}$ or d occurs already as a distance in the set $\mathcal{A}_{d-1}$.

We now prove the assertion (4). Suppose that the addition of the pair $\{\gamma_p, \gamma_p + p\}$ to the set $\mathcal{A}_{p-1}$ results in duplications. Because there are no duplications in the set $\mathcal{A}_{p-1}$ these duplications must involve the points from the pair under discussion. It follows, because of (3), that both points of the pair are larger than any of the points in $\mathcal{A}_{p-1}$, and so the possibilities are:

$$\begin{aligned}
&\text{(i)} & (\gamma_p + p) - \gamma_p &= a - a' \\
&\text{(ii)} & (\gamma_p + p) - a &= \gamma_p - a' \\
&\text{(iii)} & (\gamma_p + p) - a &= a' - a'' \\
&\text{(iv)} & \gamma_p - a &= a' - a''
\end{aligned}$$

where the numbers a, a', a'' are elements of the set $\mathcal{A}_{p-1}$. In cases (i) and (ii) then p is a difference of some elements in $\mathcal{A}_{p-1}$, hence (4) holds. The possibilities (iii) and (iv) cannot occur because the largest element of $\mathcal{A}_{p-1}$ is at most $\gamma_{p-1} + (p - 1)$ and from (3) then $\gamma_p > 2(\gamma_{p-1} + p - 1)$. But, if either (iii) or (iv) were true, then either γ_p or $\gamma_p + p$ would be at most twice the largest element of $\mathcal{A}_{p-1}$. Thus our claim (4) is shown and the theorem is proved. ■

Thus, thin infinite perfect Golomb rulers do exist. The construction in Theorem 2 does not give a formula for the nth mark—it just constructs these marks one by one.

It would be interesting to know how *thick* an infinite perfect Golomb can be. In particular is it possible to have

$$\delta_{\mathcal{G}}(x) = \#\{\, k \mid k \in \mathcal{G}, k \le x\} \sim \sqrt{x}\,? \tag{5}$$

By arguments similar to the discussion of finite perfect Golomb rulers, it is easy to see that $\delta_{\mathcal{G}}(x) < \sqrt{2x}$, and (5) is motivated by the Erdős-Turan conjecture about B_2 bases (it does not follow from nor does it imply the conjecture).

Minimal spanning rulers Next we return to rulers of finite length and discuss those that can be used to measure every distance. They differ from Golomb rulers in that there might be a distance that can be measured in two different ways, but we require that every eligible distance can be measured. We call such rulers *spanning*.

DEFINITION 6. Let $\mathcal{S} = \{0 = a_1 < a_2 < \cdots < a_p = N\}$ be a set of integers. We say that $\mathcal{S}$ is a *spanning ruler on* $[0, N]$ if every integer $1 \le d \le N$ can be expressed as $d = a - a'$, with a and $a' \in \mathcal{S}$.

We say that a spanning ruler $\mathcal{M}$ is *minimal* on $[0, N]$, if whenever $\mathcal{M}'$ is a proper subset of $\mathcal{M}$ then the set $\mathcal{M}'$ is not a spanning ruler on $[0, N]$.

Minimal spanning rulers obviously exist. Just start with $\{0, 1, \ldots, N\}$ and remove one point at a time until you can't do it anymore.

However, minimal rulers cannot be very *thin*. If $\mathcal{M}$ is a minimal ruler of length N and $p = \#\mathcal{M}$, then $\binom{p}{2} = \frac{1}{2}p(p-1) \ge N$; so p is roughly at least $\sqrt{2N}$. We now show that we can come fairly close to this lower bound.

THEOREM 3. *For every integer* $N \ge 4$ *there is a minimal spanning ruler* $\mathcal{M}_N \subset [0, N]$ *such that*

$$2\sqrt{N} - 1 \le \#\mathcal{M}_N < 2\sqrt{N} \tag{6}$$

and the equality on the left side holds only when N *is a perfect square.*

Proof. The basic idea of the proof can best be seen by an example of a thin minimal ruler for $N = 100$. The ruler $\mathcal{M}_{100}$ is in this case taken to be

$$\mathcal{M}_{100} = \{0, 1, 2, 3, \ldots, 9, 20, 30, 40, \ldots, 90, 100\}.$$

Notice that the number 10 is *not* included. The number of elements in $\mathcal{M}_{100}$ is 19 which is equal to $2\sqrt{100} - 1$. Every distance $1 \le d \le 100$ is realizable: Represent $d = 10$ as $30 - 20$ and other multiples of 10 are just $d = d - 0$. If $d = q \cdot 10 + j$, $1 \le q, j \le 9$, then $d = (q + 1) \cdot 10 - (10 - j)$.

Finally, if $1 \le d \le 9$ then $d = d - 0$. None of the numbers can be removed. For example, $d = 7$ cannot be removed because then $13 = 20 - 7$ would not be realizable. The number 30 cannot be removed because then $21 = 30 - 9$ would not be realizable. If 10 is included the ruler is not minimal. The actual proof is based on this example although some care must be taken when N is not a perfect square. Here are the details.

By inspection, when $N \in \{5, 6, 7, 8\}$ then the minimal spanning rulers satisfying (6) are, respectively:

$$\{0, 1, 3, 5\}, \{0, 1, 4, 6\}, \{0, 1, 4, 5, 7\}, \{0, 1, 4, 6, 8\}.$$

Incidentally, there are no minimal spanning rulers satisfying the condition (6) for $N \in \{1, 2, 4\}$ and there is one for $N = 3$, namely $\{0, 1, 3\}$.

Let $\xi = \lfloor\sqrt{N}\rfloor$ so that $\xi^2 \le N < (\xi+1)^2 = (\xi+2)\xi + 1$. We assume that $\xi \ge 3$. There are two possibilities:

$$(\alpha)\ N = 0 \bmod \xi \text{ so that } N = K\xi,\ K = \xi,\ \xi+1, \text{ or } \xi+2; \tag{7}$$

$$(\beta)\ N \ne 0 \bmod \xi \text{ so that } N = K\xi + \eta,\ K = \xi \text{ or } \xi+1, \text{ and } 1 \le \eta < \xi.$$

In case (α) take

$$\mathcal{M}_N = \{0, 1, \ldots, \xi - 1, 2\xi, 3\xi, \ldots, K\xi\} \quad (\xi \text{ is not included})$$

with K as in (7).

Every distance $1 \le d \le N$ is realized as the following analysis shows: If $1 \le d \le \xi - 1$ then $d = d - 0$; When $d = \xi$, then $d = 3\xi - 2\xi$ since $2\xi, 3\xi \in \mathcal{M}_N$ for $\xi \ge 3$; For $d = q\xi$, $q > 1$ then $d = q\xi - 0$; Finally if $d = q\xi + \eta$, $1 \le q < K$, $1 \le \eta < \xi$, then $d = (q+1)\xi - (\xi - \eta)$.

Also, we see that none of the marks can be removed: The endpoints 0 and N cannot be deleted because $N = N - 0$. The points $1 \le d < \xi$ cannot be deleted because of the distance $\xi + (\xi - d) = 2\xi - d$. Finally, the points $q\xi$, $2 \le q \le K$, cannot be deleted because of the distance $(q-1)\xi + 1 = q\xi - (\xi - 1)$.

Counting the marks gives $\#\mathcal{M}_N = \xi + K - 1$. Thus, to show (6) we must prove that for $t = 0, 1, 2$ then

$$2\sqrt{\xi(\xi+t)} - 1 \le 2\xi + t - 1 < 2\sqrt{\xi(\xi+t)}$$

with equality holding on the left side only when $t = 0$. This is done in a straightforward manner by squaring each side of the inequality to eliminate radical expressions.

In case (β) take

$$\mathcal{M}_N = \{0, 1, \ldots, \xi - 1, 2\xi, 3\xi, \ldots, K\xi, K\xi + \eta\} \quad (\xi \text{ not included})$$

where K, η are as specified in (7). Again, all the distances $1 \le d \le N = K\xi + \eta$ can be realized: If $1 \le d \le K\xi$ then the argument is the same as in case (α); If $d = K\xi + \delta$, $1 \le \delta \le \eta$, $d = (K\xi + \eta) - (\eta - \delta)$.

None of the marks can be removed: The endpoints cannot be removed because of the distance $N = N - 0$. The points $\xi - c$, $1 \le c < \xi$, $c \ne \eta$ cannot be removed because of the distance $\xi + c = 2\xi - (\xi - c)$.

The point $\xi - \eta$ cannot be removed because of the distance

$$(K-1)\xi + \eta = K\xi - (\xi - \eta). \tag{8}$$

However also $(K-1)\xi + \eta = (K\xi + \eta) - \xi$ and we see that (8) is the only way to realize the distance $(K-1)\xi + \eta$ since $\xi \notin \mathcal{M}_N$.

The points $2\xi, 3\xi, \ldots, K\xi$ cannot be removed for the following reason: Let τ be such that $\tau \ne \eta$, $1 \le \tau < \xi$. If $k\xi$ is removed then the distance $(k-1)\xi + \tau = k\xi - (\xi - \tau)$ is not realizable. (It can't be realized using the mark $K\xi + \eta$.)

Finally, $\#\mathcal{M}_N = \xi + K$ and to show (6) we must prove the inequality

$$2\sqrt{\xi(\xi+t)+\eta} - 1 < 2\xi + t < 2\sqrt{\xi(\xi+t)+\eta}.$$

for $t = 0, 1$ and $1 \le \eta < \xi$. Again, squaring both sides of each inequality to eliminate radicals and some algebra does the job. ∎

Minimal spanning rulers can also be quite *thickly* marked.

THEOREM 4. *For any $N > 0$ there is a minimal spanning ruler $\mathcal{M}_N$ with*

$$\#\mathcal{M}_N > \frac{1}{2}N.$$

Proof. A moment's reflection shows that if $N = 2n$ or $N = 2n + 1$ then $\mathcal{M}_N = \{0, 1, \ldots, n, N\}$ works. ∎

Acknowledgment We thank the referees for many valuable suggestions, including a very nice reformulation of the statement of Theorem 3.

REFERENCES

1. F. Biraud, E. J. Blum, and J. C. Ribes, On optimum synthetic linear arrays with application to radio astronomy, *IEEE Transactions on Antennas and Propagation* **AP-22** (1974) 108–109. doi:10.1109/TAP.1974.1140732
2. Gary S. Bloom and Solomon W. Golomb, Applications of numbered undirected graphs, *Proc. IEEE* **65** (April 1977) 562–570. doi:10.1109/PROC.1977.10517
3. A. K. Dewdney, *The Armchair Universe*, W. H. Freeman, New York, 1988.
4. A. Dimitromanolakis, Analysis of the Golomb ruler and the Sidon set problem, and determination of large near-optimal Golomb rulers, Diploma thesis, Technical University of Crete (Greece) 2002. (The English version can be downloaded from http://www.cs.toronto.edu/~apostol/golomb.)
5. Martin Gardner, *The Magic Numbers of Dr. Matrix*, Prometheus Books, Amherst, NY, 1985.
6. Martin Gardner, Mathematical Games, *Sci. Amer.* **226** (March 1972) 108–112. doi:10.1038/scientificamerican0372-108
7. Martin Gardner, *Wheels, Life, and Other Mathematical Amusements*, W. H. Freeman, New York, 1985.
8. Solomon W. Golomb, The use of combinatorial structures in communication signal design, pp. 59–78 in C. Mitchell (ed.), *Applications of Combinatorial Mathematics*, Oxford University Press, Oxford, UK, 1997.
9. H. Halberstam and K. F. Roth, *Sequences*, Springer, New York, 1983.
10. Alan T. Moffet, Minimum-redundancy linear arrays, *IEEE Transactions on Antennas and Propagation* **AP-16** (March 1968) 172–175. doi:10.1109/TAP.1968.1139138
11. Ed Pegg, Math Games: Rulers, Arrays, and Gracefulness, http://www.maa.org/editorial/mathgames/mathgames_11_15_04.html.
12. William T. Rankin, Optimal Golomb Rulers: An exhaustive parallel search implementation, M.S. thesis, Duke University, 1993. (Available at http://people.ee.duke.edu/~wrankin/golomb/golomb.html).
13. James B. Shearer, Golomb rulers webpage, http://www.research.ibm.com/people/s/shearer/grule.html.
14. Eric W. Weisstein, Golomb Ruler, MathWorld: A Wolfram Web Resource, http://mathworld.wolfram.com/GolombRuler.html.

Summary What is the smallest number of inch marks on a ruler that allow us to measure all integral distances? This question motivates our survey of Golomb rulers, perfect rulers, and minimal rulers—different types of rulers for allowing the most measurements with the smallest number of marks.

Pascal's Hexagon Theorem Implies the Butterfly Theorem

GREG MARKOWSKY
Pohang Mathematics Institute
POSTECH
Pohang, 790-784
Republic of Korea
gmarkowsky@gmail.com

It is a shame to pass up any opportunity to apply Pascal's Hexagon Theorem. The purpose of this paper is to use this beautiful theorem to give a quick proof of another theorem in plane geometry, known as the Butterfly Theorem. We also discuss how the method used here can be extended to general conics. In what follows, ab denotes the line passing through a and b, $\overline{ab}$ denotes the line segment between a and b, and $|ab|$ denotes the length of $\overline{ab}$.

BUTTERFLY THEOREM. *Let $\overline{ab}$ be a chord of a circle with midpoint m. Suppose $\overline{rs}$ and $\overline{uv}$ are two other chords that pass through m. Suppose that rv and us intersect ab, and let $p = rv \cap ab$ and $q = us \cap ab$. Then $|pm| = |qm|$.*

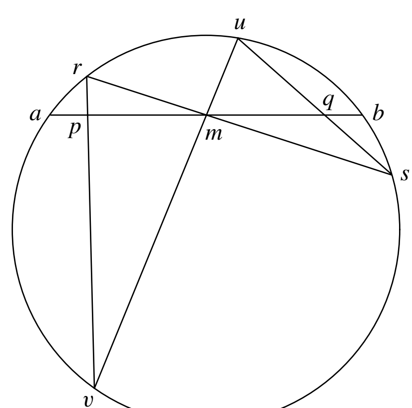


Figure 1

Note that it is not required that the points r, s, u, v lie in the configuration shown in FIGURE 1. For instance, u and v can be interchanged, and though this would result in p and q lying outside the circle the conclusion of the theorem would still hold. Numerous proofs of this theorem have been developed over time. For examples see [**1**], [**4**, p. 45], or [**5**]. The website `http://www.cut-the-knot.org` contains an excellent discussion of this theorem as well. For the proof in this paper, we will need only the following theorem as prerequisite.

PASCAL'S HEXAGON THEOREM. *Let a, b, c, d, e, f be six points on a conic. Suppose that ab intersects de at u, bc intersects ef at w, and cd intersects fa at v. Then u, v, and w are collinear.*

Math. Mag. **84** (2011) 56–62. doi:10.4169/math.mag.84.1.056.

FIGURE 2 gives an example for an ellipse. The theorem states that uvw is a straight line.

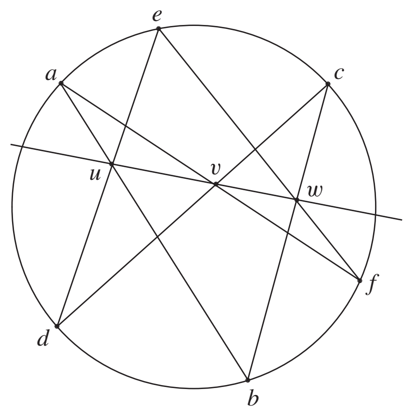


Figure 2

The reader may not be used to thinking of the lines connecting *a, b, c, d, e, f* in the above configuration as a hexagon. Nevertheless, Pascal's Theorem applies, as it applies to a standard hexagon and to a host of other cases, such as when some of the six points lie on different branches of a hyperbola, and even when two consecutive points are allowed to coincide (here the line connecting the two points is replaced by the tangent to the conic). The configuration above was chosen for ease of drawing, since in many cases some or all of the points of intersection lie outside the ellipse. Naturally, numerous proofs of this theorem exist as well; [**10**] contains a number of them, and http://www.cut-the-knot.org is another great reference. Let us assume Pascal's theorem and take on the Butterfly Theorem.

PROPOSITION 1. *Let rs be the diameter of a semicircle, and let a and b be points on the semicircle not equal to r or s. Let* $c = as \cap br$. *Let d be a point on rs. Then the following are equivalent:*

(i) *cd is perpendicular to rs.*

(ii) $\angle adr = \angle bds$.

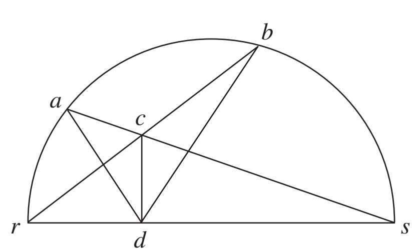


Figure 3

Proof. In FIGURE 4 points a' and b' are the reflections of a and b across rs.

If $i = ab' \cap a'b$ it is clear by symmetry that $i \in rs$ and $\angle air = \angle bis$. Furthermore, Pascal's Theorem applied to the hexagon $asa'brb'$ shows that cic' is a straight line, and symmetry shows that it must be perpendicular to rs. Thus, (i) and (ii) both hold if, and only if, $d = i$. ■

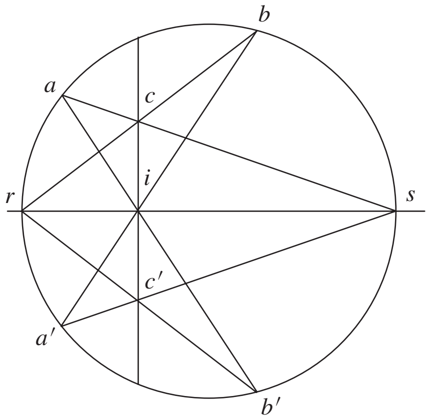


Figure 4

Let us refer to such a pair (a, b) as a *reflected pair* about d. The reader may be interested to note that the diagram in the proof of Proposition 1 furnishes a proof of the Butterfly Theorem in case one of the chords $\overline{rs}$, $\overline{uv}$ is a diameter. That is, cc' extends to a chord of the circle with midpoint i, and rs and ab' pass through d. The proposition shows that $|ic| = |ic'|$. Another application of Pascal's Theorem allows us to prove the next proposition and extend this special case to the general Butterfly Theorem for a circle.

PROPOSITION 2. *Let (a, b), (a', b') be distinct reflected pairs about a point d on rs. Let e be chosen on the semicircle so that de is perpendicular to rs. Then ab' and $a'b$ intersect at a point on de.*

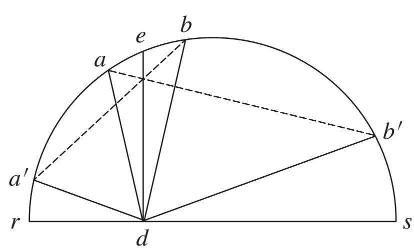


Figure 5

Proof. Connect as, $a's$, br, $b'r$ (FIGURE 6). By Proposition 1, $as \cap br$ lies on de, as does $a's \cap b'r$. Thus, by Pascal's Theorem applied to the hexagon $asa'brb'$, $ab' \cap a'b$ does as well. ∎

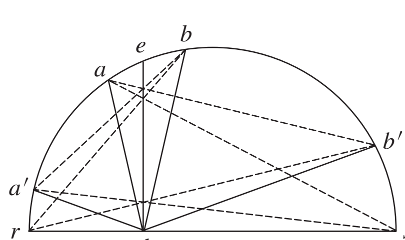


Figure 6

Proof of Butterfly Theorem. In FIGURE 1, reflect r and v across the diameter passing through m to points r' and v'. This gives the picture in FIGURE 7.

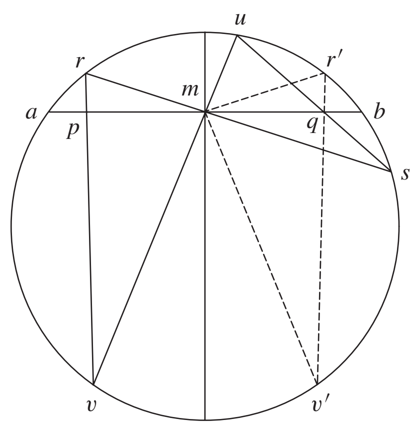


Figure 7

Now r', s and u, v' are each reflected pairs around m, so by Proposition 2, $r'v'$ and us intersect on mb. This point of intersection is q, but it is also the reflection of p, and it follows from this that $|pm| = |qm|$. ■

Pascal's Theorem is valid for all conics, as is the Butterfly Theorem, so since mathematics abhors a coincidence we should suspect that this method of proof works for general conics. Indeed it does, provided we appropriately adjust the notion of reflection. We will need the following proposition.

PROPOSITION 3. *Let a chord $\overline{ab}$ of a conic C be given, and let F be the family of chords of C parallel to $\overline{ab}$. Then the midpoints of all chords in F lie on a line ℓ.*

The line ℓ is called the conjugate diameter of F. The reason for this terminology is that in the case of ellipses it can be shown that ℓ always passes through the center of the ellipse, and ℓ is in that sense a diameter. The same is true with hyperbolas, although ℓ need not necessarily intersect the hyperbola. Conjugate diameters for parabolas can be shown to be lines parallel to the axis of the parabola. The situation for parabolas and hyperbolas indicates that it might be a slight abuse of terminology to refer to ℓ as a "diameter" of any type, but this term is common and we will use it. Conjugate diameters seem to have fallen out of style with modern textbook writers, but http://www.wikipedia.com and http://www.cut-the-knot.org contain brief illustrations of the concept for ellipses. The more classical texts [**6**, pp. 66, 68, 82] and [**9**, p. 138] both cover the concepts in great detail, and give proofs of existence for the various conics. In any event, given the initial chord $\overline{ab}$ we let F be the family of chords of the conic parallel to $\overline{ab}$ and let ℓ be the conjugate diameter of F. We define the *affine reflection* of any point p across ℓ to be the point p' such that pp' is parallel to $\overline{ab}$ and the midpoint of pp' is on ℓ. We also state that the line k' is the reflection of the line k across ℓ if the affine reflection across ℓ of every point on k lies on k'. It is not hard to show that, if j, k are lines which meet at a point p, then the affine reflections j', k' of these lines meet at the point p', the affine reflection of p. For more on affine reflections, see [**8**, pp. 44–46] or [**3**, pp. 203, 207, 208].

For our purposes, note that Proposition 3 implies that the affine reflection of any point on the conic across a conjugate diameter is another point on the conic. Let d be a point on ℓ. We define a pair of points r and s on the ellipse to be an *affine reflected pair* about d if rds' is a straight line, where s' is the reflection of s over ℓ.

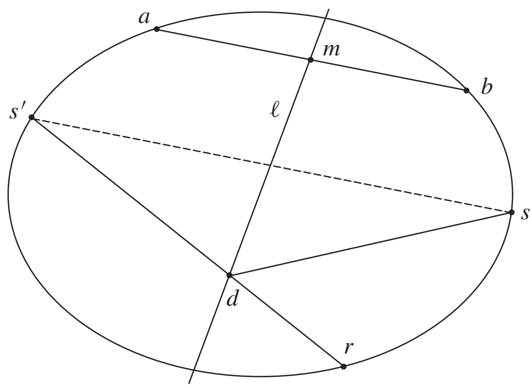


Figure 8

We see from the preceding discussion and definitions that the relevant properties of reflections relative to a circle are possessed by affine reflections relative to a general conic. Thus, the analogs of Propositions 1 and 2 as well as the Butterfly Theorem follow for conics in exactly the same manner as in the case of the circle, with one exception. Suppose that the initial chord $\overline{ab}$ contains one point from each branch of a hyperbola. The conjugate diameter ℓ of F exists, but in this case does not intersect the hyperbola, depriving us of the six points necessary to form a hexagon.

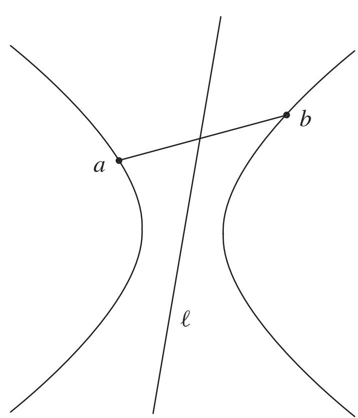


Figure 9

Rectifying this problem requires a bit of projective geometry and a deeper version of Pascal's Theorem than was given at the beginning of the paper. Note that in the given statement, it is assumed that the pairs of opposite sides of the hexagon intersect. This need not occur, though, as some pairs of opposite sides may be parallel, and it would seem that in this case Pascal's Theorem does not apply. However, the natural setting for Pascal's Theorem is in fact not the plane, but rather the *projective plane*. The projective plane is formed by adding an extra "line at infinity" to the plane. Each point on this line at infinity corresponds to a direction that a line in the plane may take. In this context, every pair of lines intersect at exactly one point, since parallel lines intersect at a point which lies on the line at infinity. For more on the projective plane, see [**8**, chap. 5] or [**2**, chap. 5]. It is a beautiful fact that Pascal's Theorem still holds in the projective plane, and we have the following special case:

PASCAL'S HEXAGON THEOREM WITH ONE PAIR OF PARALLEL SIDES. *Suppose the six points a, b, c, d, e, f lie on a conic, and that ab and de are parallel. If bc intersects ef and cd intersects fa then the line passing through $bc \bigcap ef$ and $cd \bigcap fa$ is parallel to ab and de.*

Proof. $ab \bigcap de$ lies on the line at infinity. Pascal's Theorem states that the line through $bc \bigcap ef$ and $cd \bigcap fa$ passes through $ab \bigcap de$ as well. This is equivalent to the statement of the theorem. ■

BUTTERFLY THEOREM FOR ARBITRARY CONICS. *Let $\overline{ab}$ be a chord of a conic with midpoint m. Suppose $\overline{rs}$ and $\overline{uv}$ are two other chords that pass through m. Suppose that ru and vs intersect ab, and let $p = ru \cap ab$ and $q = vs \cap ab$. Then $|pm| = |qm|$.*

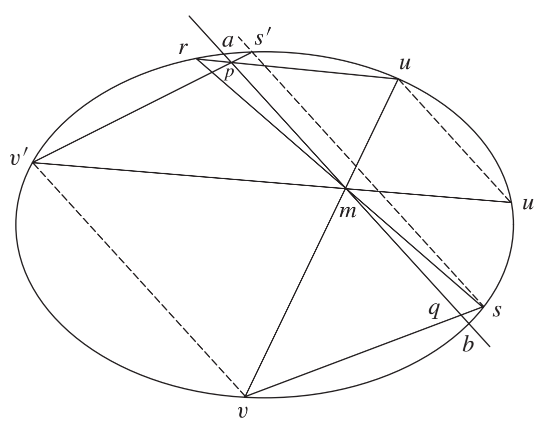


Figure 10

Proof. Let s', u', v' be the affine reflections of s, u, v across the conjugate diameter ℓ of the family of chords parallel to ab. Note that $u'v'$ is the affine reflection of uv across ℓ. The affine reflection across ℓ fixes points on ℓ, so since uv passes through m, so must $u'v'$. Thus, $m = rs \bigcap u'v'$. Consider the hexagon $rss'v'u'u$. Since uu' and ss' are parallel, we can apply our special case of Pascal's Hexagon Theorem to conclude that $m = rs \bigcap u'v'$ and $ru \bigcap s'v'$ lie on a line parallel to uu' and ss', namely ab. Since $ru \bigcap ab = p$, we must have $ru \bigcap s'v' = p$. Now $s'v'$ is the affine reflection of sv across ℓ, and it follows that p and q are affine reflections of each other across ℓ as well. Therefore, $|pm| = |qm|$. ■

Extending the theorem The Butterfly Theorem is a special case of a more general theorem, in which the point m is no longer required to be the midpoint of $\overline{ab}$. This was proved by Candy in 1896, is discussed in [**1**, p. 207], and can be quite naturally deduced from Pascal's Theorem in a similar manner to the above. Recently this theorem has been extended to a theorem in the complex projective plane in [**7**].

Acknowledgment The author was supported by the Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (Grant #2009-0094070).

REFERENCES

1. L. Bankoff, The metamorphosis of the Butterfly Problem, *Mathematics Magazine* **60** (1987) 195–210. doi:10.2307/2689339
2. M. Bennett, *Affine and Projective Geometry*, John Wiley, New York, 1995.
3. H. Coxeter, *An Introduction to Geometry*, 2nd ed., John Wiley, New York, 1989.
4. H. Coxeter and S. Greitzer, *Geometry Revisited*, Mathematical Association of America, Washington, DC, 1967.
5. R. Honsberger, The Butterfly Problem and other delicacies from the Noble Art of Euclidean Geometry I, *Two-Year College Mathematics Journal* **14** (1983) 2–8.
6. F. S. Macauley, *Geometrical Conics*, Cambridge University Press, Cambridge, 1895.
7. G. Markowsky, Pascal's Hexagon Theorem implies the Butterfly Theorem in the complex projective plane, preprint, http://arxiv.org/PS_cache/arxiv/pdf/0910/0910.4792v1.pdf.
8. P. Ryan, *Euclidean and Non-Euclidean Geometry: An Analytic Approach*, Cambridge University Press, Cambridge, 1986.
9. G. Salmon, *A Treatise on Conic Sections*, 6th ed., Longmans, Green, Reader, and Dyer, 1879.
10. K. Tan, Various proofs of Pascal's Theorem, *Mathematics Magazine* **38** (1965) 22–30. doi:10.2307/2688012

Summary Pascal's Hexagon Theorem is used to prove the Butterfly Theorem for conics, a well known result in Euclidean geometry. In the course of the proof, some basic concepts in projective geometry are introduced.

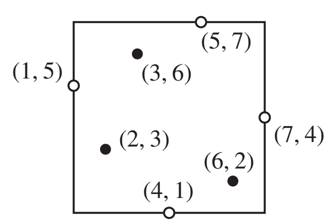
PROBLEMS

PROPOSALS

To be considered for publication, solutions should be received by July 1, 2011.

1861. *Proposed by Emeric Deutsch, Polytechnic Institute of New York University, Brooklyn, NY.*

Let $n \geq 2$ be an integer. A permutation $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ can be represented in the plane by the set of n points $P_\sigma = \{(k, \sigma(k)) : 1 \leq k \leq n\}$. The smallest square bounding P_σ, with sides parallel to the coordinate axis, has at least 2 and at most 4 points of P_σ on its boundary. The figure below shows a permutation σ with 4 points on its bounding square. For every $m \in \{2, 3, 4\}$, determine the number of permutations σ of $\{1, 2, \ldots, n\}$ having m points of P_σ on the boundary of their bounding square.



1862. *Proposed by H. A. ShahAli, Tehran, Iran.*

Let n be a positive integer. Suppose that the nonnegative real numbers $a_1, b_1, a_2, b_2, \ldots, a_n, b_n$ satisfy that $a_1 \leq a_2 \leq \cdots \leq a_n$ and $\sum_{i=1}^k a_i \geq \sum_{i=1}^k b_i$ for all $1 \leq k \leq n$. Prove that $\prod_{i=1}^k a_i \geq \prod_{i=1}^k b_i$ for all $1 \leq k \leq n$.

Math. Mag. **84** (2011) 63–71. doi:10.4169/math.mag.84.1.063. © Mathematical Association of America

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. Submitted problems should not be under consideration for publication elsewhere.

Solutions should be written in a style appropriate for this MAGAZINE.

Solutions and new proposals should be mailed to Bernardo M. Ábrego, Problems Editor, Department of Mathematics, California State University, Northridge, 18111 Nordhoff St, Northridge, CA 91330-8313, or mailed electronically (ideally as a LaTeX or pdf file) to mathmagproblems@csun.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.

1863. *Proposed by Duong Viet Thong, Department of Economics and Mathematics, National Economics University, Hanoi, Vietnam.*

Let f be a continuously differentiable function on $[a, b]$ such that $\int_a^b f(x)\,dx = 0$. Prove that

$$\left|\int_a^b xf(x)\,dx\right| \le \frac{(b-a)^3}{12}\max\{|f'(x)| : x \in [a, b]\}.$$

1864. *Proposed by Cosmin Pohoata, Princeton University, Princeton, NJ.*

Let ABC be a scalene triangle, I its incenter, and X, Y, and Z the tangency points of its incircle $\mathcal{C}$ with the sides BC, CA, and AB, respectively. Denote by $X' \neq A$, $Y' \neq B$, and $Z' \neq C$ the intersections of $\mathcal{C}$ with the circumcircles of triangles AIX, BIY, and CIZ, respectively. Prove that the lines AX', BY', and CZ' are concurrent.

1865. *Proposed by Erwin Just (Emeritus), Bronx Community College of the City University of New York, Bronx, NY.*

In the solution to Problem 1790 (this MAGAZINE **82** (2009) 67–68), it was proved that if R is a ring such that for each element $x \in R$,

$$x + x^2 + x^3 + x^4 = x^{11} + x^{12} + x^{13} + x^{28},$$

then for each element $x \in R$, $x = x^{127}$. Under the same hypothesis, prove that for each element $x \in R$, $6x = 0$ and $x = x^7$.

Quickies

Answers to the Quickies are on page 71.

Q1007. *Proposed by Daniel Goffinet, St. Etienne, France.*

What is the 1007th term of the Mclaurin series expansion of

$$f(t) = \frac{\cos^2 t - \frac{1}{2}}{\sin^2 t - \frac{1}{2}}?$$

Q1008. *Proposed by Steve Butler, UCLA, Los Angeles, CA.*

For every positive real a, find the value of

$$\int_0^1 ((1 - x^a)^{1/a} - x)^2\,dx.$$

Solutions

Non-congruent triangles in a regular n-gon **February 2010**

1836. *Proposed by Michael Wolterman, Washington and Jefferson College, Washington, PA.*

Let $n \ge 3$ be a natural number. Find how many pairwise non-congruent triangles are there among the $\binom{n}{3}$ triangles formed by selecting three vertices of a regular n-gon.

Solution by Bob Tomper, Mathematics Department, University of North Dakota, Grand Forks, ND.

Label the vertices of the n-gon as $v_0, v_1, v_2, \ldots, v_{n-1}$. Each triangle formed by selecting three vertices v_i, v_j, and v_k with $0 < i < j < k < n$ can be encoded as a triple of positive integers $(j-i, k-j, n+i-k)$ with sum n, and every triple of positive integers with sum n is the encoding of some such triangle. Since two triangles are congruent if and only if their encodings have the same set of entries, it follows that the number of triangles is $P(n,3)$, the number of partitions of n into three positive parts. It is well-known that $P(n,3)$ is the nearest integer to $n^2/12$.

Editor's Note. The fact that the number of partitions of n into three positive parts is the nearest integer to $n^2/12$ appears as Exercise 15.1 in J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*, Cambridge Univ. Press, Cambridge, 1992. As noted by Rob Pratt, this problem appeared as Problem 15B in the previous text and also as Problem 3893 in *Amer. Math. Monthly* **45** (1938) 631–632.

Also solved by Armstrong Problem Solvers, Tom Beatty, Jany C. Binz (Switzerland), Elton Bojaxhiu (Albania) and Enkel Hysnelaj (Australia), Bruce S. Burdick, Robert Calcarreta, William J. Cowieson, Chip Curtis, Robert L. Doucette, Dmitry Fleischman, G.R.A.20 Problem Solving Group (Italy), Lucyna Kabza, Elias Lampakis (Greece), Masao Mabuchi (Japan), Missouri State University Problem Solving Group, José Heber Nieto (Venezuela), Rob Pratt, Joel Schlosberg, John H. Smith, James Swenson, Marian Tetiva (Romania), Mark S. Weisser, G. Gerard Wojnar, and the proposer.

Rolle's theorem in action **February 2010**

1837. *Proposed by Duong Viet Thong, Nam Dinh University of Technology Education, Nam Dinh City, Vietnam.*

Let $f : [1,2] \to \mathbb{R}$ be a continuous function such that $\int_1^2 f(x)\,dx = 0$. Prove that there exists a real number c in the open interval $(1,2)$, such that $cf(c) = \int_c^2 f(x)\,dx$.

Solution by Ángel Plaza and Sergio Falcón, Department of Mathematics, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain.

Define the function $F(t) = t\int_t^2 f(x)\,dx$, for $t \in [1,2]$. Clearly F is continuous on $[1,2]$, differentiable on $(1,2)$, and $F(1) = F(2) = 0$. Therefore Rolle's theorem implies that there is $c \in (1,2)$ such that $F'(c) = 0$. Because $F'(c) = \int_c^2 f(x)\,dx - cf(c)$, it follows that $cf(c) = \int_c^2 f(x)\,dx$.

Also solved by Arkady Alt, Michel Bataille (France), Elton Bojaxhiu (Albania) and Enkel Hysnelaj (Australia), Michael W. Botsko, Paul Bracken, Paul Budney, Robert Calcaterra, Minh Can, John Christopher, William J. Cowieson, Chip Curtis, Robert L. Doucette, Brad Emmons, G.R.A.20 Problem Solving Group (Italy), Jeff Groah, Lee O. Hagglund, Timothy Hall, Eugene A. Herman, Alan D. Hetzel Jr. and Alin A. Stancu, Irina Ilioaea (Romania), Kamil Karayilan (Turkey), Elias Lampakis (Greece), David P. Lang, Longxiang Li (China) and Yi Zheng (China), Tianyu Li (China) and Xiaoxiang Wang (China), Charles Lindsey, Rick Mabry (Germany), Northwestern University Math Problem Solving Group, Paolo Perfetti (Italy), Phuong Pham, Rob Pratt, Henry Ricardo, Kevin Roper, Achilleas Sinefakopoulos (Greece), Bob Tomper, Haohao Wang and Jerzy Wojdyło, "Why So Series?", and the proposer. There was one incorrect submission.

A weakly convergent series of logs **February 2010**

1838. *Proposed by Costas Efthimiou, University of Central Florida, Orlando, FL.*

Compute the sum

$$\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}(-1)^{n+m}\frac{\ln(m+n)}{m+n}.$$

I. *Solution by Tiberiu Trif, Babeş-Bolyai University, Cluj-Napoca, Romania.*

The answer is $\frac{1}{2}\ln(\pi/2)$. By Leibniz' test the alternate series $\sum_{k=1}^{\infty}(-1)^k \frac{\ln k}{k}$ converges. Let σ denote its sum. Set

$$s := \sum_{n=0}^{\infty}\sum_{m=1}^{\infty}(-1)^{n+m}\frac{\ln(n+m)}{n+m} \text{ and } s_n := \sum_{k=0}^{n}\sum_{m=1}^{\infty}(-1)^{k+m}\frac{\ln(k+m)}{k+m}.$$

Because

$$\sum_{m=1}^{\infty}(-1)^{k+m}\frac{\ln(k+m)}{k+m} = \sigma - \sum_{j=1}^{k}(-1)^j\frac{\ln j}{j},$$

it follows that

$$\begin{aligned} s_{2n} &= (2n+1)\sigma - \sum_{k=1}^{2n}\sum_{j=1}^{k}(-1)^j\frac{\ln j}{j} = (2n+1)\sigma - \sum_{k=1}^{2n}(2n+1-k)(-1)^k\frac{\ln k}{k} \\ &= (2n+1)\sigma - (2n+1)\sum_{k=1}^{2n}(-1)^k\frac{\ln k}{k} + \sum_{k=1}^{2n}(-1)^k\ln k \\ &= \frac{\sigma - \sum_{k=1}^{2n}(-1)^k\frac{\ln k}{k} + \frac{1}{2n+1}\ln\frac{(2n)!!}{(2n-1)!!}}{\frac{1}{2n+1}}. \end{aligned}$$

Here $m!!$ denotes the double factorial defined recursively as $m!! = m((m-2)!!)$ and $1!! = 0!! = 1$. By the Cesáro-Stolz Theorem for a quotient of sequences, where both numerator and denominator approach 0, we have that

$$\lim_{n\to\infty} s_{2n} = \lim_{n\to\infty}\frac{\frac{\ln(2n+1)}{2n+1} - \frac{\ln(2n+2)}{2n+2} + \frac{1}{2n+3}\ln\frac{(2n+2)!!}{(2n+1)!!} - \frac{1}{2n+1}\ln\frac{(2n)!!}{(2n-1)!!}}{\frac{1}{2n+3} - \frac{1}{2n+1}}.$$

A long but elementary computation leads to

$$\begin{aligned} \lim_{n\to\infty} s_{2n} &= -\frac{1}{2}\lim_{n\to\infty}\frac{1}{2n+2}\ln\frac{((2n-1)!!)^{2(2n+2)}(2n+1)^{2(2n+2)}}{((2n)!!)^{2(2n+2)}(2n+2)^{2n+1}} \\ &= \frac{1}{2}\lim_{n\to\infty}\left(\ln\frac{((2n)!!)^2}{((2n-1)!!)^2(2n+1)} + \ln\frac{2n+2}{2n+1} - \frac{\ln(2n+2)}{2n+2}\right). \end{aligned}$$

The last two terms go to 0 when $n \to \infty$ and according to Wallis' formula

$$\lim_{n\to\infty}\frac{((2n)!!)^2}{((2n-1)!!)^2(2n+1)} = \frac{\pi}{2}.$$

Thus $\lim_{n\to\infty} s_{2n} = \frac{1}{2}\ln(\pi/2)$. Finally, since $s_{2n+1} = s_{2n} + \sum_{k=2n+2}^{\infty}(-1)^k\ln k/k$ and $\lim_{n\to\infty}\sum_{k=2n+2}^{\infty}(-1)^k\ln k/k = 0$, we conclude that $s = \lim_{n\to\infty} s_n = \frac{1}{2}\ln(\pi/2)$.

II. *Solution by the proposer.*

Using the fact the function $\ln s/s$ is the Laplace transform of $-\ln t + \gamma$ (where γ is the Euler-Mascheroni constant), we can rewrite the sum as an integral:

$$\begin{aligned}\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}(-1)^{n+m}\frac{\ln(m+n)}{m+n} &= \sum_{n=0}^{\infty}\sum_{m=1}^{\infty}(-1)^{n+m}\int_0^{\infty} e^{-(m+n)t}(\ln t+\gamma)\,dt\\ &= -\int_0^{\infty}(\ln t+\gamma)\left(\sum_{n=0}^{\infty}(-e^{-t})^n\right)\left(\sum_{m=1}^{\infty}(-e^{-t})^m\right)dt\\ &= -\int_0^{\infty}(\ln t+\gamma)\frac{-e^{-t}}{(1+e^{-t})^2}dt.\end{aligned}$$

We now use the known integral $\int_0^1(\ln\ln(u^{-1}))/(u+1)^2\,du = \frac{1}{2}(\ln(\pi/2)-\gamma)$ and the change of variable $u = e^{-t}$ to get

$$\begin{aligned}\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}(-1)^{n+m}\frac{\ln(m+n)}{m+n} &= \int_0^1\frac{\ln\ln u^{-1}+\gamma}{(u+1)^2}du\\ &= \int_0^1\frac{\ln\ln u^{-1}}{(u+1)^2}du+\gamma\int_0^1\frac{1}{(u+1)^2}du\\ &= \frac{1}{2}\left(\ln\frac{\pi}{2}-\gamma\right)+\frac{\gamma}{2}=\frac{1}{2}\ln\frac{\pi}{2}.\end{aligned}$$

Editor's Note. Some readers tried to solve the problem by adding over the diagonals $m+n=d$; unfortunately the resulting rearrangement of the series is not convergent. Ovidiu Furdui and Huizeng Qin note that the generalization of finding $\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(-1)^{n+m}\ln^p(m+n)/(m+n)$ for every $p\geq 1$ has been published in Problema 149, *La Gaceta de la RSME* **13** (2010) 78–79.

Also solved by Mark S. Ashbaugh and Francisco Vial (Chile), Bruce S. Burdick, Ovidiu Furdui (Romania) and Huizeng Qin (China), G.R.A.20 Problem Solving Group (Italy), Timothy Hall, Eugene A. Herman, Enkel Hysnelaj (Australia) and Elton Bojaxhiu (Albania), Nicholas C. Singer, and Bob Tomper. There were six incorrect submissions.

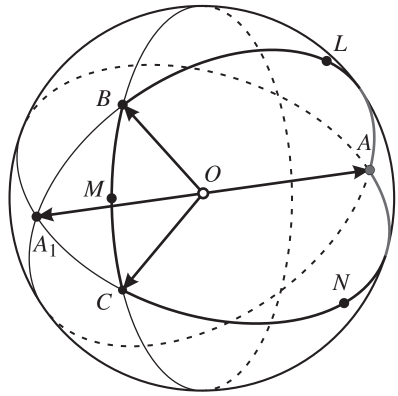
A characterization of an octant of a sphere **February 2010**

1839. *Proposed by Robert A. Russell, New York, NY.*

Consider a sphere of radius 1 and three points A, B, and C on its surface, such that the area of the convex spherical triangle ABC is π. Let L, M, and N be the midpoints of the shortest arcs AB, BC, and CA. Give a characterization of the spherical triangle LMN.

Solution by Elton Bojaxhiu (Albania) and Enkel Hysnelaj (Australia)

The spherical triangle LMN is an octant of the sphere, i.e., a right equilateral triangle. Denote by O the center of the unit sphere. For any three points X, Y, and Z in space, denote by $\angle XYZ$ the angle XYZ. For any three points X, Y, and Z on the surface of the sphere, denote by $\sphericalangle XYZ$ the spherical angle XYZ. Let $\alpha = \angle BOC$, $\beta = \angle AOC$, $\gamma = \angle AOB$, $\alpha' = \sphericalangle BAC$, $\beta' = \sphericalangle ABC$, and $\gamma' = \sphericalangle ACB$. Let A_1 be the point on the sphere antipodal to A.



From the given condition, we have that

$$\text{Spherical Area}(ABC) = \alpha' + \beta' + \gamma' - \pi = \pi,$$

i.e., $\alpha' + \beta' + \gamma' = 2\pi$. This implies that $\alpha' = (\pi - \beta') + (\pi - \gamma')$, and thus

$$\sphericalangle BA_1C = \sphericalangle A_1BC + \sphericalangle A_1CB.$$

This means that there exists a point P in the (shortest) arc BC such that $\sphericalangle A_1CB = \sphericalangle CA_1P$ and $\sphericalangle A_1BC = \sphericalangle BA_1P$. However these identities imply that $\angle A_1OP = \angle COP$ and $\angle A_1OP = \angle BOP$, thus P must coincide with M and $\angle A_1OM = \frac{1}{2}\angle BOC = \alpha/2$.

Let $a = \overrightarrow{OA}$, $b = \overrightarrow{OB}$, and $c = \overrightarrow{OC}$. Now, $\overrightarrow{OA_1} = -a$. Since OM is the bisector of $\angle BOC$ and $\angle A_1OM = \alpha/2$, it follows that

$$\begin{aligned} -\cos\gamma - \cos\beta &= -a\cdot b - a\cdot c = (-a)\cdot(b+c) = |-a|\,|b+c|\,\cos(\alpha/2) \\ &= \sqrt{(b+c)\cdot(b+c)}\,\cos(\alpha/2) = \sqrt{2+2(b\cdot c)}\,\cos(\alpha/2) \\ &= \sqrt{(2+2\cos\alpha)\left(\frac{1+\cos\alpha}{2}\right)} = 1+\cos\alpha. \end{aligned}$$

Therefore $\cos\alpha + \cos\beta + \cos\gamma + 1 = 0$. Finally, since

$$\begin{aligned} (a+b)\cdot(a+c) &= (b+c)\cdot(b+a) = (c+a)\cdot(c+b) \\ &= a\cdot b + b\cdot c + c\cdot a + 1 = \cos\alpha + \cos\beta + \cos\gamma + 1 = 0, \end{aligned}$$

it follows that $OL \perp ON$, $OM \perp OL$, and $ON \perp OM$. Thus $\angle LON = \angle MOL = \angle NOM = \frac{\pi}{2}$ and $\sphericalangle LMN = \sphericalangle MNL = \sphericalangle NLM = \frac{\pi}{2}$, which is the desired characterization.

Editor's Note. Jim Delany observes that this problem is a particular case of a problem that appears in J. D. H. Donnay, Spherical Trigonometry, Interscience Publishers, Inc., New York, 1945, 62–65. Michel Bataille indicates that the problem also follows from the note by J. Brooks and J. Strantzen, Spherical Triangles of Area π and Isosceles Tetrahedra, this MAGAZINE **78** (2005) 311–314.

Also solved by Michel Bataille (France), Jim Delany, Raúl A. Simón, and the proposer.

A generalization of Nesbitt's Inequality **February 2010**

1840. *Proposed by Tuan Le, 12th grade, Fairmont High School, Anaheim, CA.*

Let a, b, and c be nonnegative real numbers such that no two of them are equal to zero. Prove that

$$\frac{a}{b+c}+\frac{b}{c+a}+\frac{c}{a+b}+\frac{3\sqrt[3]{abc}}{2(a+b+c)}\geq 2.$$

Solution by Arkady Alt, San Jose, CA.

Let us assume first that one of the numbers is zero. Suppose that $a=0$, then $b, c\neq 0$ and the inequality becomes $b/c+c/b\geq 2$ which is true by the Arithmetic Mean–Geometric Mean Inequality. From now on, assume that $abc\neq 0$. By the Cauchy–Schwarz Inequality,

$$\left(\frac{a^2}{ab+ca}+\frac{b^2}{bc+ab}+\frac{c^2}{ca+bc}\right)((ab+ca)+(bc+ab)+(ca+bc))$$
$$\geq (a+b+c)^2,$$

thus,

$$\sum_{cyc}\frac{a}{b+c}=\sum_{cyc}\frac{a^2}{ab+ca}\geq\frac{(a+b+c)^2}{2(ab+bc+ca)}\,.$$

Therefore, it suffices to prove that

$$\frac{(a+b+c)^2}{2(ab+bc+ca)}+\frac{3\sqrt[3]{abc}}{2(a+b+c)}\geq 2\,.$$

This is equivalent to

$$(a+b+c)^3+3\sqrt[3]{abc}(ab+bc+ca)-4(a+b+c)(ab+bc+ca)\geq 0\,,$$

and further equivalent to

$$\big((a+b+c)^3-4(a+b+c)(ab+bc+ca)+9abc\big)$$
$$+\,3\sqrt[3]{abc}\left(ab+bc+ca-3\sqrt[3]{a^2b^2c^2}\right)\geq 0. \qquad (1)$$

By Schur's inequality $\sum_{cyc}a(a-b)(a-c)\geq 0$. This inequality can be rewritten in the form $(a+b+c)^3\geq 4(a+b+c)(ab+bc+ca)-9abc$, so the expression in the first line of Inequality (1) is nonnegative. By the Arithmetic Mean–Geometric Mean Inequality, $ab+bc+ca\geq 3\sqrt[3]{a^2b^2c^2}$, so the expression in the second line of Inequality (1) is also nonnegative. This completes the proof. Equality is achieved if and only if $a=b=c$ or two of a, b, and c are equal to each other and the other is equal to 0.

Editor's Note. Some of the solutions submitted used either Schur's inequality or a combination of Schur's inequality with Muirhead's Theorem. Michael Vowe makes a substitution that transform the inequality for the positive numbers a, b, and c, into an inequality for the sides of a triangle. Then he reduces this inequality to the well-known geometric inequality $R\geq 2r$, where R is the circumradius of the triangle and r is the inradius.

Also solved by George Apostolopoulos (Greece), Michel Bataille (France), Elton Bojaxhiu (Albania) and Enkel Hysnelaj (Australia), Chip Curtis, Kee-Wai Lau (Hong Kong), Michael Neubauer, Peter Nüesch (Switzerland), Paolo Perfetti (Italy), Marian Tetiva (Romania), Michael Vowe (Switzerland), and the proposer. There were four incorrect submissions.

A property of a tangential quadrilateral **February 2010**

1834 (corrected). *Proposed by Cosmin Pohoata, student, National College "Tudor Vianu," Bucharest, Romania.*

Let $ABCD$ be a cyclic quadrilateral that also has an inscribed circle with center I, and let ℓ be a line tangent to the incircle. Let A', B', C', and D', respectively, be the projections of A, B, C, and D onto ℓ. Prove that

$$\frac{AA' \cdot CC'}{AI \cdot CI} = \frac{BB' \cdot DD'}{BI \cdot DI}.$$

Composite solution by George Apostolopoulos, Messolonghi, Greece; and Omran Kouba, Higher Institute for Applied Sciences and Technology, Damascus, Syria.

We identify the plane with the set of complex numbers $\mathbb{C}$. The complex number representing a point, denoted by a capital letter, will be denoted by the corresponding lower-case letter. Let us start by proving the following lemma.

LEMMA. *Let $\mathcal{S} = \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle in the complex plane.*

(a) *For every two non-diametrically opposite points U and V in $\mathcal{S}$, the point of intersection of the tangents to $\mathcal{S}$ from U and V is the point represented by the complex number $2uv/(u+v)$.*

(b) *For every point W in $\mathcal{S}$ and every point Z in $\mathbb{C}$, if Z' denotes the projection of Z onto the tangent to $\mathcal{S}$ from W, then $ZZ' = |\mathrm{Re}(z\overline{w}) - 1|$.*

Proof. (a) The point of intersection of the tangents to $\mathcal{S}$ from U and from V is the inverse point with respect to S of the arithmetic mean of U and V; the inverse point of $z = (u+v)/2$ is

$$\frac{1}{\overline{z}} = \frac{2}{\overline{u} + \overline{v}} = \frac{2}{\frac{1}{u} + \frac{1}{v}} = \frac{2uv}{u+v}.$$

(b) The equation of the tangent line to $\mathcal{S}$ from W is $\mathrm{Re}(z\overline{w}) = 1$, then the distance of Z to Z', the projection of Z onto the tangent, is $|\mathrm{Re}(z\overline{w}) - 1|$. ■

Now, we come to our problem. Without loss of generality, we may suppose that the incircle of $ABCD$ is the unit circle $\mathcal{S} = \{z \in \mathbb{C} : |z| = 1\}$. Let S, T, U, and V be the points of contact of circle $\mathcal{S}$ with BC, CD, DA, and AB, respectively. We may also assume that the equation of ℓ is $\mathrm{Re}(z) = -1$, which is the tangent to $\mathcal{S}$ at the point W corresponding to $w = -1$.

Since A is the point of intersection of the tangents to $\mathcal{S}$ from U and V, we conclude using the lemma that $a = 2uv/(u+v)$, and by the second part of the lemma, if A' is the projection of A onto ℓ, then

$$AA' = \left|\mathrm{Re}\left(\frac{2uv}{u+v}(-1)\right) - 1\right| = \left|\mathrm{Re}\left(\frac{2uv}{u+v}\right) + 1\right| = \left|\frac{uv}{u+v} + \frac{\overline{uv}}{\overline{u+v}} + 1\right|$$
$$= \left|\frac{uv}{u+v} + \frac{\frac{1}{uv}}{\frac{1}{u} + \frac{1}{v}} + 1\right| = \left|\frac{(u+1)(v+1)}{u+v}\right|.$$

Similarly, we obtain,

$$BB' = \left|\frac{(v+1)(s+1)}{v+s}\right|, \; CC' = \left|\frac{(s+1)(t+1)}{s+t}\right|, \text{ and } DD' = \left|\frac{(t+1)(u+1)}{t+u}\right|.$$

But, $AI = |a| = |2uv/(u+v)| = 2/|u+v|$, and similarly $BI = 2/|v+s|$, $CI = 2/|s+t|$, and $DI = 2/|t+u|$. Therefore

$$\frac{AA' \cdot CC'}{AI \cdot CI} = \frac{1}{4}\,|(u+1)(v+1)(s+1)(t+1)| = \frac{BB' \cdot DD'}{BI \cdot DI},$$

as we wanted to prove.

Editor's Note. In our haste to fix this problem we corrected the identity to be proved but reintroduced an unnecessary hypothesis; the quadrilateral does not need to be cyclic. Elias Lampakis indicates that the problem's result is a well known theorem attributed to the French mathematician Michel Chasles (1793–1880).

Also solved by Herb Bailey, Elton Bojaxhiu (Albania) and Enkel Hysnelaj (Australia), Robin Chapman (United Kingdom), Dmitry Fleischman, Victor Y. Kutsenok, Elias Lampakis (Greece), Joel Schlosberg, and the proposer.

Answers

Solutions to the Quickies from page 64.

A1007. The answer is 0. The sum of the numerator and denominator is 0, so this fraction is equal to -1 whenever the denominator is nonzero. Therefore all terms of the Maclaurin series except the first are equal to 0.

A1008. Suppose that $f : [0, 1] \to [0, 1]$ is a continuous function such that $f(0) = 1$ and $f(x) = f^{-1}(x)$ for all $x \in [0, 1]$. Note that $f(x) = (1 - x^a)^{1/a}$ satisfies these conditions. Rotate the region under $y = f(x)$ from $x = 0$ to $x = 1$ around the x-axis. Computing the volume V of the resulting solid using the disc method with respect to x and the shell method with respect to y shows that

$$V = \pi \int_0^1 (f(x))^2\,dx = 2\pi \int_0^1 y f^{-1}(y)\,dy = 2\pi \int_0^1 y f(y)\,dy.$$

Therefore,

$$\begin{aligned}\int_0^1 (f(x) - x)^2\,dx &= \int_0^1 (f(x))^2\,dx - 2\int_0^1 x f(x)\,dx + \int_0^1 x^2\,dx \\ &= \int_0^1 x^2\,dx = \frac{1}{3}.\end{aligned}$$

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.

Aron, Jacob, Mathematical immortality? Name that theorem, *New Scientist* (3 December 2010) 6–7, http://www.newscientist.com/article/dn19809-mathematical-immortality-give-a-theorem-your-name.html.

There are companies that offer to name a star after you, even though they don't own it and have no authority to do so. Well, the same spirit of entrepeneurship has arisen in the mathematical world. Unfortunately, this news came out too late for me to notify you about this new source of last-minute mathematics-related winter holiday presents (but there's always the occasion of Gauss's birthday, April 30): You can now give a brand-new theorem and even name it after the recipient. You don't even have to prove the theorem (and can't choose which one)—that is taken care of by the folks at TheoryMine (http://theorymine.co.uk/), a company dedicated to using artificial intelligence to automate theorem proving. "The theorems that we produce are guaranteed to be unique, because we generate them from unique theories." The theorems are not cheaper by the dozen but cost £15 apiece, for which you can download a color PDF certificate of registration of the theorem statement, its name as specified by you, and a brief proof outline. "In the near future, there will be a range of products available to go with your theorem, like T-shirts and mugs"—not to mention an electronic journal in which you can elect to publish your theorem (presumably together with the complete proof).

Aigner, Martin, and Ehrhard Behrends (eds.), *Mathematics Everywhere*, American Mathematical Society, 2010; xiv + 330 pp, $49 ($39.20 for AMS members). ISBN 978-0-8218-4349-9.

In Germany, 2008 was the Year of Mathematics, and this book consists of 21 public lectures at a popular Berlin venue then and earlier (it is a translation of the third edition of *Alles Mathematik*). Curiously, the book begins with a prologue that notes "Since the previous editions ... mathematics has become popular. No longer is there a painful silence at a party when someone says he is a mathematician: admiration instead. Pretending to have no understanding of mathematics is no longer in fashion... it has become cool to be a mathematician." Well, that's definitely not so in the U.S. but maybe in Germany, where "no other branch of research has approached the public with such good humor and as creatively"! (The book neglects to mention that the author of the prologue is not a mathematician but a well-known (in Germany) journalist, editor of the weekly *Die Zeit*, and winner of the Media Prize of the Deutsche Mathematische Vereinigung.) Replete with color illustrations and photographs, this attractive volume of well-written essays highlights case studies (the Reed-Solomon codes behind audio CDs, image processing in liver surgery, shortest-path problems, Turing's morphogenesis in the context of Romeo and Juliet, designing materials through mathematics, tomography, kaleidoscopes), "current topics" (electronic money, the Kepler conjecture, quantum computation, Fermat's Last Theorem, Nash equilibrium, climate change), and "the central theme" (secret codes, knots, soap bubbles, Poincaré conjecture, randomness). The book concludes with Philip J. Davis meditating on "the prospects for mathematics in a multi-media civilization."

Math. Mag. **84** (2011) 72–73. doi:10.4169/math.mag.84.1.072. © Mathematical Association of America

McOwan, Peter, Paul Curzon, and Jonathan Black, *The Magic of Computer Science: Card Tricks Special* (available also in Welsh and in Italian), and *The Magic of Computer Science: Now We Have Your Attention...*, School of Electronic Engineering and Computer Science, Queen Mary, University of London; 64 pp, 60 pp, http://www.cs4fn.org/magic/magicdownload.php.

KFC [initials], Card trick leads to new bound on data compression, The Physics arXiv Blog (Technology Review blog), (26 November 2010), http://www.technologyreview.com/blog/arxiv/26078/.

Gagie, Travis, Bounds from a card trick, http://arxiv.org/abs/1011.4609.

Although authors McCowan et al. use card tricks to arouse the motivation of potential computer scientists, the magic in these beautifully laid-out booklets is in the mathematics involved. Perhaps that angle will inspire your students, too, especially when you mention that the inspiration can also go the other way: Gagie was inspired by card tricks to prove lower bounds for data compression.

Szpiro, George G., *A Mathematical Medley: Fifty Easy Pieces on Mathematics*, American Mathematical Society, 2010; x+ 236 pp, $35 ($28 for AMS members) (P). ISBN 978-0-8218-4928-6.

George Szpiro is the author of several popular topical books about mathematics—*Poincaré's Prize* (2008), *Kepler's Conjecture* (2003)—as well as a previous collection similar to this one, *The Secret Life of Numbers: 50 Easy Pieces on How Mathematicians Work and Think* (2006). Like the essays there, those here are mostly translations of Szpiro's articles for the Sunday edition of a Swiss newspaper. Each tells, in a relaxed but breezy style, a short tale that lies behind a mathematical or scientific concept, result, or paper. The tales weave together mathematical theorems, applications in the everyday world, personal histories, personalities, and recreational mathematics. You will find mock theta functions, Lévy distributions, Braess's paradox, edge-colorings in networks, sequences of primes, Benford's distribution, Nash equilibria, the Li copula, and the axiom of choice; but also algorithms for boarding aircraft, the foraging behavior of albatrosses, Franklin magic squares, Olympic sprint times, Tic-Tac-Toe in high dimensions, and cellphone ringtones. But there is not a single equation. I couldn't put the book down.

Caithamer, Peter, *Probability, Statistics, & Financial Math*, available from psfm@psfm.net, 2010; vi + 667 pp, $180. ISBN 978-0-9830011-0-2.

The author, a professor in a department of mathematics and actuarial science, asserts that this book can be used as a text for "6 to 8 complete courses in stochastics and finance." That claim is an exaggeration; for example, the 44 pp and 12 exercises on single-variable regression and time series are a far cry from standard courses on those topics. However, the book does provide in well-organized, concise, and easily-searched fashion the bare bones "theoretical underpinnings" (careful theorems and proofs) for the first four actuarial exams (covering probability, financial mathematics, and actuarial models) plus applied statistics. That is indeed a service, compared to the 1,000+-page practice manuals available for each exam. (There are 470 pp of exposition, examples, and exercises; 176 pp of solutions to exercises; and 20 pp of tables, references, and index. More details about the content are at http://psfm.net/.)

Grabiner, Judith V., *A Historian Looks Back: The Calculus as Algebra and Selected Writings*, MAA, 2010; xv + 287 pp, $62.95 ($49.95 for MAA members). ISBN 978-0-88385-527-0.

Here, *by one individual*, are seven(!) papers that won awards for excellence in exposition: three in this MAGAZINE (Allendoerfer Awards) and four in the *American Mathematical Monthly* (Ford Awards). (The other three essays in this volume did not appear in MAA journals, so could not qualify.) In addition, the volume includes a reprint of author Grabiner's book *The Calculus as Algebra: J. L. Lagrange, 1736–1813* (Garland, 1990). The main thread of her research traces the origins of calculus from Newton through Lagrange to Cauchy, in authoritative but understandable fashion. The essays also include "The centrality of mathematics in the history of Western thought" (worth sending a copy to your dean) and "Why should historical truth matter to mathematicians?" (worth bringing to the attention of new members of the profession).

NEWS AND LETTERS

71st Annual William Lowell Putnam Mathematical Competition

Editor's Note: Additional solutions will be printed in the *Monthly* later this year.

PROBLEMS

A1. Given a positive integer n, what is the largest k such that the numbers $1, 2, \ldots, n$ can be put into k boxes so that the sum of the numbers in each box is the same? [When $n = 8$, the example $\{1, 2, 3, 6\}$, $\{4, 8\}$, $\{5, 7\}$ shows that the largest k is *at least* 3.]

A2. Find all differentiable functions $f: \mathbb{R} \to \mathbb{R}$ such that

$$f'(x) = \frac{f(x+n) - f(x)}{n}$$

for all real numbers x and all positive integers n.

A3. Suppose that the function $h: \mathbb{R}^2 \to \mathbb{R}$ has continuous partial derivatives and satisfies the equation

$$h(x, y) = a\,\frac{\partial h}{\partial x}(x, y) + b\,\frac{\partial h}{\partial y}(x, y)$$

for some constants a, b. Prove that if there is a constant M such that $|h(x, y)| \le M$ for all (x, y) in $\mathbb{R}^2$, then $h(x, y)$ is identically zero.

A4. Prove that for each positive integer n, the number $10^{10^{10^n}} + 10^{10^n} + 10^n - 1$ is not prime.

A5. Let G be a group, with operation $*$. Suppose that

(i) G is a subset of $\mathbb{R}^3$ (but $*$ need not be related to addition of vectors);

(ii) For each $\mathbf{a}, \mathbf{b} \in G$, either $\mathbf{a} \times \mathbf{b} = \mathbf{a} * \mathbf{b}$ or $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ (or both), where $\times$ is the usual cross product in $\mathbb{R}^3$.

Prove that $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ for all $\mathbf{a}, \mathbf{b} \in G$.

A6. Let $f: [0, \infty) \to \mathbb{R}$ be a strictly decreasing continuous function such that $\lim_{x\to\infty} f(x) = 0$. Prove that $\int_0^\infty \frac{f(x) - f(x+1)}{f(x)}\,dx$ diverges.

B1. Is there an infinite sequence of real numbers $a_1, a_2, a_3, \ldots$ such that

$$a_1^m + a_2^m + a_3^m + \cdots = m$$

for every positive integer m?

Math. Mag. **84** (2011) 74–80. doi:10.4169/math.mag.84.1.074.

B2. Given that A, B, and C are noncollinear points in the plane with integer coordinates such that the distances AB, AC, and BC are integers, what is the smallest possible value of AB?

B3. There are 2010 boxes labeled $B_1, B_2, \ldots, B_{2010}$, and $2010n$ balls have been distributed among them, for some positive integer n. You may redistribute the balls by a sequence of moves, each of which consists of choosing an i and moving *exactly* i balls from box B_i into any one other box. For which values of n is it possible to reach the distribution with exactly n balls in each box, regardless of the initial distribution of balls?

B4. Find all pairs of polynomials $p(x)$ and $q(x)$ with real coefficients for which

$$p(x)\,q(x+1) - p(x+1)\,q(x) = 1.$$

B5. Is there a strictly increasing function $f : \mathbb{R} \to \mathbb{R}$ such that $f'(x) = f(f(x))$ for all x?

B6. Let A be an $n \times n$ matrix of real numbers for some $n \geq 1$. For each positive integer k, let $A^{[k]}$ be the matrix obtained by raising each entry to the k^{th} power. Show that if $A^k = A^{[k]}$ for $k = 1, 2, \ldots, n+1$, then $A^k = A^{[k]}$ for all $k \geq 1$.

SOLUTIONS

Solution to A1. The largest k is $\lfloor (n+1)/2 \rfloor$. The examples

$$\{1, n\},\ \{2, n-1\},\ \ldots,\ \{n/2, n/2+1\}$$

for n even and

$$\{n\},\ \{1, n-1\},\ \{2, n-2\},\ \ldots,\ \{(n-1)/2, (n+1)/2\}$$

for n odd show that $k = \lfloor (n+1)/2 \rfloor$ is possible.

On the other hand, in any distribution, some box must contain n, so the sum in each box is at least n, so

$$kn \leq 1 + 2 + \cdots + n = \frac{n(n+1)}{2},$$

and $k \leq (n+1)/2$. Because k is an integer, $k \leq \lfloor (n+1)/2 \rfloor$.

Solution to A2. Observe that $f(x+n) - f(x) = nf'(x)$ for all real x and all positive integers n. (We need this fact only for $n = 1$ and $n = 2$.) We have, for all x,

$$\begin{aligned} f'(x+1) &= \frac{f(x+2) - f(x+1)}{1} \\ &= (f(x+2) - f(x)) - (f(x+1) - f(x)) \\ &= 2f'(x) - f'(x) = f'(x). \end{aligned}$$

But $f'(x+1) - f'(x)$ is the derivative of the function $f(x+1) - f(x) = f'(x)$, so $f''(x)$ exists and is 0 for all x. Integrating twice shows that there exist $a, b \in \mathbb{R}$ such that $f(x) = ax + b$ for all x.

Conversely, it is easy to check that any linear function $f(x) = ax + b$ satisfies the condition in the problem.

Solution to A3. If $a = b = 0$, then $h(x, y) = 0$ for all (x, y). Otherwise, fix an arbitrary point (x_0, y_0) and define a function g of one variable by

$$g(t) = h(x_0 + at, y_0 + bt).$$

By the chain rule,

$$\begin{aligned} g'(t) &= a\,\frac{\partial h}{\partial x}(x_0 + at, y_0 + bt) + b\,\frac{\partial h}{\partial y}(x_0 + at, y_0 + bt) \\ &= h(x_0 + at, y_0 + bt) = g(t). \end{aligned}$$

Therefore, $g(t) = g(0)e^t$. But $|h(x, y)| \le M$ for all (x, y), so $|g(t)| \le M$ for all t. Thus $g(0) = 0$, and $h(x_0, y_0) = g(0) = 0$.

Solution to A4. Let N be the number, let 2^m be the highest power of 2 dividing n, and let $x = 10^{2^m}$. Then 10^{10^n} is divisible by 10^n, which is divisible by 2^n, which is greater than 2^m (because $2^n > n \ge 2^m$). Thus the first two exponents, 10^{10^n} and 10^n, are even multiples of 2^m, whereas by definition of m, the third exponent, n, is an odd multiple of 2^m. So $N = x^{2a} + x^{2b} + x^{2c+1} - 1$ for some nonnegative integers a, b, c. We have $x \equiv -1 \pmod{(x + 1)}$, so

$$N \equiv (-1)^{2a} + (-1)^{2b} + (-1)^{2c+1} - 1 = 0 \pmod{(x + 1)}.$$

Thus N is divisible by the integer $x + 1 = 10^{2^m} + 1 > 1$, but N and $x + 1$ are not congruent modulo 10, so they are not equal. Therefore, N is not prime.

Solution to A5. Suppose not, so that G is not contained in a 1-dimensional subspace. Let $\mathbf{u}$ be any element of G not equal to $\mathbf{0}$. Choose $\mathbf{v} \in G$ outside $\mathbb{R}\mathbf{u}$. Then $\mathbf{u} \times \mathbf{v} \notin \mathbb{R}\mathbf{u}$, and in particular $\mathbf{u} \times \mathbf{v} \ne \mathbf{0}$, so $\mathbf{u} \times \mathbf{v} = \mathbf{u} * \mathbf{v} \in G$. Replacing $\mathbf{v}$ with $\mathbf{u} \times \mathbf{v}$, we may assume that $\mathbf{u}$ and $\mathbf{v}$ are perpendicular.

Now $\mathbf{u} \times \mathbf{v}$ and $\mathbf{u} \times (\mathbf{u} \times \mathbf{v})$ are nonzero, so

$$\mathbf{u} \times (\mathbf{u} \times \mathbf{v}) = \mathbf{u} * (\mathbf{u} \times \mathbf{v}) = \mathbf{u} * (\mathbf{u} * \mathbf{v}) = (\mathbf{u} * \mathbf{u}) * \mathbf{v}.$$

Therefore, $(\mathbf{u} * \mathbf{u}) * \mathbf{v}$ is a negative real multiple of $\mathbf{v}$. Because $(\mathbf{u} * \mathbf{u}) \times \mathbf{v}$ cannot be a nonzero multiple of $\mathbf{v}$, it follows that $(\mathbf{u} * \mathbf{u}) \times \mathbf{v} = \mathbf{0}$. Thus $\mathbf{u} * \mathbf{u}$ is a multiple of $\mathbf{v}$.

Repeating the previous paragraph with $\mathbf{u} \times \mathbf{v}$ in place of $\mathbf{v}$ shows that $\mathbf{u} * \mathbf{u}$ is also a multiple of $\mathbf{u} \times \mathbf{v}$, so $\mathbf{u} * \mathbf{u} = \mathbf{0}$.

In particular, if the identity $\mathbf{e}$ of G is not $\mathbf{0}$, setting $\mathbf{u} = \mathbf{e}$ in the above yields $\mathbf{e} * \mathbf{e} = \mathbf{0}$, which shows that $\mathbf{e} = \mathbf{0}$ in any case.

Fix perpendicular nonzero $\mathbf{u}, \mathbf{v} \in G$ as above. Then

$$(\mathbf{u} * \mathbf{u}) * \mathbf{v} = \mathbf{e} * \mathbf{v} = \mathbf{v},$$

contradicting the fact that it is a negative multiple of $\mathbf{v}$.

Solution to A6.

First proof. For a given positive integer n, choose a positive integer k large enough that $f(n + k) \le f(n + 1)/2$. For $n \le x \le n + 1$, we have

$$f(x + k) \le f(n + k) \le f(n + 1)/2 \le f(x)/2.$$

Thus

$$\int_n^{n+k} \frac{f(x)-f(x+1)}{f(x)}\,dx = \int_n^{n+1} \sum_{j=0}^{k-1} \frac{f(x+j)-f(x+j+1)}{f(x+j)}\,dx$$

$$\geq \int_n^{n+1} \sum_{j=0}^{k-1} \frac{f(x+j)-f(x+j+1)}{f(x)}\,dx = \int_n^{n+1} \frac{f(x)-f(x+k)}{f(x)}\,dx$$

$$\geq \int_n^{n+1} \frac{f(x)-f(x)/2}{f(x)}\,dx = \frac{1}{2}.$$

We have infinitely many disjoint intervals of the form $[n, n+k]$, so the integral diverges.

Second proof, based on a student paper: Note that the integrand is positive, so if the integral were convergent, there would be a lower bound N such that

$$\int_a^{\infty} \frac{f(x)-f(x+1)}{f(x)}\,dx < \frac{1}{2} \qquad \text{for all } a \geq N.$$

This would imply, for all $a \geq N$,

$$\frac{1}{2} > \int_a^{\infty} \frac{f(x)-f(x+1)}{f(x)}\,dx > \int_a^{\infty} \frac{f(x)-f(x+1)}{f(a)}\,dx$$

$$= \frac{1}{f(a)} \lim_{A\to\infty} \int_a^{A} [f(x)-f(x+1)]\,dx$$

$$= \frac{1}{f(a)} \lim_{A\to\infty} \left[\int_a^{A} f(x)\,dx - \int_{a+1}^{A+1} f(x)\,dx\right]$$

$$= \frac{1}{f(a)} \lim_{A\to\infty} \left[\int_a^{a+1} f(x)\,dx - \int_{A}^{A+1} f(x)\,dx\right]$$

$$= \frac{1}{f(a)} \int_a^{a+1} f(x)\,dx > \frac{f(a+1)}{f(a)}.$$

That is, $f(a+1) < \frac{1}{2} f(a)$ for all $a \geq N$. But then

$$\frac{f(x)-f(x+1)}{f(x)} > \frac{\frac{1}{2}f(x)}{f(x)} = \frac{1}{2}$$

for $x \geq N$, and so the integral diverges after all.

Solution to B1. No.

First proof. If $|a_n| > 1$ for some n, then

$$m = \sum_{i=1}^{\infty} a_i^m \geq |a_n|^m > m$$

for sufficiently large even positive integers m, a contradiction. If $|a_n| \leq 1$ for all n, then

$$4 = \sum_{n=1}^{\infty} a_n^4 \leq \sum_{n=1}^{\infty} a_n^2 = 2,$$

again a contradiction.

Second proof, based on student papers: There isn't even a sequence such that the equation holds for all of $m = 2$, $m = 3$, and $m = 4$. Otherwise, we would have

$$\begin{aligned}0 &= 4 - 2 \cdot 3 + 2 \\ &= \sum_{i=1}^{\infty} a_i^4 - 2\sum_{i=1}^{\infty} a_i^3 + \sum_{i=1}^{\infty} a_i^2 \\ &= \sum_{i=1}^{\infty} a_i^2(a_i - 1)^2.\end{aligned}$$

Because all terms are nonnegative, this implies that for each i, $a_i = 0$ or $a_i = 1$, so $a_i^2 = a_i^3$. But then $2 = \sum_{i=1}^{\infty} a_i^2 = \sum_{i=1}^{\infty} a_i^3 = 3$, a contradiction.

Solution to B2. The smallest possible distance is 3.

The 3-4-5 triangle with $A = (0, 0)$, $B = (3, 0)$, $C = (0, 4)$ shows that $AB = 3$ is possible.

Now suppose that $AB < 3$. Without loss of generality, assume that $A = (0, 0)$. Because neither $x^2 + y^2 = 1^2$ nor $x^2 + y^2 = 2^2$ has a solution in positive integers, we may also assume that $B = (k, 0)$ where k is 1 or 2. If $AB = 1$, then the lengths of the other two sides differ by less than 1, so $AC = BC$, and C has x-coordinate $1/2$, a contradiction. If $AB = 2$, then AC and BC differ by less than 2; on the other hand, if $C = (m, n)$, then $AC^2 = m^2 + n^2$ and $BC^2 = (m - 2)^2 + n^2$ have the same parity, so AC and BC have the same parity. Hence $AC = BC$, so $C = (1, n)$, where $n > 0$ without loss of generality. But $AC^2 = n^2 + 1$ lies strictly between two consecutive squares n^2 and $(n + 1)^2$, a contradiction.

Solution to B3. Answer: for $n \geq 1005$.

If $n \leq 1004$, then

$$2010n < 1 + 2 + \cdots + 2009,$$

so one can place the balls so that box B_i contains at most $i - 1$ balls. Then no move is allowable, and the number of balls in box B_1 is 0, not n.

If $n \geq 1005$, the following algorithm produces the desired distribution:

1. Repeat the following until $B_2, \ldots, B_{2010}$ are all empty:
   (a) Move all balls from B_1 into some other nonempty box.
   (b) Choose $j \in \{2, \ldots, 2010\}$ such that B_j contains at least j balls.
   (c) Repeatedly move j balls from B_j to B_1 until B_j has fewer than j balls.
   (d) Move balls from B_1 to B_j until B_j has exactly j balls.
   (e) Empty B_j by moving its j balls to B_1.
2. Move n balls from B_1 to B_k for each $k \in \{2, \ldots, 2010\}$.

Explanation:

- Step 1(b) is always possible, because otherwise the total number of balls would be at most

$$0 + 1 + 2 + \cdots + 2009 = \frac{2010 \cdot 2009}{2} < 2010 \cdot 1005 \leq 2010n.$$

- Each iteration of Step 1 increases the number of empty boxes among $B_2, \dots, B_{2010}$ by 1, so eventually we reach Step 2.
- Step 2 is possible, because all the balls are in B_1 when the step starts.

Solution to B4. Answer: All solutions have the form $p(x) = a + bx, q(x) = c + dx$ with $ad - bc = 1$.

If $p(x) = p_m x^m + \cdots$ and $q(x) = q_n x^n + \cdots$, then the coefficient of x^{m+n-1} on the left-hand side works out to $(n - m)p_m q_n$, so either $m = n$ or $m + n = 1$. If $m = n$, then replacing $q(x)$ by $q(x) - \frac{q_n}{p_m} p(x)$ yields a solution with the new n less than m, but we just showed that this is impossible unless $m = 1$. Thus $m, n \le 1$ in any case. If $p(x) = a + bx$ and $q(x) = c + dx$, then the condition is

$$(a + bx)[c + d(x + 1)] - [a + b(x + 1)](c + dx) = 1,$$

which is equivalent to $ad - bc = 1$.

Solution to B5. No. Suppose f is such a function. Then we have $f' \ge 0$, so $f''(x) = f'(f(x))f'(x) \ge 0$, so f' is nondecreasing, so the formula for f'' shows that f'' is nondecreasing too. If $f''(x) = 0$ for all x, then $f(x) = ax + b$ for some a and b, and then $f'(x) = f(f(x))$ implies $a = b = 0$, contradicting the assumption that f is strictly increasing. Otherwise, if $f''(r) = s > 0$, say, then $f''(x) \ge s$ for all $x \ge r$, so there exist t, u such that $f(x) \ge (s/2)x^2 + tx + u$ for all $x \ge r$. Then there exists $k > 0$ such that $f(x) \ge kx^2$ for sufficiently large positive x. This implies that $f'(x) = f(f(x)) \ge kf(x)^2$ for sufficiently large positive x. Integrating $f'/f^2 \ge k$ shows that there exists C such that $C - 1/f(x) \ge kx$ for sufficiently large positive x. This contradicts $f(x) \ge kx^2$ for sufficiently large positive x.

Solution to B6. The hypothesis implies that $A^{[k]}A = A^{[k+1]}$ for $k = 1, \dots, n$. Taking the (i, j) entry of both sides shows that

$$a_{i1}^k a_{1j} + \cdots + a_{in}^k a_{nj} = a_{ij}^{k+1}.$$

Thus, for fixed i and j, $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$ is a solution to the linear system

$$B \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{ij}^2 \\ \vdots \\ a_{ij}^{n+1} \end{pmatrix},$$

where B is the matrix $\begin{pmatrix} a_{i1} & \cdots & a_{in} \\ & \ddots & \\ a_{i1}^n & \cdots & a_{in}^n \end{pmatrix}$. The Vandermonde determinant formula implies that distinct nonzero columns of B are linearly independent. On the other hand, the vector on the right-hand side of the system is a_{ij} times the jth column of B. Let S denote the *set* $\{a_{i1}, a_{i2}, \dots, a_{in}\} - \{0\}$. Then every solution to the system satisfies

$$\sum_{r:\, a_{ir}=s} x_r = \begin{cases} a_{ij}, & \text{if } s = a_{ij} \\ 0, & \text{if } s \ne a_{ij} \end{cases}$$

for all s in S. It follows that

$$a_{i1}^k a_{1j} + \cdots + a_{in}^k a_{nj} = \sum_{s \in S} \left(s^k \sum_{r\,:\,a_{ir}=s} a_{rj} \right) = a_{ij}^{k+1}$$

for all $k \geq 1$. Thus $A^{[k]}A = A^{[k+1]}$ for all k, and the result follows by induction.

Letter to the Editor

I read with interest Parker's Note [**2**] in the February 2010 issue of this MAGAZINE, presenting a remarkable property of the catenary, namely that the area under this curve is proportional to its length. This property already appears in the classical treatise on curves by Loria [**1**, p. 576], where it is enunciated for the cumulative area $A(x)$ and chord length $s(x)$ above an interval from the origin to an arbitrary x-value. However, if $A(x)$ and $s(x)$ are proportional (i.e., their ratio is constant), this property extends to their increments between two x-values. Therefore, the proportionality between area and length applies to any horizontal interval, as stated in the Note [**2**].

JAVIER SÁNCHEZ-REYES
University of Castilla-La Mancha, Spain
Javier.SanchezReyes@uclm.es

REFERENCES

1. G. Loria, *Spezielle Algebraische und Transzendente Ebene Kurven: Theorie und Geschichte*, Teubner, Leipzig, 1911.
2. E. Parker, A Property Characterizing the Catenary, *Math.Mag.* **83**(1) (2010) 63–64.

To appear in *College Mathematics Journal*, March 2011

Articles

Singular Vectors' Subtle Secrets, *by David James, Michael Lachance, and Joan Remski*

Faulhaber's Triangle, *by Mohammad Torabi-Dashti*

How Spherical Are the Archimedean Solids and Their Duals? *by P. K. Aravind*

The Symmedian Point: Constructed and Applied, *by Robert K. Smither*

Dynamics of Folds in the Plane, *by Nikolai A. Krylov and Edwin L. Rogers*

Lattice Cubes, *by Richard Parris*

The Symmetry Group of the Permutahedron, *by Karl-Dieter Crisman*

A Midsummer Knot's Dream, *by A. Henrich, N. MacNaughton, S. Narayan, O. Pechenik, R. Silversmith, J. Townsend*

Student Research Projects

Golden Matrix Families, *by Anne Fontaine and Susan Hurley*

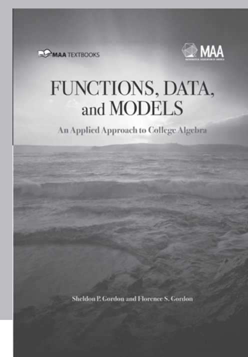
Classroom Capsules

An Application of Sylvester's Rank Inequality, *by Sidney H. Kung*

Book Reviews

Logical Labyrinths, by Raymond M. Smullyan, *reviewed by Kenneth Schilling*

Roads to Infinity, The Mathematics of Truth and Proof, by John Stillwell, *reviewed by Stan Wagon*



MAA TEXTBOOKS
MAA
FUNCTIONS, DATA, and MODELS
An Applied Approach to College Algebra
Sheldon P. Gordon and Florence S. Gordon