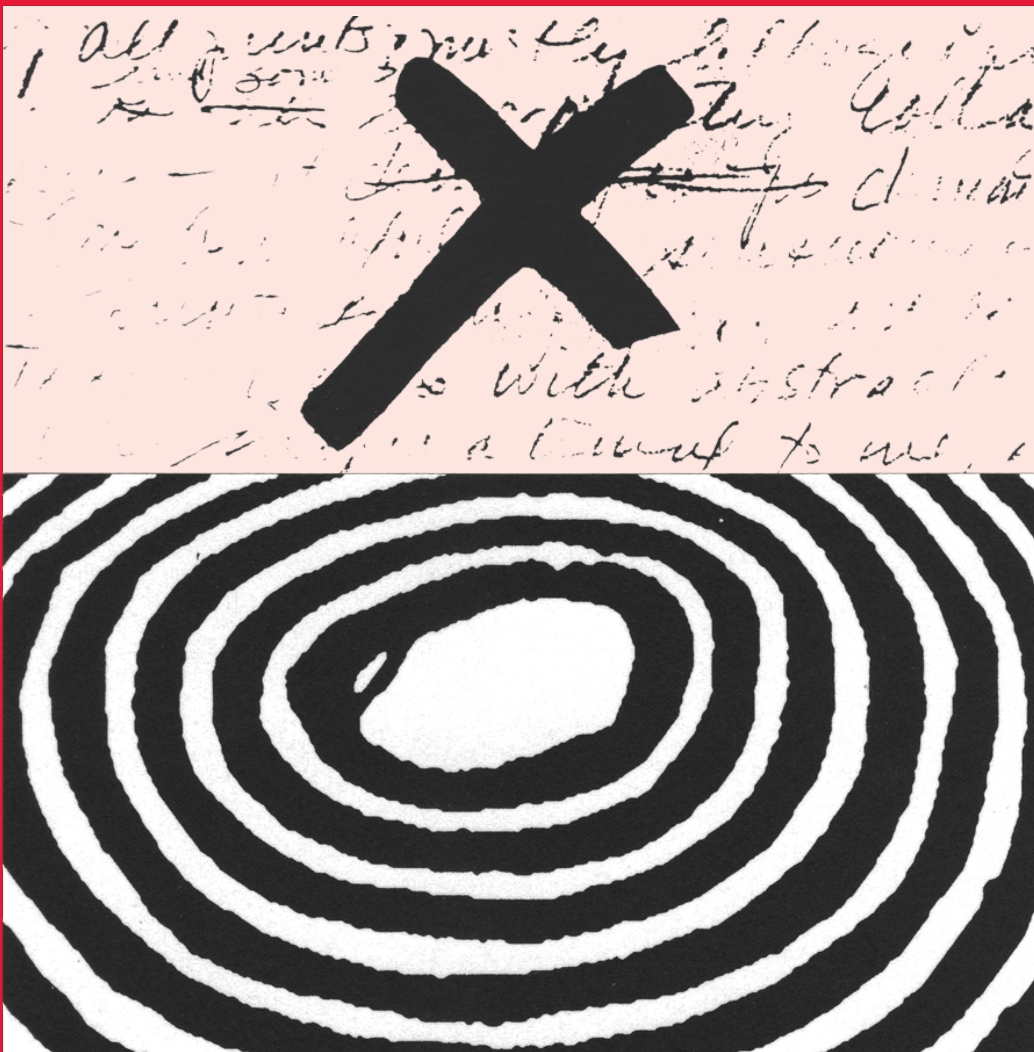
Vol. 85, No. 4, October 2012

MAA

MATHEMATICS MAGAZINE



- Fourier's 17 Lines
- Stirling Numbers of the Second Kind
- Skunk Redux

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

EDITORIAL POLICY

Mathematics Magazine aims to provide lively and appealing mathematical exposition. The *Magazine* is not a research journal, so the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for the *Magazine*. Articles should include examples, applications, historical background, and illustrations, where appropriate. They should be attractive and accessible to undergraduates and would, ideally, be helpful in supplementing undergraduate courses or in stimulating student investigations. Manuscripts on history are especially welcome, as are those showing relationships among various branches of mathematics and between mathematics and other disciplines.

A more detailed statement of author guidelines appears in this *Magazine*, Vol. 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: "Six by Six: Meditations (36)," screenprint and collage by Annie Stromquist, 2011.

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: (877) 622-2373
E-mail: tmarmor@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. 85, No. 4, October 2012

MAA

MATHEMATICS MAGAZINE

LETTER FROM THE EDITOR

When you are faced with an intractable problem, one way you can make progress is to master a particular example. To keep one's focus, it helps to choose a famous example, or one with historical significance. That's why we pay so much attention to factoring the RSA Challenge numbers, or finding Traveling Salesman routes through actual German cities. This issue's first article is about classifying arrangements of lines in a plane—a problem that can be posed in many ways, some quite intractable. So Jerry Alexanderson and Jack Wetzel have attacked a test case first posed by Fourier in 1788, and previously addressed in this MAGAZINE in 1980. Have they mastered the example? See the adventure unfold. (The article has a supplement by D. Lichtblau and W. Wichiramala; you can find it at our website.)

Next, Khristo Boyadzhiev takes us on a tour of the Stirling Numbers of the Second Kind. He traces them from Stirling's book in 1730, to their uses in Newton series, to the contributions of Grünert and Euler, and beyond. If you make these numbers your own, you'll start to see them everywhere!

Here's something different: David Kong and Peter Taylor treat us to a dialog, in which they discover the core principle of dynamic programming and apply it to the classroom game of SKUNK. As often happens, first impressions are misleading. The truth lies deeper, and the proof deeper still.

Football fans: Have you seen a sudden-death overtime yet this season? There is a new rule, designed to reduce the influence of the coin toss at the start of sudden death. Will it work? Chris Jones tells us, in the Notes Section. There is good news there, too, for those of you who have always wanted to be Brahmagupta Triangles, but lack the integrity required. Are you are at least rational? Then you can aspire to the related class of triangles described by Herb Bailey and William Gosnell. Frank Sandomierski shows us an instant proof of a matrix theorem. And, Jørn Olsson and James Sellers have given us a quick reaction to the "Remarkable Identity" that appeared in our June issue—they offer a combinatorial proof, and it turns out to involve Stirling Numbers. Did I say you would see them everywhere? But these Stirling Numbers are of the First Kind.

Speaking of quick reactions, credit is due to the committee members who have prepared the problems and solutions to this year's USAMO, USAJMO, and IMO problems.

The issue finishes with this year's Allendoerfer Award citation. The winners for 2012 are Mark Kayll and John Adam, for articles that appeared in this MAGAZINE during 2011.

Walter Stromquist, Editor

ARTICLES

Perplexities Related to Fourier's 17 Line Problem

GERALD L. ALEXANDERSON
Santa Clara University
Santa Clara, CA 95053
galexand@math.scu.edu

JOHN E. WETZEL
University of Illinois at Urbana-Champaign
Urbana, IL 61801
jewetz@comcast.net

Fourier's 17 line problem

In 1788, the 20-year-old Joseph Fourier posed the following problem in a letter to his friend and teacher C. L. Bonard, as reported by John Herivel in his 1975 biography of Fourier [**4**, p. 244]:

> Here is a little problem of a rather singular nature: it occurred to me in connection with certain propositions in Euclid we discussed on several occasions. Arrange 17 lines in the same plane so that they give 101 points of intersection. It is to be assumed that the lines extend to infinity, and that no point of intersection belongs to more than two lines.

Fourier asked for an example, and no doubt Bonard, who was Professor of Mathematics at the École Royale Militaire of Auxerre where Fourier had been a student until the previous year, was able to provide one.

There seems to have been no particular reason for the choice of the numbers 17 and 101, for in his letter Fourier went on to add, "The problem must be reduced purely to analysis so that given m and n one can arrive at the necessary equations."

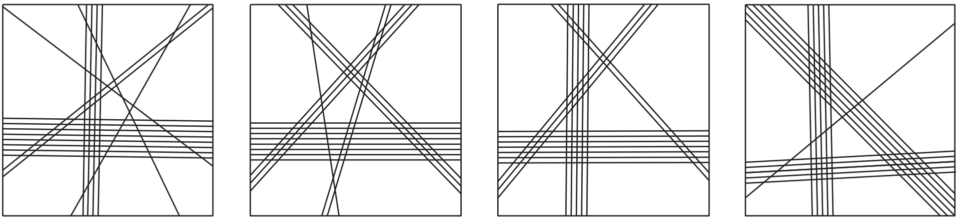
Shortly after Herivel's biography appeared, Turner [**8**] and Webster [**9**] independently found the four different arrangements pictured in FIGURE 1. (The drawings are based on those in Turner [**8**, p. 218].) The captions give the numbers of lines in the parallel families, using superscripts to code repetitions.

Although their reasoning differs in detail, both Turner and Webster suppose the 17 lines are sorted into p parallel families π_j of $\mu_j \geq 1$ lines and find that there are just four sets of possible values for $\mu_1, \mu_2, \ldots, \mu_p$ that satisfy the conditions $\sum_{j=1}^{p} \mu_j = 17$ and $\sum_{1 \leq i < j \leq p} \mu_i \mu_j = 101$, or, equivalently, $\sum_{j=1}^{p} \mu_j^2 = 87$.

Parallel equivalence. Both Turner and Webster found four different solution arrangements, but precisely what is it that there are just four of?

Let us agree to call a set of 17 lines in the plane that intersect to form 101 points a *Fourier arrangement.* Euclidean arrangements of lines having no point of intersection

Math. Mag. **85** (2012) 243–251. doi:10.4169/math.mag.85.4.243. © Mathematical Association of America



(a) $\langle 8, 4, 2, 1^3 \rangle$ (b) $\langle 8, 3^2, 2, 1 \rangle$ (c) $\langle 7, 5, 3, 2 \rangle$ (d) $\langle 6, 5^2, 1 \rangle$

Figure 1 The four solutions

lying on more than two lines are commonly called *simple* in the literature. We write $\mathcal{F}_s$ for the collection of all simple Fourier arrangements.

For the moment we regard a line having no parallel partners as a parallel family of order 1.

DEFINITION. Suppose the p parallel families π_j of a simple arrangement $\mathcal{A}$ are indexed so that $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_p \geq 1$, where $\mu_i = \#(\pi_i)$. We call the p-tuple

$$\langle \mu_1, \mu_2, \ldots, \mu_p \rangle$$

the *parallel data* of $\mathcal{A}$.

In 1826 Steiner [**7**] showed by induction that if σ_1 and σ_2 are the first two elementary symmetric functions on the parallel data $\mu_1, \mu_2, \ldots, \mu_p$, i.e., if

$$\sigma_1(\mu_1, \mu_2, \ldots, \mu_p) = \sum_{j=1}^{p} \mu_j,$$

$$\sigma_2(\mu_1, \mu_2, \ldots, \mu_p) = \sum_{1 \leq i < j \leq p} \mu_i \mu_j,$$

then the simple arrangement $\mathcal{A}$ forms

$$\begin{aligned} C &= 1 + \sigma_1 + \sigma_2 \quad \text{regions}, \\ E &= \sigma_1 + 2\sigma_2 \quad \text{segments and rays}, \\ V &= \sigma_2 \quad \text{points}. \end{aligned} \tag{1}$$

Although Steiner gave only the first of these formulas, it is convenient to call them all *Steiner's formulas.*

Now we can define the appropriate equivalence relation on the collection $\mathcal{F}_s$ of all simple Fourier arrangements.

DEFINITION. Two simple Fourier arrangements are *parallel-equivalent,* and for our purposes, *essentially the same,* if they have the same parallel data—that is, if they have the same number of parallel families having respectively the same cardinalities.

This relation is clearly an equivalence relation on $\mathcal{F}_s$, and it partitions $\mathcal{F}_s$ into equivalence classes. What Turner and Webster found is that this equivalence relation has precisely four equivalence classes.

It follows from Steiner's formulas (1) that, in addition to forming $V = 101$ points of intersection, these arrangements all form $C = 119$ regions and $E = 219$ segments, including rays.

A generalization. Fourier asked that his problem be "reduced purely to analysis so that given m and n one can arrive at the necessary equations." This would require that the parallel counters satisfy $\sum_{j=1}^{p} \mu_j = m$ and $\sum_{1\leq i<j\leq p} \mu_i \mu_j = n$, or, equivalently, $\sum_{j=1}^{p} \mu_j^2 = m^2 - 2n$, because

$$\left(\sum_{j=1}^{p} \mu_j\right)^2 = \sum_{j=1}^{p} \mu_j^2 + \left(\sum_{1\leq i<j\leq p} \mu_i \mu_j\right).$$

This seems to have provoked remarkably little response; in addition to [**8**] and [**9**], the only followup article of which we are aware is Woeginger [**12**], who recast the question as a decision problem:

> Fourier's general problem: For given positive integers S and Q, decide whether there exist positive integers p and $\mu_1, \mu_2, \ldots, \mu_p$ so that $\sum_{j=1}^{p} \mu_j = S$ and $\sum_{j=1}^{p} \mu_j^2 = Q$.

Woeginger's principal result is a polynomial time algorithm for determining whether such integers exist for given S and Q. He also opines (in §4) that the problem of enumerating such solution tuples is likely to be NP-hard.

Two related 17 line problems

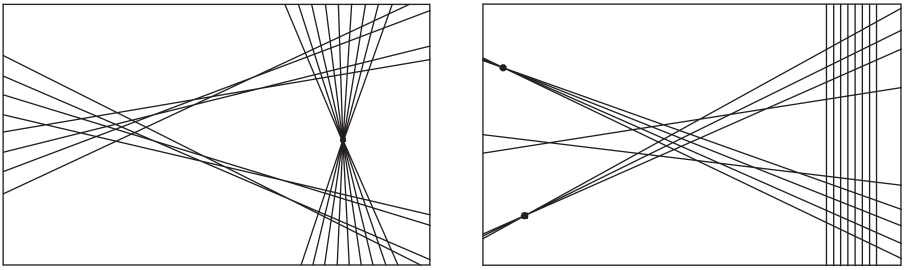
More than 30 years ago, in an unfortunate moment of idle curiosity, we wondered what happens if multiple oints—points that lie on more than just two lines—are permitted in Fourier's 17 line problem. There are, then, two natural questions that parallel Fourier's challenge to Bonard:

1. (No parallels) Arrange 17 lines in the plane so as to form 101 points of intersection, assuming that no two lines are parallel. How many essentially different such arrangements are there?
2. (No restrictions) Arrange 17 lines in the plane so as to form 101 points of intersection, but impose no restrictions on the number of multiple points or the number of parallel families. How many essentially different such arrangements are there?

In each case we agree that a multiple point contributes just one point to the point count. FIGURE 2(a) shows an example of a "parallel free" Fourier arrangement having a multiple point of order (multiplicity) 9, as demanded by Question 1; and FIGURE 2(b) shows an example with multiple points of orders 3 and 4 and a parallel family of order 8, as sought in Question 2.

The enumeration part of each question requires an understanding as to when two arrangements are to be regarded as "essentially the same." We simply agree to call two arrangements "essentially the same" if they have the same numbers of multiple points having the same orders and the same numbers of parallel families having the same orders. Without question this minimal notion of "sameness" has many weaknesses, some of which we examine in the last section, but it seems appropriate as a start.

Although it was convenient when we considered Steiner's formulas to count a singleton line—that is, a line having no parallel partners—as a parallel family, this language is inappropriate in much of what follows. *Let us agree from now on that "a*



(a) Parallel-free Fourier arrangement

(b) Fourier arrangement with both parallels and multiple points

Figure 2 Two Fourier arrangements

parallel" is a line having at least one parallel partner, and "a parallel family" has at least two parallel lines.

In 1982 we announced our findings on these two questions in an abstract [**1**]. We found that there are just 20 essentially different Fourier arrangements having no parallels, and 900 essentially different Fourier arrangements having at least one parallel family and at least one multiple point, for a total of 924 unrestricted Fourier arrangements.

Clearly both Question 1 and Question 2 can be asked for m lines and n points. We know of no literature whatsoever on these more general problems.

Roberts' formulas. Formulas based on subtracting off the various parts that are not formed because the lines of the arrangement fail to be in general position were given in 1889 by Roberts [**6**]. This paper, an amazing example of "intuitive" geometric reasoning, is concerned in large part with the heuristic development of formulas for the number of cells, faces, edges, and vertices that are formed by a completely arbitrary arrangement of planes in 3-space. More recently, a definitive algebraic investigation of similar face-count formulas for arrangements in d-dimensional Euclidean and projective spaces was given in 1975 by Zaslavsky [**13**]. We shall need Roberts' formula in the plane for the number of points of intersection.

First we recall the well-known formulas for arrangements of n lines in "general position," i.e., simple arrangements having no parallel pairs. These formulas have been known at least since the early nineteenth century and may well date back another century or more:

$$C = 1 + n + \binom{n}{2} = \frac{1}{2}(n^2 + n + 2) \quad \text{regions},$$

$$E = n + 2\binom{n}{2} = n^2 \quad \text{segments and rays},$$

$$V = \binom{n}{2} = \frac{1}{2}(n^2 - n) \quad \text{points}.$$

They are immediate consequences of Steiner's formulas (1) and are readily proved by induction. Other proofs are possible; see, for example, [**2**] and [**3**].

Roberts' formulas in the plane are the following, in which we agree that empty sums (sums from 1 to 0) have the value 0. Suppose the n lines of an arrangement $\mathcal{A}$ in the plane form m points $P_1, P_2, \ldots, P_m$ of orders $\lambda_1, \lambda_2, \ldots, \lambda_m$ and p families of parallel lines of orders $\mu_1, \mu_2, \ldots, \mu_p$. Then $\mathcal{A}$ forms

$$C = 1 + n + \binom{n}{2} - \sum_{i=1}^{m} \left[\binom{\lambda_i - 1}{2}\right] - \sum_{j=1}^{p} \binom{\mu_j}{2} \quad \text{regions,}$$

$$E = n^2 - \sum_{i=1}^{m} \lambda_i (\lambda_i - 2) - 2 \sum_{j=1}^{p} \binom{\mu_j}{2} \quad \text{segments and rays,} \tag{2}$$

$$V = \binom{n}{2} - \sum_{i=1}^{m} \left[\binom{\lambda_i}{2} - 1\right] - \sum_{j=1}^{p} \binom{\mu_j}{2} \quad \text{points.}$$

An elaborate heuristic discussion of these formulas together with accessible proofs can be found in [**11**].

Question 1: No parallels. Let $\mathcal{A}$ be a Fourier arrangement having no parallel families. For each point P, let $\lambda(P)$ be the number of lines of $\mathcal{A}$ that pass through P. If $\lambda(P) > 2$, then we call P a *multiple* point of *order* $\lambda(P)$. Since $\binom{17}{2} > 101$, the Fourier arrangement $\mathcal{A}$ must have at least one multiple point. We write m for the number of multiple points formed by $\mathcal{A}$, and we index the multiple points $P_1, P_2, \ldots, P_m$ so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 3$, where λ_i is the order of P_i. It will be convenient to call the m-tuple $\langle \lambda_1, \lambda_2, \ldots, \lambda_m \rangle$ the *point data* of the parallel-free arrangement $\mathcal{A}$.

In the absence of parallels, Roberts' formula for V requires that

$$\sum_{i=1}^{m} \left[\binom{\lambda_i}{2} - 1\right] = \binom{17}{2} - 101 = 35. \tag{3}$$

Evidently the multiple point of largest order must lie on no more than 9 lines, and if $\lambda_1 = 9$, no other multiple points are possible. An example is pictured in FIGURE 2(a). It follows from Roberts' formulas that this arrangement forms $C = 126$ regions and $E = 226$ segments and rays, and, of course, $V = 101$ points.

Suppose next that $\lambda_1 = 8$. Formula (3) then demands

$$\sum_{i=2}^{m} \left[\binom{\lambda_i}{2} - 1\right] = 35 - \left[\binom{8}{2} - 1\right] = 8,$$

and the only possibility is $m = 5$ and $\lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = 3$. Such an arrangement has point data $\langle 8, 3^4 \rangle$, and an example is pictured in FIGURE 3(a). It forms $C = 129$ cells and $E = 229$ edges and, of course, $V = 101$ points.

Examining successively the cases with λ_i equal to 7, 6, 5, and 4 (we omit the repetitive and tedious details), we find the results collected in TABLE 1, which shows for each arrangement the largest order, the point data, and the numbers C and E of cells and edges it forms. Plainly $\lambda_1 = 3$ is impossible, because then Formula (3) would demand that 35 be even.

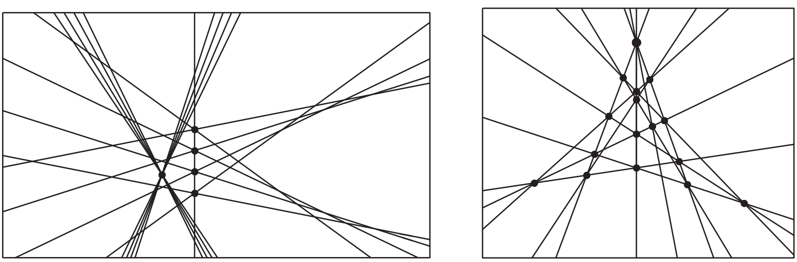
Solution to Question 1. It follows that the answer to the enumeration part of Question 1 is that there are just 20 equivalence classes of arrangements of 17 lines that form 101 points of intersection, as we announced [**1**].

Somewhat remarkably, our investigations have shown that apart from the first, $\langle 9 \rangle$, each entry listed in the table can be realized by a suitable Fourier arrangement in which every line lies on at least one multiple point. FIGURES 3(a) and 3(b) show such arrangements for the second entry, $\langle 8, 3^4 \rangle$, and for the last entry, $\langle 4, 3^{15} \rangle$. Finding examples of the other cases would be a challenging exercise for the interested reader.

TABLE 1: Realizable Steiner data

	λ_1	Point data	C	E
1	9	$\langle 9\rangle$	126	226
2	8	$\langle 8, 3^4\rangle$	129	229
3	7	$\langle 7, 5, 3^3\rangle$	130	230
4		$\langle 7, 4^3\rangle$	130	230
5		$\langle 7, 4, 3^5\rangle$	131	231
6	6	$\langle 6^2, 4, 3\rangle$	130	230
7		$\langle 6, 5, 3^6\rangle$	132	232
8		$\langle 6, 4^3, 3^3\rangle$	132	232
9		$\langle 6, 4, 3^8\rangle$	133	233
10		$\langle 6, 5, 4^2, 3\rangle$	131	231
11	5	$\langle 5^3, 3^4\rangle$	132	232
12		$\langle 5^2, 4^3, 3\rangle$	132	232
13		$\langle 5^2, 4, 3^6\rangle$	133	233
14		$\langle 5, 4^4, 3^3\rangle$	133	233
15		$\langle 5, 4^2, 3^8\rangle$	134	234
16		$\langle 5, 3^{13}\rangle$	135	235
17	4	$\langle 4^7\rangle$	133	233
18		$\langle 4^5, 3^5\rangle$	134	234
19		$\langle 4^3, 3^{10}\rangle$	135	235
20		$\langle 4, 3^{15}\rangle$	136	236

Question 2: No restrictions. Similar methods can be employed for Question 2, but the details are more complicated. Having already considered the cases in which there are no multiple points and no parallel families, we turn next to arrangements having both. We write $\mathcal{F}^*$ for the collection of all Fourier arrangements having both multiple points and parallel families, and as before we introduce an equivalence relation in $\mathcal{F}^*$



(a) $\langle 8, 3^4\rangle$ (b) $\langle 4, 3^{15}\rangle$

Figure 3 Two parallel-free Fourier arrangements by agreeing to call arrangements $\mathcal{A}_1$ and $\mathcal{A}_2$ *essentially the same* if they have the same numbers of multiple points and parallel families of respectively the same orders.

Let $\mathcal{A}$ be an arrangement in $\mathcal{F}^*$, with multiple points of orders $\lambda_1, \lambda_2, \ldots, \lambda_m$ and parallel families of orders $\mu_1, \mu_2, \ldots, \mu_p$, both of which we assume are indexed in nonincreasing order: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 3$ and $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_p \geq 2$. We call the $(m+p)$-tuple

$$\sigma = \langle \lambda_1, \lambda_2, \ldots, \lambda_m \mid \mu_1, \mu_2, \ldots, \mu_p \rangle \tag{4}$$

the *data tuple of the Fourier arrangement* $\mathcal{A}$, and we observe that the entries in σ satisfy the conditions $1 \leq m \leq 17$, $1 \leq p \leq 8$, $8 \geq \lambda_1 \geq \lambda_m \geq 3$, and $8 \geq \mu_1 \geq \mu_p \geq 2$.

The converse question, whether a given $(m+p)$-tuple (4) whose entries satisfy the above inequalities is the data tuple of some Fourier arrangement, is a difficult one.

Necessary conditions for realizability. Three necessary conditions help. First, if σ is to be the data tuple of a realizable arrangement, then Roberts' formula (2) for V must hold:

$$\sum_{i=1}^{m} \left[\binom{\lambda_i}{2} - 1 \right] + \sum_{j=1}^{p} \binom{\mu_j}{2} = 35. \tag{5}$$

Second, there being only 17 lines available,

$$\sum_{j=1}^{p} \mu_j \leq 17. \tag{6}$$

Third, at most one line from each of the p families of parallels together with the $17 - \sum_{j=1}^{p} \mu_j$ lines lacking parallel partners are available to form a multiple point, so the multiplicity λ_1 of the largest multiple point must satisfy

$$\lambda_1 \leq p + 17 - \sum_{j=1}^{p} \mu_j. \tag{7}$$

Independent computer-assisted counts by W. Wichiramala and D. Lichtblau show that there are 2195 tuples that satisfy (5). Of these, just 901 also satisfy (6). And only one of these 901 fails test (7)! Indeed, Lichtblau found that just $\delta = \langle 6 \mid 4^2, 3^3 \rangle$ failed test (7); since there are only 5 parallel families and no lines lacking parallel partners, there are not enough lines to build a multiple point of order greater than 5.

Solution to Question 2. Thirty years ago we announced (in [**1**]) that just 900 data tuples are realizable, so that our answer to the enumeration part of Question 2 was $900 + 4 + 20 = 924$. Although we believed the question too special to justify publishing the details, we had, and we continue to have, considerable confidence in our count of 900 realizable tuples, which we found by writing and running numerous small FORTRAN programs to check a variety of specific configuration-based sufficient conditions, followed by hand-checking the remaining cases. The details of that work have been long since discarded.

With considerable confidence, however, we repeat our contention that there are just 924 unrestricted Fourier arrangements, and we leave the re-verification of this count as a challenge for others.

Much additional information about the above computer counts is reported in the article supplement [**5**], "Perplexed calculations," by D. Lichtblau and W. Wichiramala, which can be found at this MAGAZINE's website.

Trouble, trouble, trouble

Throughout we have taken the point of view that two Fourier arrangements should be regarded as "essentially the same" if they have the same number of multiple points of the same order and the same number of parallel families of the same order. This seems reasonable at first glance, but how well do these "minimal" requirements capture the intuitive meaning of "essentially the same"? Not very well, it seems to us.

The difficulty is apparent already in the initial Fourier challenge to Bonard, which may be why Fourier asked only for an example. Suppose, for example, the three lines lacking parallel partners in FIGURE 1(a) are relocated so as to form a triangle having all of the 56 points in which the parallel families of orders 8, 4, and 2 intersect in its interior. Does one want to call that Fourier arrangement "essentially the same" as the one pictured in FIGURE 1(a)? We wonder.

Even the enumeration question for lines in general position ("In how many different ways can n lines in general position be arranged?") is unresolved, according to West [**10**].

The situation is even more perplexing for parallel-free arrangements. Call a line in such an arrangement *lonely* if it does not pass through any multiple points. Then the number s of lonely lines needed to bring the line count up to 17 depends not only on the number and orders of the multiple points, but also on their configuration. For example, FIGURE 4 shows schematically nine Fourier arrangements having point data $\langle 8, 3^4 \rangle$ that require various numbers of lonely lines. Evidently, different configurations of the same multiple points can require different numbers of lonely lines, and different configurations of the same multiple points can require the same numbers of lonely lines. It seems clear that arrangements in which the multiple points are configured differently should somehow be distinguished in any intuitively reasonable enumeration, and simply incorporating the lonely line count into the definition of the equivalence relation would not accomplish this result.

We confess that we know of no way systematically to enumerate all the possible distinct configurations of the multiple points and the number of lonely lines they require. Clearly the situation is even more confusing if parallel families are permitted. At this point we throw up our hands and admit defeat.

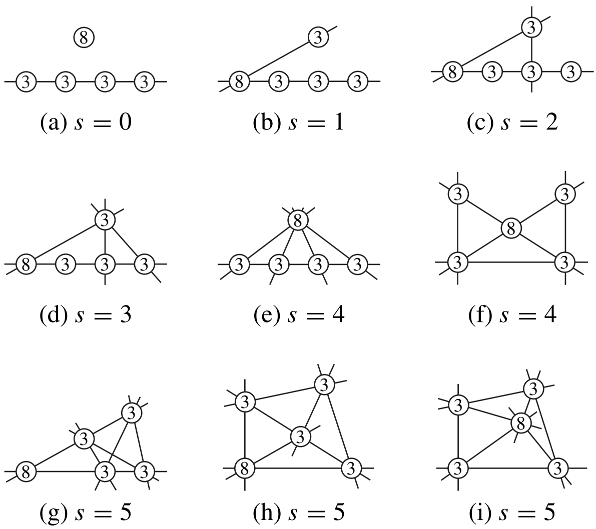


Figure 4 Some lonely line possibilities for $\langle 8, 3^4 \rangle$

Conclusion

Until someone formulates a truly intuitively satisfactory definition of "essentially the same" for arrangements of 17 lines intersecting in exactly 101 points—something we seem quite unable to do—a satisfactory solution of all these Fourier problems will remain a tantalizing mystery. Somewhere the ghost of Fourier is chuckling.

Acknowledgments We are grateful to Wacharin Wichiramala (Chulalongkorn University, Bangkok, Thailand) and Daniel Lichtblau (Wolfram Research, Champaign, Illinois) for their significant contributions to the enumeration part of Question 2.

REFERENCES

1. G. L. Alexanderson and John E. Wetzel, On a problem of Fourier, *Abstracts Amer. Math. Soc.* **3** (1982) 361.
2. G. L. Alexanderson and John E. Wetzel, Simple partitions of space, *Math. Mag.* **51** (1978) 220–225. http://dx.doi.org/19300980
3. Bro. U. Alfred ⟨Brousseau⟩, A mathematician's progress, *Math. Teacher* **59** (1966) 722–727.
4. John Herivel, *Joseph Fourier: The Man and the Physicist*, Clarendon Press, Oxford, 1975.
5. Daniel Lichtblau and Wacharin Wichiramala, Perplexed calculations. http://www.maa.org/pubs/mm_supplements/index.html
6. Samuel Roberts, On the figures formed by the intercepts of a system of straight lines in a plane, and on analogous relations in space of three dimensions, *Proc. London Math. Soc.* **19** (1889) 405–422. http://dx.doi.org/10.1112/plms/s1-19.1.405
7. Jacob Steiner, Einige Gesetze über die Theilung der Ebene und des Raumes, *J. reine angew. Math.* **1** (1826) 349–364. http://dx.doi.org/10.1515/crll.1826.1.349
8. Barbara Turner, Fourier's seventeen lines problem, *Math. Mag.* **53** (1980) 217–219. http://dx.doi.org/10.2307/2689614
9. R. J. Webster, A problem of Fourier, *Math. Gazette* **64** (1980) 270–271. http://dx.doi.org/10.2307/3616728
10. Douglas West, personal communication, September 13, 2011.
11. John E. Wetzel, On the division of the plane by lines, *Amer. Math. Monthly* **85** (1978) 647–656. http://dx.doi.org/10.2307/2320333
12. Gerhard J. Woeginger, Seventeen lines and one-hundred-and-one points, *Theoret. Comput. Sci.* **321** (2004) 415–421. http://dx.doi.org/10.1016/j.tcs.2004.04.006
13. Thomas Zaslavsky, Facing up to arrangements: Face-count formulas for partitions of space by hyperplanes, *Mem. Amer. Math. Soc. No. 154*, Amer. Math. Soc., Providence, RI, 1975.

Summary In 1788 a young Joseph Fourier posed to his teacher the following little problem in geometry: arrange 17 lines in the plane so as to form 101 points of intersection, assuming there are no "multiple points"; that is, no more than two lines are concurrent. It is not difficult to show that there are just four essentially different such arrangements.

In this note we recap results we found 30 years ago and show more generally that if multiple points are permitted but no two lines are allowed to be parallel, then there are 20 essentially different such arrangements; and if both parallels and multiple points are permitted, there are 924 essentially different such arrangements.

The perplexing question is what, exactly, is meant for two arrangements of lines to be *essentially the same*. We consider this matter briefly and conclude that we do not know how to formulate a suitable definition.

GERALD L. ALEXANDERSON is the Valeriote Professor of Science at Santa Clara University. He is a former Secretary and President of the Mathematical Association of America and a former editor of this MAGAZINE. Long interested in combinatorics and geometry, he and Wetzel have collaborated on some problems in those areas over many years. He is also known for collections of interviews and profiles of mathematicians, coauthored with Donald J. Albers, in a series of books on mathematical people, the most recent addition to the list being *Fascinating Mathematical People* (Princeton, 2011). For fun he reads books produced with ink on paper from his sizable collection of antiquarian books and ostentatiously abjures the possibility of owning a Kindle or a Nook.

JOHN E. WETZEL spent his childhood in Goshen, Indiana, did his undergraduate work and earned his B.S. at Purdue, and did his graduate work and earned his Ph.D. in mathematics at Stanford. Employed by the University of Illinois, he retired as Professor Emeritus in 1999 after 38 years of service. He and his wife Rebecca share a condo in Champaign with a miscellany of hippopotamic *objets d'art,* including an aggregation of some 75 hippo cookie jars. Although his graduate work was in complex analysis, he has always been interested in problems in classical geometry, most recently in the ways in which one shape can fit into another.

Close Encounters with the Stirling Numbers of the Second Kind

KHRISTO N. BOYADZHIEV
Ohio Northern University
Ada, Ohio 45810
k-boyadzhiev@onu.edu

Let $n = 4$, and consider the terms in row n of Pascal's triangle, with alternating signs: $(1, -4, 6, -4, 1)$. Treat this list as a vector and take its scalar product with a vector of consecutive integer squares:

$$(1, -4, 6, -4, 1) \cdot (0, 1, 4, 9, 16) = 0.$$

Next, try cubes:

$$(1, -4, 6, -4, 1) \cdot (0, 1, 8, 27, 64) = 0.$$

So far, this is getting us nothing. Vectors of first powers and of zeroth powers also give scalar products of zero. We get something more when we try fourth powers:

$$(1, -4, 6, -4, 1) \cdot (0, 1, 16, 81, 256) = 24,$$

which is equal to $n!$.

These are all instances of the strange evaluation,

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k k^m = \begin{cases} 0 & \text{if } m < n, \\ (-1)^n n! & \text{if } m = n, \end{cases} \tag{1}$$

which we see from time to time in books and articles. For example, Katsuura [**12**] gives the following theorem, extending (1) a little bit: For any two real or complex numbers x, y and for any positive integers m and n,

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k (xk + y)^m = \begin{cases} 0 & \text{if } m < n, \\ (-1)^n x^n n! & \text{if } m = n. \end{cases} \tag{2}$$

From this form we see that the mth powers do not need to be of consecutive integers; the identity holds for the mth powers of consecutive terms in any arithmetic sequence. This, indeed, is a very strange result!

Is it just a curious fact, or is there something bigger behind it? Also, what happens when $m > n$? The theorem obviously deserves further elaboration. Therefore, we want to fill this gap now and also to provide some related historical information.

Identity (2) is not new. It appears in a more general form in H. W. Gould's *Combinatorial Identities* [**7**]. Namely, if $f(t) = c_0 + c_1 t + \cdots + c_m t^m$ is a polynomial of degree m, then Gould's entry (Z.8) says that

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k f(k) = \begin{cases} 0 & \text{if } m < n, \\ (-1)^n n!\, c_n & \text{if } m = n, \end{cases} \tag{3}$$

Math. Mag. **85** (2012) 252–266. doi:10.4169/math.mag.85.4.252. © Mathematical Association of America

which implies (2). As Gould writes on p. 82: "Relation (Z.8) is very useful; we have numerous interesting cases by choosing $f(t)$ ...." Identity (2) was later rediscovered by Ruiz [**16**], who proved it by induction.

Here is a simple observation (made also by Katsuura)—expanding the binomial $(xk + y)^m$ in (2) and changing the order of summation, we find that (2) is based on the more simple identity (1) (which, by the way, is entry (1.13) in [**7**]). In his paper [**8**] Gould provides a nice and thorough discussion of identity (1), calling it *Euler's formula,* as it appears in the works of Euler on nth differences of powers. See also Schwatt's book [**17**, pp. 18–19, 48].

The mystery of identity (1) is revealed by its connection to the Stirling numbers. An old result in classical analysis (also discussed in [**8**]) says that

$$(-1)^n n!\, S(m,n) = \sum_{k=0}^{n} \binom{n}{k} (-1)^k k^m, \tag{4}$$

where $S(m,n)$ are the *Stirling numbers of the second kind* [**2, 5, 8, 9, 11, 19**]. They have the property $S(m,n) = 0$ when $m < n$, and $S(m,m) = 1$. We can define these numbers in combinatorial terms: $S(m,n)$ counts the number of ways to partition a set of m elements into n nonempty subsets. Thus we can read $S(m,n)$ as "m subset n." An excellent combinatorial treatment of the Stirling numbers can be found in [**9**]. Pippenger's recent article in this MAGAZINE [**14**] mentions, among other things, their probabilistic interpretation.

A simple combinatorial argument ([**9**, p. 259]) provides the important recurrence

$$S(m,n) = nS(m-1,n) + S(m-1,n-1), \tag{5}$$

valid for $m > 0$ and for all integers n, which together with the initial conditions $S(0,0) = 1$ and $S(0,n) = 0$ for $n \neq 0$, gives an alternative definition for these numbers. Using this recurrence we can compute

$$\begin{aligned} S(n+1,n) &= nS(n,n) + S(n,n-1) \\ &= n + S(n,n-1) \\ &= n + (n-1) + S(n-1,n-2) \\ &= n + (n-1) + (n-2) + \cdots + 1 = \frac{n(n+1)}{2}; \end{aligned}$$

i.e., $S(n+1,n) = \frac{n(n+1)}{2}$; and now from (4) we find

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k k^{n+1} = \frac{(-1)^n n}{2}(n+1)!,$$

which extends (1) to the case of $m = n+1$. A proof of (5) using finite differences is presented in [**8**].

The alternative notation

$$S(m,n) = \left\{ {m \atop n} \right\}$$

suggested in 1935 by the Serbian mathematician Jovan Karamata (see [**9**, p. 257]) fits very well with the combinatorial interpretation. With this notation, the recursion (5) becomes

$$\left\{ {m \atop n} \right\} = n \left\{ {m-1 \atop n} \right\} + \left\{ {m-1 \atop n-1} \right\} \quad (n \leq m),$$

which parallels the well-known property of binomial coefficients

$$\binom{m}{n} = \binom{m-1}{n} + \binom{m-1}{n-1}.$$

With the help of equation (4) we can fill the gap in (2) for $m > n$. Namely, we have

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k (xk+y)^m = (-1)^k n! \sum_{j=n}^{m} \binom{m}{j} x^j y^{m-j} S(j,n).$$

For the proof we just need to expand $(xk+y)^m$, change the order of summation, and apply (4).

Identity (3) has a short and nice extension beyond polynomials. The representation

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k f(k) = (-1)^n n! \sum_{m=0}^{\infty} c_m S(m,n) \tag{6}$$

is true for any $n \geq 0$ and any function $f(t) = c_0 + c_1 t + \cdots$ that is analytic on a disk with radius $R > n$. To prove this, we multiply (4) by c_m and sum for m from zero to infinity.

A combinatorial proof of (4) based on the combinatorial definition of $S(m,n)$ can be found in [**5**, pp. 204–205]; a proof based on finite differences is given in [**11**, p. 169; see also 177–178 and 189–190]. We shall present here two proofs of (1) and (4). For the first one we shall visit the birthplace of the Stirling numbers.

James Stirling and his table

The name "Stirling numbers" comes from the Danish mathematician Niels Nielsen (1865–1931). On p. 68 of his book [**13**] Nielsen attributed these numbers to James Stirling, a Scottish mathematician (1692–1770), who worked on Newton series.

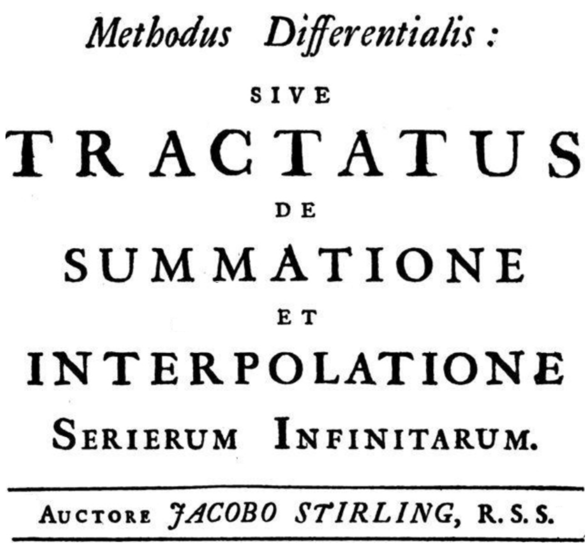

Methodus Differentialis:
SIVE
TRACTATUS
DE
SUMMATIONE
ET
INTERPOLATIONE
SERIERUM INFINITARUM.
AUCTORE *JACOBO STIRLING*, R. S. S.

Figure 1 Part of the front page of Stirling's book

Stirling studied at Oxford, then went to Italy for political reasons and almost became a professor of mathematics in Venice. In 1718 he published through Newton a paper titled *Methodus Differentialis Newtoniana Illustrata.* In 1725 Stirling returned to England and in 1730 published his book *Methodus Differentialis* (The Method of Differences) [**18**]. The book was written in Latin, as were most scientific books of that time. An annotated English translation was published recently by Ian Tweddle [**19**].

At that time mathematicians realized the importance of series expansion of functions, and various techniques were gaining momentum. A *Newton series* is an expansion of a function, say f, in terms of the *difference polynomials,* $P_0(z) = 1$, $P_1(z) = z$, $P_2(z) = z(z-1)$, $P_3(z) = z(z-1)(z-2)$, and in general $P_k(z) = z(z-1)\cdots(z-k+1)$. That is,

$$\begin{aligned} f(z) &= \sum_{k=0}^{\infty} a_k z(z-1)(z-2)\cdots(z-k+1) \\ &= a_0 + a_1 z + a_2 z(z-1) + a_3 z(z-1)(z-2) + \cdots . \end{aligned} \tag{7}$$

The difference polynomials are also called *falling powers,* and they are a basis for the space of polynomials. In this way a Newton series resembles a Taylor series, which is an expansion of f in terms of another basis, the *power polynomials* $p_k(z) = z^k$, $k = 0, 1, \ldots$. The attention paid to both series raised the question of the relationships between the difference polynomials and the power polynomials.

At the beginning of his book Stirling studied carefully the coefficients A^n_m in the representations

$$\begin{aligned} z^m &= A^m_1 z + A^m_2 z(z-1) + A^m_3 z(z-1)(z-2) \\ &\quad + \cdots + A^m_m z(z-1)\cdots(z-m+1) \end{aligned} \tag{8}$$

where $m = 1, 2, \ldots$. On p. 8 he presented a table containing many of these coefficients, reproduced here as FIGURE 2.

Tabulam priorem.

1	1	1	1	1	1	1	1	1	&c.
	1	3	7	15	31	63	127	255	&c.
		1	6	25	90	301	966	3025	&c.
			1	10	65	350	1701	7770	&c.
				1	15	140	1050	6951	&c.
					1	21	266	2646	&c.
						1	28	461	&c.
							1	36	&c.
								1	&c.
									&c.

Figure 2 Stirling's first table

In the table m changes horizontally, left to right, and n changes vertically, from top to bottom. Therefore, by following the columns of the table we find

$$\begin{aligned} z &= z, \\ z^2 &= z + z(z-1), \\ z^3 &= z + 3z(z-1) + z(z-1)(z-2), \\ z^4 &= z + 7z(z-1) + 6z(z-1)(z-2) + z(z-1)(z-2)(z-3), \\ z^5 &= z + 15z(z-1) + 25z(z-1)(z-2) + 10z(z-1)(z-2)(z-3) \\ &\quad + z(z-1)(z-2)(z-3)(z-4), \text{ etc.} \end{aligned}$$

The coefficients A_k^m are exactly the numbers which we call today Stirling numbers of the second kind. For completeness, we add to this sequence also $A_0^0 = 1$ and $A_0^m = 0$ when $m > 0$. The following is true.

THEOREM 1. *Let the coefficients A_n^m be defined by the expansion* (8). *Then*

$$A_n^m = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} k^m, \tag{9}$$

and the right side is zero when $n > m$.

It is not obvious that (8) implies (9). For the proof of the theorem we need some preparation.

Stirling's technique for computing this table is presented on pp. 24–29 in Ian Tweddle's translation [**19**]. As Tweddle comments on p. 171, had Stirling known the recurrence relation (5), the computation of the table would have been much easier.

Newton series and finite differences

The theory of Newton series, like (7), also resembles the theory of Taylor series. First of all, one needs to find a formula for the coefficients a_k. In the case of Taylor series, the function f is expanded on the power polynomials and the coefficients are expressed in terms of the higher derivatives of the function evaluated at zero. In the case of Newton series, instead of derivatives, one needs to use finite differences. This is suggested by the very form of the series, as the function is expanded now on the difference polynomials.

For a given function $f(z)$ we set

$$\Delta f(z) = f(z+1) - f(z).$$

Then

$$\begin{aligned} \Delta^2 f(z) &= \Delta(\Delta f)(z) = f(z+2) - 2f(z+1) + f(z), \\ \Delta^3 f(z) &= \Delta(\Delta^2 f)(z) = f(z+3) - 3f(z+2) + 3f(z+1) - f(z), \text{ etc.} \end{aligned}$$

We notice the binomial coefficients appearing here with alternating signs. Following this pattern we arrive at the representation

$$\Delta^n f(z) = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} f(z+k).$$

In particular, with $z = 0$,

$$\Delta^n f(0) = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} f(k). \tag{10}$$

We use this formula now to compute the coefficients a_k in the Newton series (7). With $z = 0$ we see that $a_0 = f(0)$. A simple computation shows that $\Delta P_k(z) = k P_{k-1}(z)$, $k \geq 1$ and so

$$\Delta f(z) = a_1 + 2a_2 z + 3a_3 z(z-1) + \cdots ,$$

which yields $a_1 = \Delta f(0)$. Also,

$$\Delta^2 f(z) = 2a_2 + 2 \cdot 3a_3 z + 3 \cdot 4z(z-1) + \cdots$$

and $2a_2 = \Delta^2 f(0)$. Continuing this way we find $3!\, a_3 = \Delta^3 f(0)$, $4!\, a_4 = \Delta^4 f(0)$, etc. The general formula is $k!\, a_k = \Delta^k f(0)$, $k = 0, 1, \ldots$. Thus (7) becomes

$$f(z) = \sum_{k=0}^{\infty} \frac{\Delta^k f(0)}{k!} z(z-1)(z-2)\cdots(z-k+1). \tag{11}$$

Proof of Theorem 1. Take $f(z) = z^m$ in (11) to obtain (in view of (10))

$$z^m = \sum_{n=0}^{\infty} \left\{ \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} k^m \right\} z(z-1)(z-2)\cdots(z-n+1). \tag{12}$$

Comparing this to (8) yields the representation (9). Note also that the series (12) truncates, as on the left-hand side we have a polynomial of degree m. The summation on the right-hand side stops with $n = m$, and

$$\frac{1}{m!} \sum_{k=0}^{m} \binom{m}{k} (-1)^{m-k} k^m = 1. \qquad \blacksquare$$

This is not the end of the story, however. In the *curriculum vitae* of the Stirling numbers there is another remarkable event.

Grünert's polynomials

Amazingly, the same Stirling numbers appeared again, one hundred years later, in a very different setting. They appeared in the work [**10**] of the German mathematician Johann August Grünert (1797–1872), professor at the University of Greifswald, Germany. He taught there from 1833 until his death. Grünert, a student of Pfaff and Gauss, was interested in many topics, not only in mathematics, but also in physics. He wrote a number of books on such diverse subjects as conic sections, the loxodrome, optics, and the solar eclipse. Some of his books, including *Optische Untersuchungen* (Studies in Optics) and *Theorie der Sonnenfinsternisse* (Theory of the Solar Eclipse) are available now as Google books on the Internet. In 1841 Grünert started to edit and publish the highly respected *Archiv der Mathematik und Physik* (known also as "Grünert's Archiv"). His biography, written by his student Maximus Curtze, appeared in volume 55 of that journal.

Grünert came to the numbers $S(m,n)$ by repeatedly applying the operator $x\frac{d}{dx}$ to the exponential function e^x. This procedure generates a sequence of polynomials

$$x\frac{d}{dx}e^x = xe^x,$$
$$\left(x\frac{d}{dx}\right)^2 e^x = (x^2+x)e^x,$$
$$\left(x\frac{d}{dx}\right)^3 e^x = (x^3+3x^2+x)e^x,$$
$$\vdots$$
$$\left(x\frac{d}{dx}\right)^m e^x = (B_0^m + B_1^m x + B_2^m + \cdots + B_m^m x^m)e^x, \tag{13}$$

with certain coefficients B_k^m. We shall see that these coefficients are exactly the Stirling numbers of the second kind. This fact follows from the theorem below.

THEOREM 2. (GRÜNERT) *Let the coefficients* B_n^m *be defined by equation* (13). *Then*

$$B_n^m = \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}(-1)^{n-k}k^m. \tag{14}$$

Proof. From the expansion

$$e^x = \sum_{k=0}^{\infty}\frac{x^k}{k!},$$

we find

$$\left(x\frac{d}{dx}\right)^m e^x = \sum_{k=0}^{\infty}\frac{k^m x^k}{k!},$$

for $m = 0, 1, \ldots$, and from (13),

$$\begin{aligned} B_0^m + B_1^m x + B_2^m x^2 + \cdots + B_m^m x^m &= e^{-x}\sum_{k=0}^{\infty}\frac{k^m x^k}{k!} \\ &= \left\{\sum_{j=0}^{\infty}\frac{(-1)^j x^j}{j!}\right\}\left\{\sum_{k=0}^{\infty}\frac{k^m x^j}{k!}\right\}. \end{aligned} \tag{15}$$

Multiplying the two power series on the right-hand side yields

$$B_0^m + B_1^m x + B_2^m x^2 + \cdots + B_m^m x^m = \sum_{n=0}^{\infty} x^n \left\{\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}(-1)^{n-k}k^m\right\},$$

and again, comparing coefficients we see that the series on the right-hand side is finite and (14) holds. The theorem is proved! ■

From (14), (9), and (4) we conclude that $A_n^m = B_n^m = S(m,n)$.

Intermediate summary and the exponential polynomials

We summarize the story so far.

The coefficients A_n^m defined by the representation (8) are the same as the coefficients B_n^m defined by equation (13), and also the same as the Stirling numbers of the second kind $S(m, n)$:

$$\begin{aligned} A_n^m = B_n^m &= S(m, n) \\ &= \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} k^m. \end{aligned}$$

Also, $S(m, n) = 0$ when $m < n$, and $S(n, n) = 1$. In particular, this proves (1).

The Stirling numbers of the second kind are used in combinatorics, often with the notation $S(m, n) = \left\{ {m \atop n} \right\}$. The number $S(m, n)$ gives the number of ways by which a set of m elements can be partitioned into n nonempty subsets. Thus $\left\{ {m \atop n} \right\}$ is naturally defined for $n \le m$ and $\left\{ {n \atop n} \right\} = 1$. When $m < n$, $\left\{ {m \atop n} \right\} = 0$. The numbers $\left\{ {m \atop n} \right\}$ equal A_n^m because they satisfy (4) as proven in [**5**].

The polynomials

$$\phi_n(x) = S(n, 0) + S(n, 1)x + \cdots + S(n, n)x^n,$$

$n = 0, 1, \dots,$ appearing in Grünert's work, are called *exponential polynomials*. They have been rediscovered and used by several authors. These polynomials are defined by equation (13), i.e.,

$$\phi_n(x) = e^{-x} \left(x \frac{d}{dx} \right)^n e^x \tag{16}$$

or by the generating function (see [**2**]),

$$e^{x(e^t - 1)} = \sum_{n=0}^{\infty} \frac{\phi_n(x)}{n!} t^n.$$

Here are the first five of them.

$$\begin{aligned} \phi_0(x) &= 1 \\ \phi_1(x) &= x \\ \phi_2(x) &= x^2 + x \\ \phi_3(x) &= x^3 + 3x^2 + x \\ \phi_4(x) &= x^4 + 6x^3 + 7x^2 + x \end{aligned}$$

A short review of these polynomials is given in [**2**]. Replacing x by ax in (16), where a is any constant, we see that (16) can be written as

$$\left(x \frac{d}{dx} \right)^n e^{ax} = \phi_n(ax) e^{ax}. \tag{17}$$

This form is useful in some computations.

The exponential generating function for $S(m, n)$

For any integer $n \geq 0$, let us expand the function $f(x) = (e^x - 1)^n$ in a Taylor series about $x = 0$ (i.e., a Maclaurin series):

$$f(x) = \sum_{m=0}^{\infty} \frac{f^{(m)}(0)}{m!} x^m.$$

For this purpose we first write

$$(e^x - 1)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} e^{kx},$$

and then, according to (14) or (4) we compute $f^{(m)}(0)$,

$$\left(\frac{d}{dx}\right)^m (e^x - 1)^n \bigg|_{x=0} = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} k^m = n!\, S(m, n).$$

Therefore,

$$\frac{1}{n!}(e^x - 1)^n = \sum_{m=0}^{\infty} S(m, n) \frac{x^m}{m!}. \tag{18}$$

This is the exponential generating function for the Stirling numbers of the second kind $S(m, n)$. The summation, in fact, can be limited to $m \geq n$, as $S(m, n) = 0$ when $m < n$. Equation (18) is often used as the definition of $S(m, n)$.

Euler and the derivatives game

Let $|x| < 1$. We want to show that the numbers $S(m, n)$ naturally appear in the derivatives

$$\left(x \frac{d}{dx}\right)^m \frac{1}{1 - x} = \sum_{n=0}^{\infty} n^m x^n, \tag{19}$$

where $m = 0, 1, \ldots$. To show this we first write

$$\frac{1}{1 - x} = \int_0^\infty e^{-(1-x)t}\, dt = \int_0^\infty e^{xt} e^{-t}\, dt.$$

Then, in view of (17),

$$\begin{aligned} \left(x \frac{d}{dx}\right)^m \frac{1}{1 - x} &= \int_0^\infty \phi_m(xt) e^{xt} e^{-t}\, dt \\ &= \sum_{n=0}^{m} S(m, n) x^n \int_0^\infty t^n e^{-(1-x)t}\, dt \\ &= \sum_{n=0}^{m} S(m, n) x^n \frac{n!}{(1 - x)^{n+1}} \\ &= \frac{1}{1 - x} \sum_{n=0}^{m} S(m, n) n! \left(\frac{x}{1 - x}\right)^n. \end{aligned} \tag{20}$$

For the third equality we use the well-known formula (which defines the Laplace transform of t^α)

$$\frac{\Gamma(\alpha+1)}{s^{\alpha+1}} = \int_0^\infty t^\alpha e^{-st}\,dt.$$

Introducing the polynomials

$$\omega_m(z) = \sum_{n=0}^{m} S(m,n)n!\,z^n,$$

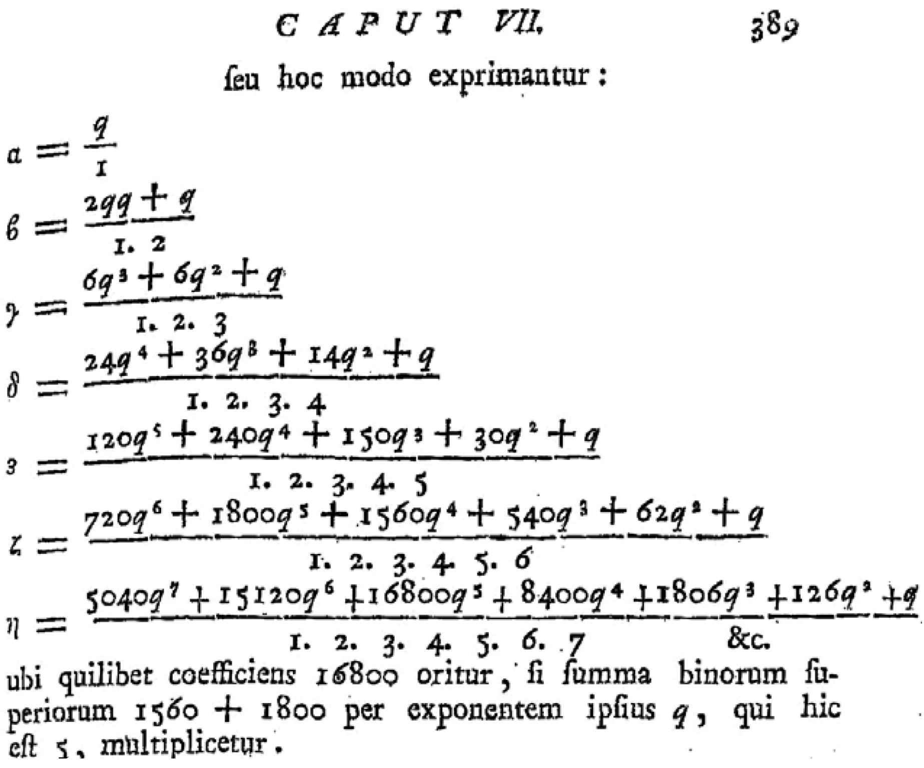
we can write (20) in the form

$$\sum_{n=0}^{\infty} n^m x^n = \frac{1}{1-x}\omega_m\left(\frac{x}{1-x}\right). \tag{21}$$

Thus we have

$$\omega_0(x) = 1,$$
$$\omega_1(x) = x,$$
$$\omega_2(x) = 2x^2 + x,$$
$$\omega_4(x) = 24x^4 + 36x^3 + 14x^2 + x, \text{ etc.}$$

The polynomials ω_n can be seen on p. 389, in Part 2, Chapter VII of Euler's book [**6**].


CAPUT VII. 389

ſeu hoc modo exprimantur :

$$\alpha = \frac{q}{1}$$
$$\beta = \frac{2qq + q}{1.\ 2}$$
$$\gamma = \frac{6q^3 + 6q^2 + q}{1.\ 2.\ 3}$$
$$\delta = \frac{24q^4 + 36q^3 + 14q^2 + q}{1.\ 2.\ 3.\ 4}$$
$$\varepsilon = \frac{120q^5 + 240q^4 + 150q^3 + 30q^2 + q}{1.\ 2.\ 3.\ 4.\ 5}$$
$$\zeta = \frac{720q^6 + 1800q^5 + 1560q^4 + 540q^3 + 62q^2 + q}{1.\ 2.\ 3.\ 4.\ 5.\ 6}$$
$$\eta = \frac{5040q^7 + 15120q^6 + 16800q^5 + 8400q^4 + 1806q^3 + 126q^2 + q}{1.\ 2.\ 3.\ 4.\ 5.\ 6.\ 7}$$
&c.

ubi quilibet coefficiens 16800 oritur, ſi ſumma binorum ſuperiorum 1560 + 1800 per exponentem ipſius q, qui hic eſt 5, multiplicetur.

Figure 3 The geometric polynomials in Euler's work

Essentially, Euler obtained these polynomials by computing the derivatives (19) directly. We shall see now how all this can be done in terms of exponentials.

Here is a good exercise. Let us expand the function

$$f(t) = \frac{1}{\mu e^{\lambda t} + 1}$$

in Maclaurin series (λ, μ are two parameters). We need to find the higher derivatives of f at zero. Assuming for the moment that $|\mu e^{\lambda t}| < 1$ we use the expansion

$$\frac{1}{\mu e^{\lambda t} + 1} = \frac{1}{1 - (-\mu e^{\lambda t})} = \sum_{n=0}^{\infty} (-\mu)^n e^{\lambda t n}$$

From this

$$\left(\frac{d}{dt}\right)^m \frac{1}{\mu e^{\lambda t} + 1} = \lambda^m \sum_{n=0}^{\infty} (-\mu)^n n^m e^{\lambda t n}$$

and in view of (21)

$$\frac{1}{\mu e^{\lambda t} + 1} = \frac{1}{\mu e^{\lambda t} + 1} \omega_m \left(\frac{-\mu e^{\lambda t}}{\mu e^{\lambda t} + 1}\right),$$

so that

$$\left(\frac{d}{dt}\right)^m \frac{1}{\mu e^{\lambda t} + 1}\Bigg|_{t=0} = \frac{\lambda^m}{\mu + 1} \omega_m \left(\frac{-\mu}{\mu + 1}\right), \tag{22}$$

which yields the desired representation

$$\frac{1}{\mu e^{\lambda t} + 1} = \frac{1}{\mu + 1} \sum_{m=0}^{\infty} \lambda^m \omega_m \left(\frac{-\mu}{\mu + 1}\right) \frac{t^m}{m!}.$$

In particular, with $\lambda = \mu = 1$,

$$\frac{2}{e^t + 1} = \sum_{m=0}^{\infty} \omega_m \left(\frac{-1}{2}\right) \frac{t^m}{m!} = \sum_{m=0}^{\infty} \left\{ \sum_{n=0}^{m} S(m, n) n! \frac{(-1)^n}{2^n} \right\} \frac{t^m}{m!}. \tag{23}$$

The polynomials ω_m appeared in the works of Euler, but they do not carry his name. In [**8**] and [**17**], they are used to evaluate the series on the right-hand side of (19) in terms of Stirling numbers. These polynomials were studied in [**4**] and called *geometric polynomials,* because of their obvious relation to the geometric series. It was shown in [**4**] that ω_m participate in a certain *series transformation formula.* In [**3**] the geometric polynomials were used to compute the derivative polynomials for $\tan x$ and $\sec x$.

One can write (21) in the form

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

where A_m are polynomials of degree m. These polynomials are known today as Eulerian polynomials and their coefficients are the Eulerian numbers [**5, 9**].

At the same time, there is a sequence of interesting and important polynomials carrying the name *Euler polynomials.* These are the polynomials $E_m(x)$, $m = 0, 1, \ldots,$ defined by the generating function

$$\frac{2e^{xt}}{e^t + 1} = \sum_{m=0}^{\infty} E_m(x) \frac{t^m}{m!}. \tag{24}$$

Using (23) we write, as in (15)

$$\begin{aligned}\frac{2e^{xt}}{e^t + 1} &= \left\{ \sum_{n=0}^{\infty} \frac{x^n t^n}{n!} \right\} \left\{ \sum_{k=0}^{\infty} \omega_k \left(\frac{-1}{2} \right) \frac{t^k}{k!} \right\} \\ &= \sum_{m=0}^{\infty} \frac{t^m}{m!} \left\{ \sum_{k=0}^{m} \binom{m}{k} \omega_k \left(\frac{-1}{2} \right) x^{m-k} \right\}. \end{aligned} \tag{25}$$

Comparing (24) and (25) yields

$$E_m(x) = \sum_{k=0}^{m} \binom{m}{k} \omega_k \left(\frac{-1}{2} \right) x^{m-k}$$

with

$$\begin{aligned} E_m(0) &= \omega_m \left(\frac{-1}{2} \right) \\ &= \sum_{n=0}^{m} S(m, n) n! \, \frac{(-1)^n}{2^n}. \end{aligned}$$

Relation to Bernoulli numbers

The Bernoulli numbers B_m, $m = 0, 1, \ldots,$ can be defined by the generating function

$$\frac{t}{e^t - 1} = \sum_{m=0}^{\infty} B_m \frac{t^m}{m!}, \quad |t| < 2\pi \tag{26}$$

[**1, 5, 9**]. From this

$$B_m = \left(\frac{d}{dt} \right)^m \frac{t}{e^t - 1} \bigg|_{t=0}. \tag{27}$$

It is tempting to evaluate these derivatives at zero by using the Leibnitz rule for the product $t \cdot \frac{1}{e^t - 1}$ and formula (22) with $\mu = -1$, $\lambda = 1$. This will not work, though, because the denominator $\mu + 1$ on the right-hand side becomes zero. To find a relation between the Bernoulli and Stirling numbers we shall use a simple trick and the generating function (18). Writing $t = \ln e^t = \ln(1 + (e^t - 1))$ we have for t small enough

$$\begin{aligned} \frac{t}{e^t - 1} &= \frac{\ln(1 + (e^t - 1))}{e^t - 1} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n + 1} (e^t - 1)^n \end{aligned}$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} \left\{ n! \sum_{m=n}^{\infty} S(m,n) \frac{t^m}{m!} \right\}$$

$$= \sum_{m=0}^{\infty} \frac{t^m}{m!} \left\{ \sum_{n=0}^{m} (-1)^n \frac{n!}{n+1} S(m,n) \right\}.$$

Comparing this to (26) we find for $m = 0, 1, \dots,$

$$B_m = \sum_{n=0}^{m} (-1)^n \frac{n!}{n+1} S(m,n). \tag{28}$$

Sums of powers

The Bernoulli numbers historically appeared in the works of the Swiss mathematician Jacob Bernoulli (1654–1705) who evaluated sums of powers of consecutive integers [**1, 9**]

$$1^m + 2^m + \cdots + (n-1)^m = \frac{1}{m+1} \sum_{k=0}^{m} \binom{m+1}{k} B_k n^{m+1-k},$$

for any $m \geq 0$, $n \geq 1$. This is the famous Bernoulli formula. It is interesting to see that sums of powers can also be evaluated directly in terms of Stirling numbers of the second kind. In order to do this, we invert the representation (4); i.e.,

$$n^m = \sum_{k=0}^{n} \binom{n}{k} S(m,k) k!. \tag{29}$$

This inversion is a property of the binomial transform [**15**]. Given a sequence $\{a_k\}$, its *binomial transform* $\{b_k\}$ is the sequence defined by

$$b_n = \sum_{k=0}^{n} \binom{n}{k} a_k, \tag{30}$$

and the inversion formula is

$$a_n = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} b_k.$$

Next, from (29),

$$1^m + 2^m + \cdots + n^m = \sum_{p=1}^{n} \left\{ \sum_{k=0}^{p} \binom{p}{k} S(m,k) k! \right\}$$

$$= \sum_{k=0}^{n} S(m,k) k! \left\{ \sum_{p=k}^{n} \binom{p}{k} \right\},$$

by changing the order of summation. Now using the well-known identity

$$\sum_{p=k}^{n} \binom{p}{k} = \binom{n+1}{k+1},$$

we finally obtain

$$1^m + 2^m + \cdots + n^m = \sum_{k=0}^{n} \binom{n+1}{k+1} S(m,k)k!,$$

which is the desired representation.

Stirling numbers of the first kind

Inverting equation (8) we have

$$z(z-1)\cdots(z-m+1) = \sum_{k=0}^{m} s(m,k)z^k$$

where the coefficients $s(m,k)$ are called *Stirling numbers of the first kind.* The following inversion property is true.

$$\sum_{k=0}^{\infty} S(m,k)s(k,n) = \delta_{mn} = \begin{cases} 0 & m \neq n, \\ 1 & m = n. \end{cases}$$

The coefficients here come from the representation ($m = 1, 2, \ldots$)

$$\frac{1}{z^{m+1}} = \sum_{k=0}^{\infty} \frac{\sigma(m+k,m)}{z(z+1)\cdots(z+m+k)},$$

following the columns of the table. The numbers $\sigma(m,k)$ are called today the Stirling cycle numbers or the unsigned Stirling numbers of the first kind [**5, 9**]. An often-used notation is

$$\sigma(m,k) = \left[\begin{matrix} m \\ k \end{matrix} \right].$$

Tabula poſterior.

1									
1	1								
2	3	1							
6	11	6	1						
24	50	35	10	1					
120	274	225	85	15	1				
720	1764	1624	735	175	21	1			
5040	13068	13132	6769	1960	322	28	1		
40320	109584	105056	67284	22449	4536	546	36	1	
&c.	&c.	&c.	&c.	&c.	&c.	&c.	&c.	&c.	&c.

Figure 4 Stirling's Second Table

We have

$$s(m,k) = (-1)^{m-k}\sigma(m,k).$$

Stirling's book [**18, 19**] contains a second table showing the values of $\sigma(m,k)$; see FIGURE 4. More properties, combinatorial interpretation, details, and generating functions can be found in the excellent books [**5, 9, 11**].

REFERENCES

1. Tom M. Apostol, A primer on Bernoulli numbers and polynomials, *Math. Mag.* **81** (2008) 178–190.
2. Khristo N. Boyadzhiev, Exponential polynomials, Stirling numbers and evaluation of some gamma integrals, *Abstract and Applied Analysis* **2009** (2009). http://dx.doi.org/10.1155/2009/168672
3. ———, Derivative polynomials for tanh, tan, sech and sec in explicit form and some applications, *Fibonacci Quarterly* **45**(4) (2007) 291–303.
4. ———, A series transformation formula and related polynomials, *Int. J. Math. Math. Sc.* **23** (2005) 3849–3866. http://dx.doi.org/10.1155/IJMMS.2005.3849
5. Louis Comtet, *Advanced Combinatorics*, Kluwer, Dordrecht, 1974.
6. Leonardo Eulero, Institutiones calculi differentialis cum ejus usu in analysi finitorum ac doctrina serierum, Impensis academiae imperialis scientiarum Petropolitanae, 1755. Also, another edition, Ticini: in typographeo Petri Galeatii superiorum permissu, 1787. (Opera Omnis Ser. I (Opera Math.), Vol. X, Teubner, 1913). Online at http://www.math.dartmouth.edu/~euler/pages/E212.html
7. Henry W. Gould, *Combinatorial Identities*, published by the author, Revised edition, 1972.
8. ———, Euler's formula for the *n*th differences of powers, *Amer. Math. Monthly* **85** (1978) 450–467. http://dx.doi.org/10.2307/2320064
9. Ronald L. Graham, Donald E. Knuth, and Oren Patashnik, *Concrete Mathematics*, Addison-Wesley, New York, 1994.
10. Johann A. Grünert, Uber die Summerung der Reihen..., *J. Reine Angew. Math.* **25** (1843) 240–279. http://dx.doi.org/10.1515/crll.1843.25.240
11. Charles Jordan, *Calculus of Finite Differences*, Chelsea, New York, 1950 (First edition: Budapest 1939).
12. Hidefumi Katsuura, Summations involving binomial coefficients, *College Math. J.* **40** (2009) 275–278. http://dx.doi.org/10.4169/193113409X458723
13. Niels Nielsen, *Die Gammafunktion*, Chelsea, New York, 1965.
14. Nicholas Pippenger, The hypercube of resistors, asymptotic expansions, and preferential arrangements, *Math. Mag.* **83** (2010) 331–346.
15. John Riordan, *Combinatorial Identities*, Robert E. Krieger, 1979.
16. Sebastián Martín Ruiz, An algebraic identity leading to Wilson's theorem, *The Math. Gazette* **80** (1996) 579–582. http://dx.doi.org/10.2307/3618534
17. I. J. Schwatt, *An Introduction to the Operations with Series*, Chelsea, New York, 1924.
18. Jacobo Stirling, *Methodus Differentialis: sive Tractatus de Summatione et Interpolatione Serierum Infinitarum*, Londini, 1730.
19. Ian Tweddle, *James Stirling's Methodus Differentialis: An Annotated Translation of Stirling's Text*, Springer, New York, 2003.

Summary This is a short introduction to the theory of Stirling numbers of the second kind $S(m,k)$ from the point of view of analysis. It is written as an historical survey centered on the representation of these numbers by a certain binomial transform formula. We tell the story of their birth in the book *Methodus Differentialis* (1730) by James Stirling, and show how they mature in the works of Johann Grünert. The paper demonstrates the usefulness of these numbers in analysis. In particular, they appear in several differentiation and summation formulas. The reader can also see the connection of $S(m,k)$ to Bernoulli numbers, to Euler polynomials, and to power sums.

KHRISTO BOYADZHIEV is a Professor of mathematics at Ohio Northern University. He received his masters and Ph.D. degrees in mathematics from the University of Sofia, Bulgaria. His main interests are in classical analysis and geometry.

Skunk Redux

DAVID KONG
Queen's University
Kingston, Ontario, Canada
x.kong@queensu.ca

PETER D. TAYLOR
Queen's University
Kingston, Ontario, Canada
peter.taylor@queensu.ca

Peter Taylor opens his linear algebra course at Queen's University by having the students play and analyze a simple dice game called Skunk Redux. This is a variation of a common game known as Skunk or Pig. The dialogue below is an account of what happened last fall, when David, one of Peter's students, asked some intriguing questions which prompted the two of them to wrestle with an unexpected problem.

Skunk is a dice game played in elementary classrooms to illustrate the fundamentals of probability [**1**]. Players are given a table with the letters SKUNK across the top like this:

S	K	U	N	K

Each column is used to record the results from one of the five identical rounds. Several players play simultaneously. The objective is to have the highest cumulative *payoff* (the sum of the payoffs from the five rounds) at the end of the game. This is how points are earned in each round:

1. At the beginning of the round, you stand.
2. Two dice are thrown.
3. If at least one 1 appears, the round is over and you have payoff 0. Otherwise you begin with a *score* equal to the total showing on the dice.
4. If you wish, you may sit down. If you do, your payoff is your score.
5. Otherwise, the dice are thrown again.
6. If at least one 1 appears, the round is over and your payoff for the round is 0.
7. Otherwise you add to your score the total showing on the dice. This gives you a new, larger score.
8. Go back to 4.

Eventually a 1 appears and the round is over.

For example, for the sequence of rolls (2, 5), (4, 2), (6, 1), if you sit after the first roll you get payoff 7, if you sit after the second roll you get payoff 13, and if you stay standing for the third roll you get payoff 0.

Extensive work has been done on variations of this game, most notably a 2-person game where an optimal strategy must take into account the opponent's score and strate-

Math. Mag. **85** (2012) 267–276. doi:10.4169/math.mag.85.4.267.

gies [**2**]. For example, the player in second place would likely use a riskier strategy than the player in first place.

This paper is concerned with a 1-person game—one person and a single round. Our analysis focuses on optimizing the expected payoff for this single round—hence, "Skunk Redux."

The first day of class

PETER: I open my linear algebra course with this game because it creates a fun environment, generates a lively discussion, and encapsulates many of the important concepts in the course—strategy, probability, movement between states, taking an average, and so on.

After playing something like 10 rounds in a row, I have the students average their 10 payoffs. This serves as an estimate of "average payoff per game." Naturally there is a tendency to see who receives the highest payoff, or more precisely, once we start talking about strategies, what *strategy* receives the highest payoff.

I want to get a good class discussion going about the different types of strategies. First of all, what is a strategy? It is a rule that tells you whether to sit or stand in any situation. Any situation in this game can be specified in terms of two variables: the number of times the dice have been rolled, and the current score.

I find that students have differing opinions on how to make use of these two variables. Some strategies are highly intuitive and students sit when they "feel" the time has come. Some sit after a certain number of rolls while others pay attention only to the score (e.g., "sit when I get above 25"). Still others use a mixture (e.g., "sit after 25 or after the third roll, whichever comes first").

DAVID: Surely the number of rolls is irrelevant and should not be a factor in any optimal strategy. The rolls are independent events! The only quantity of relevance is the current score.

PETER: David is, of course, correct. But this issue always generates a fascinating and surprising debate. A number of students will argue quite vociferously that if the dice have been thrown, say, ten times without showing a 1, the chances are increased that a 1 will appear on the next roll.

Moving on, we restrict attention to strategies that take account only of the current score. Such a strategy must specify, for each possible score s, whether you should stand or sit.

DAVID: Let's begin by defining s as your current score. If you decide to sit, your payoff will be s. If you decide to stand, your score will be either better or worse. If, on average, your new score is greater than s, you should stand for the next roll; if it is less, you should sit.

To calculate your average new score, note that with probability $25/36$ (see TABLE 1 below) your score increases by the dice sum and with probability $11/36$ your score drops to 0. Now the average dice sum, given that a 1 does not appear, is 8. (This is nicely seen in TABLE 1 by pairing each entry with its mirror image in the diagonal of 8's.) The average new score from standing is then:

$$\frac{11}{36}(0) + \frac{25}{36}(s + 8).$$

TABLE 1: Addition to the score for each of the 36 possibilities

	1	2	3	4	5	6
1						
2		4	5	6	7	8
3		5	6	7	8	9
4		6	7	8	9	10
5		7	8	9	10	11
6		8	9	10	11	12

You should remain standing when this exceeds s, and that happens when

$$25(s+8) > 36s$$
$$s < 200/11 \approx 18.2.$$

Thus you should remain standing as long as $s \leq 18$ and sit when $s \geq 19$.

Peter started rolling the dice on that first day of class. As usual, I did not bring anything to class, not even a calculator, so I had to ballpark it. "How much would I be willing to risk to get an average reward of 8?" Somehow I came up with the number 20, which in hindsight was fairly close to the actual answer. From there, I rigorously abided by my strategy, sitting when the score surpassed that critical value. It took some willpower not to allow my emotions to steer me toward the standard freshman crowd—the eternal optimists who luckily see the world as their oyster, untainted by the rationality I sometimes wish I could do away with. There were times when I would begrudgingly sit from the sidelines while the most risk-friendly participants racked up unimaginable sums. But in the long haul, my strategy paid off.

First day of class and already an interesting (yet accessible) problem. I was truly excited for university. What I did not realize at this point was that I was soon to be led to something even more interesting.

The assignment

PETER: For their first assignment, I usually give the students an extension of the game to analyze. For example, the dice may be replaced by a few coins. One of my favorite (and most demanding) extensions has been the following:

Suppose that, before each roll, you are able to specify the number of dice that are to be rolled, and you can change this number from roll to roll based on your score. As before, the round is over with zero payoff, if you are standing and *any* of the dice show a 1. A strategy must now specify, for each score s, whether to remain standing and if so, how many dice to use. Find the optimal strategy.

DAVID: Now that's an enticing problem! Rolling more dice at a time will help you increase your score more quickly, but it also increases the probability of rolling a 1. The key difference between this problem and the simpler one is that now there are two decisions to make for each value of s—whether to remain standing, and if so, how many dice to roll. But I expected the solution not to be much different than before.

PETER: Like David, most students find this problem challenging. Not many manage to come up with a good argument. But there are always a few students who produce the following solution and for some years I have always accepted it as being correct. It is based on the idea that we employed in the solution for the original game, that the correct decision at each step is the one that maximizes the expected new score.

The $A(n)$ strategy. Let $A(n)$ be your expected new score if you stay standing and choose to roll n dice. Note that the probability of not throwing a 1 is $(5/6)^n$ and (as above) the average outcome on a single die is 4. Then:

$$A(n) = (5/6)^n(s + 4n).$$

DAVID: I got the above equation for $A(n)$ without much difficulty. Now the problem was to find the maximum value of $A(n)$. When in doubt, a first-year student differentiates. The result was correct enough but it was ugly with logarithms and decimals. A bit later, I found a much nicer *algebraic* solution. I thought of it as discrete maximization, and it worked beautifully. The idea was that for $A(n)$ to be a maximum at a particular n, it must be at least as great as the neighboring $A(n)$ values, $A(n-1)$ and $A(n+1)$.

$$A(n-1) \leq A(n) \geq A(n+1)$$

The first part is:

$$\begin{aligned} A(n-1) &\leq A(n) \\ (5/6)^{n-1}(s + 4(n-1)) &\leq (5/6)^n(s + 4n) \\ s + 4(n-1) &\leq (5/6)(s + 4n) \\ 4n &\leq 24 - s \end{aligned}$$

and the same for the second:

$$\begin{aligned} A(n) &\geq A(n+1) \\ (5/6)^n(s + 4n) &\geq (5/6)^{n+1}(s + 4(n+1)) \\ (s + 4n) &\geq (5/6)s + (5/6)4(n+1) \\ 4n &\geq 20 - s \end{aligned}$$

Putting these together, the condition for a maximum $A(n)$ is that

$$20 - s \leq 4n \leq 24 - s$$

For example, given the score $s = 10$, there is only one integer value ($n = 3$) that satisfies this inequality.

PETER: David's analysis so far is the one I have always accepted, and posted on the website for the class. It says that $4n$ has to be between $20 - s$ and $24 - s$. We can summarize this condition with TABLE 2. When s is a multiple of 4, there are two values of n that give the same average score. [At $s = 20$, the "other" value is $n = 0$, which means *sit*.] And by the way, it can easily be verified directly for $s < 20$, that $A(n) > s$ for the indicated n, signifying that you gain on average by standing.

And then David came up to me after class…

DAVID: I had the solution outlined above and it seemed really elegant (isn't that table beautiful?) but it worried me. Maximizing $A(n)$ only maximizes the score *after* the

TABLE 2: The $A(n)$ strategy: Roll n dice with score s

score s	roll n dice
$s = 0$	$n = 6$
$0 \leq s \leq 4$	$n = 5$
$4 \leq s \leq 8$	$n = 4$
$8 \leq s \leq 12$	$n = 3$
$12 \leq s \leq 16$	$n = 2$
$16 \leq s \leq 20$	$n = 1$
$s \geq 20$	$n = 0$

next roll, whereas the objective of the game is to have the highest possible payoff, which is your score *at the moment you sit down.* Do we need to worry about this distinction? It is tempting to think that they lead to the same outcome—if you put yourself ahead in the immediate future, wouldn't that also put you ahead in the long run? But I could see no valid argument for this. I spent an entire night (my first university all-nighter) tangled with this question.

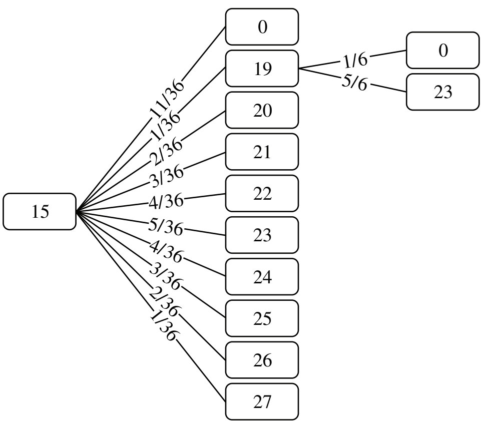
At some point I decided that my only hope was to look for a strategy that outperformed the $A(n)$ strategy. I became interested in the strategy of always using one die because it was the simplest strategy around. I decided to "put it to the test," using EXCEL to compare it with the $A(n)$-strategy I had developed so far. After 50,000 Monte Carlo iterations, the differences were insignificant and inconclusive.

The breakthrough occurred when I looked at the case of $s = 15$. I made a few calculations that put the issue to rest.

A counterexample to the $A(n)$-strategy. Take the case of $s = 15$. The $A(n)$ strategy tells you to use $n = 2$ dice. If double 2s are rolled you stand for one more round using 1 die. Otherwise, you sit. The result is summarized in TABLE 3. The average score is approximately 15.98.

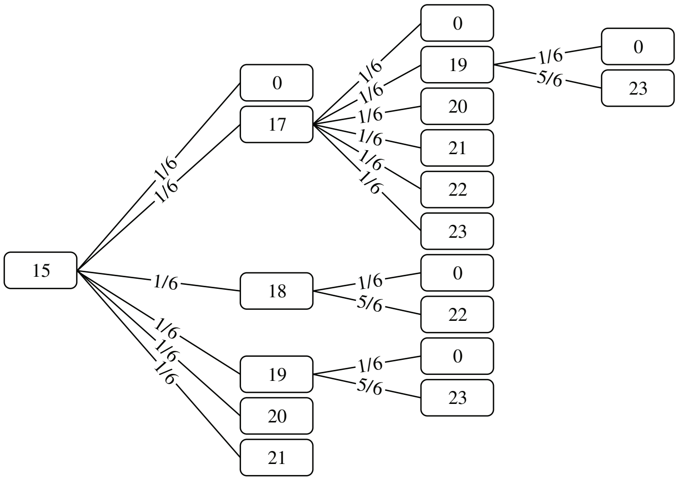
Now compare this with the strategy that uses only one die and stands whenever the score is less than 20 (TABLE 4).

TABLE 3: How the $A(n)$ strategy plays out at $s = 15$



prob.	score $= X$
5/216	23
2/36	20
3/36	21
4/36	22
5/36	23
4/36	24
3/36	25
2/36	26
1/36	27
E(X)	15.97685

TABLE 4: How the 1-die strategy plays out at $s = 15$



prob.	score $= X$
5/216	23
1/36	20
1/36	21
1/36	22
1/36	23
5/36	22
5/36	23
1/6	20
1/6	21
E(X)	16.00463

The average score is a bit above 16, and higher than was obtained starting with 2 dice. For this particular s-value, the 1-die strategy outperforms the $A(n)$ strategy!

PETER: David's 1-die strategy was a revelation to me and for a time I had a bit of trouble thinking clearly about the situation. The example above of $s = 15$ certainly shows that the $A(n)$ strategy is not optimal. But is the 1-die strategy optimal? Are there situations when it might be better to roll more than 1 die? And suppose that the 1-die strategy *is* optimal. When do we stop? Is $s = 20$ the right place to sit? I was thrown for a bit of a loop and decided to go back to the beginning.

It is surprisingly easy to get confused, particularly when there is more than one question buzzing around. What's needed is to focus on one thing at a time, and hope that it's the right thing to begin with. The next day David came to me with a ridiculously simple argument that nothing could possibly outperform the 1-die strategy.

DAVID: Peter is right—it's so easy to miss simple things. And this is one of them. Suppose your score is s and you are using a strategy that tells you to roll 3 dice. Then you would have exactly the same outcome by standing for the next 3 turns and rolling 1 die each time. The reason for this is that the condition for the game to end with a zero payoff is the same in each case—getting a 1 on any of the three dice. So the 1-die strategy will do just as well as the one you are using. But furthermore, it might even do better because it gives you the option of stopping before the third turn.

PETER: Indeed that's exactly why the 1-die strategy outperformed the $A(n)$ strategy at $s = 15$. If you happen to roll a 6 on your first die (giving you $s = 21$) the 1-die strategy lets you stop and sit down, whereas the $A(n)$ strategy rolls again. Now if you stop, your payoff is 21, but if you roll again, your average score becomes

$$s = \frac{0 + 23 + 24 + 25 + 26 + 27}{6} = \frac{125}{6} \approx 20.833$$

which is less than 21.

DAVID: Always roll one die.

The meeting

After the revelation about a pure 1-die strategy, the final challenge was to determine and prove the critical s-value for when to sit. This appeared obvious enough but a formal proof took quite a while to formulate. We sat down for a final meeting to discuss this.

PETER: So the only question left is, when do you sit?

DAVID: At $s = 20$.

PETER: How do we know?

DAVID: Use the same calculation we made above at $s = 21$. It works for any $s > 20$. Your expected score after one roll will always be less than s.

PETER: Right. It is worth emphasizing that. The $A(1)$ strategy (which is optimal) asks you to compare:

$$s \quad \text{and} \quad \frac{(s+2)+(s+3)+(s+4)+(s+5)+(s+6)}{6}$$

On the left is the payoff if you sit and on the right is your expected new score if you stand. For $s < 20$ the right side is bigger, for $s > 20$ the left side is bigger, and for $s = 20$ they are equal. So the strategy says sit when $s > 20$. But as you pointed out long ago, this only considers the next roll instead of the indefinite future. What we really need on the right is some indication of your payoff at the end of the game, given that you stand and play optimally.

DAVID: We need the notion of what a strategy is "worth." If you have score 19, you can expect to increase that on average by staying in the game, so having a score of 19 is actually worth more than 19. However, if you have 21 you can't do any better (in fact, by staying in the game you'll do worse on average), so 21 is only worth 21.

PETER: We could formalize that. Define $v(s)$, the *value* of s, to be the expected payoff for a player who currently has score s and who plays optimally. For example, $v(19) > 19$ and $v(21) = 21$.

DAVID: In fact

$$v(19) = \frac{v(21)+v(22)+v(23)+v(24)+v(25)}{6},$$

and $v(s)$ in general would be

$$v(s) = \max\left(s, \frac{v(s+2)+v(s+3)+v(s+4)+v(s+5)+v(s+6)}{6}\right).$$

The first term represents the payoff if you sit. The second term represents the average payoff if you stand and play optimally. You choose whichever one is greater. If we knew that $v(s) = s$ for big enough s, say for all $s \geq 100$, then we could use the recursive equation to work backwards. We would get $v(99) = 99$, then $v(98) = 98$, and that would keep on working all the way to $v(20) = 20$. The first time s would be less than the expression on the right would be at $s = 19$.

PETER: So what we need to do is to find some large enough s^* for which we can show that $v(s) = s$ for all $s \geq s^*$.

Pretty Black Cat. One way in which Peter creates exercises for the students is to construct variations on what happens when a 1 is rolled. One such variation seems at first quite uninteresting, but in fact it holds the key to a lovely proof of the result we are searching for.

PETER: I've been thinking about a modification called Pretty Black Cat ("PBC") in which you always roll one die, and when a 1 is rolled, the game ends but you do not lose your current score.

DAVID: Not very interesting, of course, because you'd simply always stay in the game.

PETER: Indeed. But the game is so simple that we ought to be able to calculate its $v(s)$ values easily.

DAVID: No doubt. But I'm wondering where this is headed.

PETER: I'm thinking that whatever strategy you choose to use in Skunk, the same strategy used in PBC will give you a payoff that is at least as high. It surely follows that the $v(s)$ values for PBC will always be at least as big as those for Skunk Redux, so PBC's $v(s)$ will give us an upper bound on Skunk's $v(s)$... and that might be useful.

DAVID: Indeed it might. Let's see... in PBC a player with score s would get exactly one more roll with probability $1/6$, exactly two more with probability $(5/6)(1/6)$, exactly three more with probability $(5/6)^2(1/6)$, etc., and the average payoffs would be, s, $s+4$, $s+8$, etc. We just have to add a bunch of terms.

PETER: Or perhaps we could try a recursive argument.

DAVID: Yes. I might have thought of that, as it is one of the big themes of the course. Let k be the amount you gain on average by continuing to play. Then, if your next roll is a 1, k is zero, and otherwise, you gain 4 on average and you are able to keep playing so your overall gain is on average $4+k$. This gives us the recursive equation:

$$k = (1/6)(0) + (5/6)(4+k)$$

and that solves to give $k = 20$.

PETER: Nicely done. So for Pretty Black Cat, the value of having a score s is $v(s) = s + 20$.

DAVID: We can conclude that for Skunk Redux, $v(s) \leq s + 20$.

PETER: Maybe that will be enough to find a score s for which $v(s) = s$.

DAVID: Let's see. Returning to Skunk Redux, $v(s) = s$ if

$$\frac{v(s+2)+v(s+3)+v(s+4)+v(s+5)+v(s+6)}{6} \leq s$$

and since $v(x) \le x + 20$, that will hold if

$$\frac{(s+22)+(s+23)+(s+24)+(s+25)+(s+26)}{6} \le s$$

and that simplifies to $s \ge 120$.

PETER: Wow.

DAVID: We conclude that $v(s) = s$ for every $s \ge 120$.

PETER: That elusive but utterly unsurprising conclusion is just what we need to start the backwards recursion and make all of our deductions legitimate. Finally, we can safely say that 20 is indeed the place to sit.

Epilogue

And thus the four-month journey concludes with the astounding realization that our initial reasoning is flawed. For the $(n = 2)$-dice game we discussed at the beginning, a comparison of the expected immediate gains by sitting and by standing fails to take account of the long-term possibilities. The answer to sit when $s > 200/11$ is correct but requires a more rigorous argument involving $v(s)$.

The reason the 1-die strategy is optimal in an n-dice game, as previously mentioned, is that any gain you can make by rolling n dice can be obtained by rolling 1 die n times. Also, it is important to notice that while the 1-die strategy is optimal, it is not the only optimal strategy. For example, since you will never leave the game with $s < 20$, and 3 dice can only take you to 18, you might just as well throw 4 dice at the very beginning. The same reasoning continues to apply. For example, an optional strategy allows a play of 2 dice for $8 \le s \le 14$, and so on.

An interesting problem arises if we *exclude* the option of using 1 die, that is, you can roll any number of dice except 1. In this case your effective choices become sit, stand with 2 dice, or stand with 3 dice. This is because any number $n > 3$ can be written as a linear combination of 2 and 3. The optimal strategy for this game (found with EXCEL) is displayed in TABLE 5. It has an intriguing pattern.

TABLE 5: An optimal strategy when 1 die is forbidden

score s	optimal n
$s = 0$	$n = 3$
$1 \le s \le 6$	$n = 2$
$7 \le s \le 11$	$n = 3$
$12 \le s \le 18$	$n = 2$
$s > 18$	$n = 0$

More generally, suppose there is a given set of available numbers of dice to roll: $\{n_1, n_2, \ldots, n_k, \ldots\}$, where no n_j is a nonnegative-integer linear combination of the other n_i. We invite others to conduct further research on optimal strategies for this and other variations.

REFERENCES

1. D. Brutlag, Choice and chance in life: The game of SKUNK, *Mathematics Teaching in the Middle School* **1**(1) (April 1994) 28–33. http://illuminations.nctm.org/LessonDetail.aspx?id=L248
   In the elementary classroom game, there is an additional twist: if the dice show "double ones," all points from previous rounds are also wiped out. Our version is simpler.
2. T. W. Neller and C. G. M. Presser, Optimal play of the dice game Pig, *The UMAP Journal* **25**(1) (2004) 25–47. http://cs.gettysburg.edu/~tneller/resources/pig/
   Neller's version of the game is called Pig and involves a 2-player race to a score of 100. Otherwise, the rules are the same.

Summary In the simple version of Skunk, a pair of dice is rolled again and again until either you choose to sit or at least one 1 comes up. If you sit, your payoff is the dice sum of all your previous rolls. If at least one 1 comes up while you are still standing, your payoff is zero. Here we look at an extension of the game in which you must also choose the number of dice to use and you can alter this from round to round. The simple game seems to have a straightforward enough optimal strategy but our analysis of the extension reveals flaws in our initial reasoning. Some additional extensions are considered.

DAVID KONG is a 3rd year undergraduate student at the Queen's School of Business and is pursuing a Bachelor of Science in mathematics. He has worked in both the hedge fund and private equity industries (Fore Research, ONCAP Partners) and was invited to the 2010 Canadian Mathematics Olympiad. This has led to an interest in statistics, probability, and modeling, particularly through EXCEL. David enjoys skiing, biking, cooking, and creating latte art.

PETER TAYLOR is a professor in the Department of Mathematics and Statistics at Queen's University, cross-appointed to the Department of Biology and the Faculty of Education. He is a Queen's graduate and has a Harvard Ph.D. His area of research is theoretical evolutionary ecology, particularly the evolution of cooperative behaviour. He is a 3M Teaching Fellow and is Chair of the Education Committee of the Canadian Mathematical Society. Recently he chaired the Task Force at Queen's, charged with writing the university's new Academic Plan.

NOTES

The New Rules for NFL Overtime

CHRIS JONES
Saint Mary's College of California
Moraga, CA 94575
cjones@stmarys-ca.edu

As many readers may be aware, beginning in the 2010 season the NFL has changed its overtime rules. In this article we will look at why the rules needed changing, how they were changed, and why the new system—at least mathematically—looks so appealing. We will also look at an alternative rule change that might have been even more appealing to mathematicians.

The rule

In the previous system, a coin toss took place and the team that won the toss had the choice either to kick off or receive the ball. The game was played with regular NFL rules and the first team to score won.

Most teams with the choice elected to receive the ball, because that gave them the first possession and the first good chance to score. According to ESPN.com, since 1994 about 60% of games that went into overtime under these rules were won by the team that won the toss. This advantage is statistically significant [**2**] and seems to have been increasing over time. Perhaps thinking that a coin toss should not have such a large influence on the outcome of a game, the NFL became convinced that a rule change was needed.

In April 2010 the NFL owners installed a new system that initially applied only to playoff games. In April 2012 they extended the new rules to all games. In this new system, as before, any team scoring a touchdown or safety wins immediately. So does any team scoring a field goal, *except* that if the initial receiving team scores a field goal (3 points) on its first possession, then the team that kicked off has a chance to reply. If they fail to score, they lose. If they score a field goal, the game again reverts to the old sudden-death rules, and if they score a touchdown they win the game. The game is played until one team wins, removing the possibility of a tie.

The Markov chain

We shall study the effect of this rule change using a Markov chain model. We will define the states of the system, determine the initial conditions and the transition probabilities between states, and then draw conclusions about the outcomes of games using the new rule.

Math. Mag. **85** (2012) 277–283. doi:10.4169/math.mag.85.4.277.

We begin by defining the states of the system. Throughout our work we shall refer to the team that receives the ball (and has the first possession) as Team A, and the team that kicks off as Team B. There are two absorbing states, those in which either team has won; we call these "A's winning state" and "B's winning state," or just "A wins" and "B wins." Every game eventually reaches one of these states and stays there.

There are four other, non-absorbing states. We have the initial possession for Team A in which a touchdown will win the game, but a field goal leads to the additional possession; we will refer to this state as A*. We have the possession for Team B when they are down by one field goal, a state that can occur (if it occurs at all) only after A's first possession; we will refer to this state as B*. Finally we have the two sudden-death states, in which either team has possession under the old rules; we refer to these states as A and B.

Every game begins in state A*. From that state it moves to state B*, state B, or either absorbing state. From state B* it can move to state A or to an absorbing state. Then the game continues through any number of possessions (zero or more) alternating between states A and B, before finally moving to state "A wins" or "B wins." Except for the absorbing states, each step of the Markov chain is identical with one possession, so in general we will refer to the steps of the chain as possessions. FIGURE 1 below has the state diagram for the new rules.

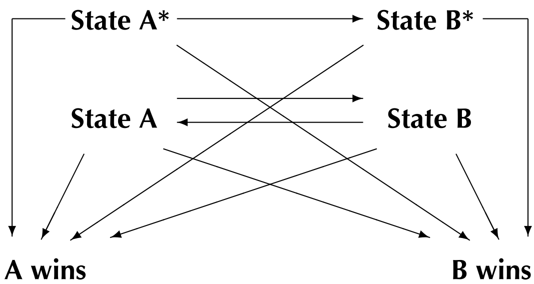


Figure 1 The state diagram

We notice that if a game reaches state A or state B, then we are back to the sudden-death game played according to the old rules. This means that there are plenty of data about what happens in these states, and we will use this data to estimate transition probabilities.

The transition matrix

TABLE 1 below shows the probabilities p_{ij} of going from a given game state (in a row) to another (in a column). We shall determine these p_{ij}'s, initially based on game data.

We note that the majority of the entries are zero. For example, after the initial possessions there is no way to return to the A* and B* states, and thus other than p_{12}, which is the probability of a field goal in the initial possession, the first two columns are all zeros. The remaining nonzero probabilities p_{ij} represent either a change in possession or a score.

In all the analysis that follows we assume two evenly matched teams. Therefore, by symmetry, in the sudden-death scenario we have $p_{43} = p_{34}$, $p_{35} = p_{46}$, and $p_{36} = p_{45}$. We are also making the assumption that drives begin from an average starting point. Clearly a team that starts from its own one-yard line has a lower probability of scoring

TABLE 1: Transition probabilities

to / from	Poss. A*	Poss. B*	Poss. A	Poss. B	A wins	B wins
Poss. A*	0	p_{12}	0	p_{14}	p_{15}	p_{16}
Poss. B*	0	0	p_{23}	0	p_{25}	p_{26}
Poss. A	0	0	0	p_{34}	p_{35}	p_{36}
Poss. B	0	0	p_{43}	0	p_{45}	p_{46}
A wins	0	0	0	0	1	0
B wins	0	0	0	0	0	1

than one that begins inside the opponents' 10-yard line, but for simplicity we treat all drives in a given state as the same.

In the A* possession we shall assume that Team A plays as it would during the first four quarters of a game. Here a team would like to score a touchdown, but will settle for a field goal. (There is speculation as to whether this will be the case in state A*, or whether teams may opt to be more bold and take risks going for a touchdown to avoid giving the opponents the right to reply [**3**].) We can find an estimate for the probabilities to fit into row one of our matrix using game data. In the 2008 season there were 5461 total possessions not ended by the completion of a half. TABLE 2 breaks down how each of these 5461 possessions ended.

TABLE 2: Outcomes of 5461 possessions in 2008

Result	Number
Offensive touchdown	1122
Field goal	845
Punt	2294
Turnover on downs	231
Missed field goals	155
Non-scoring turnovers	708
Safety	21
Turnovers for touchdowns	85

Thus, in any given possession we may assume the following probabilities:

- A field goal will be scored in $\frac{845}{5461} = 0.15$ of possessions, leading to the B* possession.
- A touchdown will be scored in $\frac{1122}{5461} = 0.21$ of possessions, resulting in A winning.
- $\frac{2294+231+708+155}{5461} = 0.62$ result in a change of possession to B.
- $\frac{21+85}{5461} = 0.02$ result in the defense winning via a turnover or safety.

If we assume that, in possession A*, the team plays as it would during a non-overtime possession, we can assign the values $p_{12} = 0.15$, $p_{14} = 0.62$, $p_{15} = 0.21$ and $p_{16} = 0.02$.

The B* row is more complex, requiring that we make some estimates as to how a team's decisions would effect the probabilities of scoring on a given possession. We note that in this scenario, down by 3 points, Team B will never punt, as that would mean losing the game. In general the team will play more aggressively, resulting in more turnovers (in particular, on downs), but also more scores. Thus we need to split the 2294 possessions that result in punts into scores and losing the ball on downs. In the NFL, around 50% of fourth-down attempts are successfully converted, although the majority of these are for relatively short yardage. Using this fact combined with the number of successful late-game drives, in which a similar situation to Team B's occurs, we estimate subjectively that, of the 42% of possessions that would normally end in a punt, 60% will result in a turnover on downs and 40% will end up with scores. Of those that end up in a score, we shall keep the touchdown to field goal ratio as it is given in our game data, $1122 : 845 \approx 57 : 43$.

All outcomes other than a field goal or touchdown result in Team B losing the game. The probabilities are therefore

- Touchdown $0.21 + (0.42 \times 0.4 \times 0.57) = 0.31$,
- Field goal $0.15 + (0.42 \times 0.4 \times 0.43) = 0.22$,
- Change of possession $1 - (0.31 + 0.22) = 0.47$.

Thus we can assign $p_{23} = 0.22$, $p_{25} = 0.47$, and $p_{26} = 0.31$.

Having established the top two rows and first two columns we can consider the lower-right four-by-four portion of the matrix, which is a significant matrix in its own right. If we reach one of these states, we are back to the sudden-death scenario of the previous rules. This portion of the matrix was analyzed in [**1**], which included a Markov-chain treatment of the previous rules. Following that treatment, initially we consider using the average game data as we did above. Recall that at this point there are only three possible outcomes from a possession: winning, giving up possession, or losing. Based on the game data and what we have already seen we could assume a team scores in $0.15 + 0.21 = 0.36$ of its possessions, and the probability of a change of possession and the defense scoring are, as before, 0.62 and 0.02, respectively. This would give a four-by-four matrix of the form

$$T_1 = \begin{bmatrix} 0 & 0.62 & 0.36 & 0.02 \\ 0.62 & 0 & 0.02 & 0.36 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

However, as noted in [**1**], this matrix does not do a good job of modeling the previous overtime system. In a typical NFL quarter, there are on average six possessions. If we take T_1 to the sixth power to model a single quarter of overtime, we get

$$T_1^6 = \begin{bmatrix} 0.06 & 0 & 0.57 & 0.37 \\ 0 & 0.06 & 0.37 & 0.57 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

which heavily overestimates the number of games that were not concluded at the end of the first quarter of overtime. Under the old rules, in the event of there being no score after one quarter, a regular season game was declared a tie; however, of the nearly 100 games that went into overtime between 2004 and 2010, only once did a game end with such a result. The obvious explanation for this is that a team can play more conservatively as they get closer to the opposition's goal line, knowing a field goal will suffice to win the game. We would therefore expect the probability of a score to rise and the probability of a change in possession to decrease. In [**1**] we show that the matrix T_2 given below, when taken to the power six, gives an excellent estimate of the true outcomes of overtime games:

$$T_2 = \begin{bmatrix} 0 & 0.5 & 0.48 & 0.02 \\ 0.5 & 0 & 0.02 & 0.48 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

When taken to the sixth power, this matrix gives

$$T_2^6 = \begin{bmatrix} 0.02 & 0 & 0.64 & 0.34 \\ 0 & 0.02 & 0.34 & 0.64 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

This matches closely with the actual results from games from 2004–09, in which the receiving team won 64% of the time and the kicking team 35% of the time, with 1% ending in a tie. (We note that the advantage for the receiving team has actually increased over the past twenty years. This gives a higher value for teams winning than the previously quoted value from ESPN.com.) It also fits with our intuition that a team, by playing conservatively, should score more often when faced with a sudden death situation.

We can now use T_2 from above to give $p_{34} = p_{43} = 0.50$, $p_{35} = p_{46} = 0.48$, and $p_{45} = p_{36} = 0.02$ in TABLE 1. Our transition matrix is now therefore

$$T_3 = \begin{bmatrix} 0 & 0.15 & 0 & 0.62 & 0.21 & 0.02 \\ 0 & 0 & 0.22 & 0 & 0.47 & 0.31 \\ 0 & 0 & 0 & 0.50 & 0.48 & 0.02 \\ 0 & 0 & 0.50 & 0 & 0.02 & 0.48 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

By raising this matrix to high powers, we can see the ultimate outcomes of each state:

$$\lim_{n\to\infty} T_3^n = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.52 & 0.48 \\ 0 & 0 & 0 & 0 & 0.61 & 0.39 \\ 0 & 0 & 0 & 0 & 0.65 & 0.35 \\ 0 & 0 & 0 & 0 & 0.35 & 0.65 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The first row of this matrix gives the probabilities of reaching the absorbing states, given that we begin the chain in state A*. We expect Team A to win 52% of the time, and Team B to win 48% of the time.

The overall advantage to the team first receiving the ball is thus reduced to 52–48. It may seem surprising that such a small change would result in apparently removing most of the advantage for the team that receives the ball first; however, we now have two advantages balancing out. Team A has the benefit of receiving the ball first and, should they score a touchdown, they win without an opportunity for reply. However, in the B* possession, Team B has the advantage of knowing what they need to do to win, or at least to prolong the game. In a situation where they might otherwise punt they will attempt to gain a first down, knowing that they have nothing to lose. These advantages seem to balance out quite nicely, creating an even field.

The mathematicians' proposal—"cut and choose"

One further proposal that is suggested by mathematicians (though has gained little traction in football circles [**2**]) is based on the cut-and-choose technique employed by cake lovers worldwide. This is the simplest form of the fair division problem in game theory. Suppose we have one piece of cake to share between two siblings. One sibling gets to cut while the second chooses the slice she wishes to eat. Any deviation from a 50-50 split will guarantee the "chooser" gets the larger share and so there is an incentive for the "cutter" to cut as accurately as possible. How can we transfer this idea to football?

In football the analogous situation would involve one team placing the ball at a point on the field, say the offensive team's own 15-yard line. The other team then chooses whether they wish to be the offensive team in that position, or let the other team have the first possession. They then proceed to play the traditional NFL overtime rules, in which the first team to score wins. As demonstrated in game situations, given the chance to start between, say, their own 20- and 30-yard line (the typical result of a normal kickoff and return) a team always chooses to have the ball. But suppose a team is told they would start with the ball on their own one-yard line: would they still take the ball? Unlikely, even for a high-powered offense. Thus there must be a break-even point, where teams are unsure as to whether they would wish to start with the ball or on defense.

Taking this back to our cake analogy, suppose that one team (the cake cutter) gets to choose where the team with the ball begins. The other team (the cake chooser), knowing where on the field the team on offense would begin, gets to choose whether to start on offense or defense. The incentive for the choosing team is to place the ball at the fairest point on the field. Too close to the offense's own goal line and they will end up with the ball there. Too far away and the opposition will take the ball and have the chance to drive down the field and score. Variations of this have been proposed in which the two teams submit sealed bids for where they would wish to begin an overtime possession and the one who offers to start closest to their own goal line gets the ball. This, however, is the same basic concept as cut and choose.

Conclusion

The analysis confirms that the new rule does balance out a large proportion of the advantage that comes from winning the toss. In [**1**] we look at the possibility of playing first to six, another possibility that was proposed as the solution to the advantage the toss provides. If we require a total of six points scored to win in overtime, there is some reduction in the advantage for the team winning the toss to a 56-44 edge. In this case, however, there is the unfortunate byproduct of a considerable increase in the average length of games, likely to be unpalatable to a league so geared to TV schedules. The cut-and-choose method might marginally extend the game, since the offense would likely start closer to their own goal line, but makes no great changes to the format of the game and would thus not significantly adjust the length of time. It is, however, perfectly fair; perhaps the stronger team has an advantage in terms of its bidding policy, but any system that gives advantage to the better team can only be good. The downside of this method is its complexity. Sports fans, notoriously resistant to change, are happy and accepting of a system in which a coin is tossed. It is harder to see them embracing a system in which two teams submit sealed bids for the yard-line they wish to start on.

The new system, the one installed by the NFL, has clear advantages over each of the alternatives we have mentioned. Although it is a little more complicated than the previous system, it is a recognizable method, with only a small change from the familiar rule. It successfully balances out most of the advantage to the team that wins the toss, while barely raising the expected length of the game. No doubt these were among the considerations of the NFL, and it appears that their selection of a new technique was a reasonable choice.

REFERENCES

1. Chris Jones, A look at overtime in the NFL, in *Mathematics and Sports*, Dolciani Mathematical Expositions #43, Mathematical Association of America, Washington, DC, 2010.
2. Carl Bialik, Should the outcome of a coin flip mean so much in NFL overtime? *Wall Street Journal*, December 23, 2003. http://faculty.haas.berkeley.edu/hender/ot_wsj.pdf
3. Kurt W. Rotthoff, Bankruptcy behavior in the NFL: Does the overtime structure change the strategy of the game? To appear in *Journal of Economics and Finance*. Available at SSRN: http://ssrn.com/abstract=1105739

Summary In 2010 the NFL changed the rules for football overtime for playoff games, and in 2012 extended the change to apply to all games. We look at how the previous rule favored the team winning the coin toss with a statistically significant advantage, and then use Markov Chains to analyze the new method. It can be demonstrated that, allowing for certain assumptions about a team's risk tolerance, the new method neatly balances out the coin toss advantage without significantly extending the expected length of the game. One other proposed technique for determining the winner of an overtime game is considered.

Combinatorial Remarks about a "Remarkable Identity"

JØRN B. OLSSON
University of Copenhagen
DK-2100 Copenhagen Ø, Denmark
olsson@math.ku.dk

JAMES A. SELLERS
Penn State University
University Park, PA 16802
sellersj@psu.edu

In the June, 2012, issue of this MAGAZINE, Frumosu and Teodorescu-Frumosu [**1**] proved that, for all integers $m \geq 2$,

$$\sum_{p=1}^{m} \left(\frac{(-1)^p}{p!} \sum_{k_1+\cdots+k_p=m} \frac{1}{k_1 \cdots k_p} \right) = 0 \tag{1}$$

where the inner sum is taken over all p-term *ordered partitions* of m. Their proof is calculus-based, relying on power series manipulations. In this note, we provide a combinatorial proof of this identity, which the authors requested at the end of their article. The proof gives a more general form of the identity, which allows us to state a number of other concrete results of the same type. Related results may be proved using character theory of the symmetric groups. This is discussed in the final section.

Math. Mag. **85** (2012) 283–288. doi:10.4169/math.mag.85.4.283.

First Combinatorial Approach—Stirling Numbers of the First Kind

The primary step in proving a generalization of (1) in a combinatorial way is to rewrite the inner sum so that the sum is taken over *partitions* rather than ordered partitions. We will utilize the *rising factorial* $x^{\overline{m}}$ which for $m \geq 1$ is defined by

$$x^{\overline{m}} := x(x+1)\cdots(x+m-1). \tag{2}$$

For example, $1^{\overline{m}} = m!$ and $2^{\overline{m}} = (m+1)!$ for all $m \geq 1$. Also, $(-1)^{\overline{m}} = 0$ when $m \geq 2$, and $(-2)^{\overline{m}} = 0$ when $m \geq 3$.

The quantity $x^{\overline{m}}$ is a polynomial for each $m \geq 1$, and so it can be written as a sum of ordinary powers:

$$x^{\overline{m}} = \sum_{p=1}^{m} s(m, p)x^p \tag{3}$$

The coefficients $s(m, p)$ which appear in (3) are called the (unsigned) *Stirling numbers of the first kind*. There is a rich theory of these numbers [**2, 4**]. The key property which we need in this note is that $s(m, p)$ counts the number of permutations of the set $\{1, 2, \ldots, m\}$ with exactly p cycles in their cycle decompositions.

We now show that

$$\frac{1}{p!}\sum_{k_1+\cdots+k_p=m}\frac{1}{k_1\cdots k_p} = \frac{1}{m!}s(m, p) \tag{4}$$

where as in (1) the sum is on p-term ordered partitions of m.

Indeed

$$\sum_{k_1+\cdots+k_p=m}\frac{1}{k_1\cdots k_p}$$

$$= \sum_{\substack{k_1\geq k_2\geq\cdots\geq k_p\geq 1\\ k_1+\cdots+k_p=m}}\frac{1}{k_1\cdots k_p}\times(\text{number of ways to permute the parts})$$

$$= \sum_{\substack{t_1,t_2,\ldots,t_m\geq 0\\ t_1\cdot 1+t_2\cdot 2+\cdots+t_m\cdot m=m\\ t_1+t_2+\cdots+t_m=p}}\frac{1}{1^{t_1}2^{t_2}\ldots m^{t_m}}\times\frac{p!}{t_1!t_2!\ldots t_m!}$$

where t_i, $1 \leq i \leq m$, is the number of occurrences of the part i in a given partition of m. By writing the partitions of m as $t_1 \cdot 1 + t_2 \cdot 2 + \cdots + t_m \cdot m$ we are able to get an explicit handle on those partitions of m which contain exactly p parts.

From the above we see that

$$\frac{1}{p!}\sum_{k_1+\cdots+k_p=m}\frac{1}{k_1\cdots k_p} = \sum_{\substack{t_1,t_2,\ldots,t_m\geq 0\\ t_1\cdot 1+t_2\cdot 2+\cdots+t_m\cdot m=m\\ t_1+t_2+\cdots+t_m=p}}\frac{1}{1^{t_1}2^{t_2}\ldots m^{t_m}}\times\frac{1}{t_1!t_2!\ldots t_m!}$$

which is equivalent to

$$\frac{1}{p!}\sum_{k_1+\cdots+k_p=m}\frac{1}{k_1\cdots k_p} = \frac{1}{m!}\sum_{\substack{t_1,t_2,\ldots,t_m\geq 0\\ t_1\cdot 1+t_2\cdot 2+\cdots+t_m\cdot m=m\\ t_1+t_2+\cdots+t_m=p}}\frac{1}{1^{t_1}2^{t_2}\ldots m^{t_m}}\times\frac{m!}{t_1!t_2!\ldots t_m!}. \tag{5}$$

Now a summand in the sum on the right-hand side of (5) counts the number of permutations of the set $\{1, 2, \ldots, m\}$ which, for each $1 \leq i \leq m$, have exactly t_i cycles of length i in their unique cycle decomposition. This fact may be deduced directly using elementary counting methods. Therefore, the sum on the right-hand side of (5) equals $s(m, p)$, and this proves(4).

Together, equations (3) and (4) imply the following significant generalization of (1):

THEOREM 1. *Let $m \geq 1$. We have the polynomial identity*

$$\sum_{p=1}^{m} \left(\frac{x^p}{p!} \sum_{k_1+\cdots+k_p=m} \frac{1}{k_1 \cdots k_p} \right) = \frac{1}{m!} x^{\overline{m}}. \tag{6}$$

In particular, Theorem 1 shows that the left-hand side of (1) equals $\frac{1}{m!}(-1)^{\overline{m}}$, which in turn equals 0, whenever $m \geq 2$, by (2). Thus we have given a combinatorial proof of (1).

Theorem 1 can also be used to prove other combinatorial identities that are related to (1). For example, we see that

$$\sum_{p=1}^{m} \left(\frac{1}{p!} \sum_{k_1+\cdots+k_p=m} \frac{1}{k_1 \cdots k_p} \right) = 1 \tag{7}$$

for each $m \geq 1$ by substituting $x = 1$ into (6). Similarly, the substitution $x = 2$ in (6) yields

$$\sum_{p=1}^{m} \left(\frac{2^p}{p!} \sum_{k_1+\cdots+k_p=m} \frac{1}{k_1 \cdots k_p} \right) = m + 1 \tag{8}$$

for $m \geq 1$. Lastly, for $m \geq 3$, (6) gives

$$\sum_{p=1}^{m} \left(\frac{(-2)^p}{p!} \sum_{k_1+\cdots+k_p=m} \frac{1}{k_1 \cdots k_p} \right) = 0 \tag{9}$$

via the substitution $x = -2$.

We return specifically to (7)–(9) below when we present our character-theoretic perspective.

Second Combinatorial Approach—Character Theory

In this section we refocus our attention to group theory, in particular to the characters of finite groups. Character theory was invented by Frobenius as a help to study the structure of finite groups, but it has developed into an area of independent interest with applications outside of group theory. See [**3**] for a very thorough introduction.

In this context, we consider the permutations of the previous section as elements of symmetric groups. We will see that from this point of view, the results in the previous section are all special cases of a more general formula involving characters. In addition, as often happens in abstraction, the more general formula quickly implies other results that aren't at all obvious from the results of the previous section.

If G is a (finite) group, then a representation T of G is a map, associating to each group element $g \in G$ an invertible square matrix $T(g)$ with complex entries such that T "respects multiplication." This means that $T(g_1g_2) = T(g_1)T(g_2)$ for all $g_1, g_2 \in G$, i.e., T is a homomorphism of groups. The character χ_T of T is defined by $\chi_T(g) = \text{trace}(T(g))$, the sum of the diagonal entries in $T(g)$. A character of G is the character of some representation. One of the amazing facts about characters is that in a sense you may recover a representation from its character. You are not losing information by looking only at traces. The fact that similar matrices have the same trace implies that character values are constant on the conjugacy classes of the group. It is known that the sum and the product of two characters is again a character. A character is called *irreducible* if it cannot be written as a sum of two other characters. The simplest irreducible character of G is the *trivial character* 1_G, which maps all elements of G to 1. Any character χ may be decomposed into a sum of (not necessarily distinct) irreducible characters and it can be shown that this decomposition is unique. We denote by $a(\chi)$ the multiplicity (number of occurrences) of the trivial character 1_G in the decomposition of the character χ. It is well-known [**3**, Theorem 14.17] that

$$a(\chi) = \frac{1}{m!} \sum_{g \in G} \chi(g) \tag{10}$$

so $a(\chi)$ is really just the *average of all the values of the character* χ.

The set of all permutations of $\{1, 2, \ldots, m\}$ considered above forms a group, which is called the symmetric group S_m. The characters of S_m turn out to be non-zero integer-valued functions on the elements of S_m, which take the same value on permutations having the same cycle decomposition. This is because these elements are in the same conjugacy class of S_m.

To involve the characters of S_m we start by replacing the term $(-1)^p$ in the left-hand side of (1) with a "weight function" $w(k_1, k_2, \ldots, k_p)$. That is, we define

$$\sigma(m, w) := \sum_{p=1}^{m} \left(\frac{1}{p!} \sum_{k_1 + \cdots + k_p = m} \frac{w(k_1, \ldots, k_p)}{k_1 \cdots k_p} \right)$$

for $m \geq 1$ and suitable choices of the function w. Choosing w to be identically 1 or to be $(-1)^{m-p}$, we see that $\sigma(m, w)$ equals 1 or 0, respectively, by (7) and (1). These are also examples of a special kind of weight function which we now consider.

Any character of S_m gives rise to a weight function. Thus, if χ is any character of S_m, then we may define a weight function w_χ as follows:

> $w_\chi(k_1, \ldots, k_p)$ is the value of χ on an element that is a product of disjoint cycles of lengths $k_1, \ldots, k_p$.

Since disjoint cycles in symmetric groups commute, this value is independent of the order of the arguments. Apart from the trivial character 1_{S_m} the simplest irreducible character is the sign character sgn_{S_m}. It maps an even permutation to 1 and an odd permutation to -1. If a permutation in S_m is a product of p disjoint cycles then it is an even permutation exactly when $m - p$ is even, i.e., when $(-1)^{m-p} = 1$. Thus $w_{\text{sgn}_{S_m}}(k_1, \ldots, k_p) = (-1)^{m-p}$.

We can now prove a more general form of Theorem 1.

THEOREM 2. *For any character* χ *of* S_m,

$$\sigma(m, w_\chi) = a(\chi).$$

Proof. By definition $w_\chi(k_1, \ldots, k_p)$ is independent of the ordering of $k_1, \ldots, k_p$. Therefore the calculation in the previous section shows that

$$\sigma(m, w_\chi) = \sum_{p=1}^{m} \left(\frac{1}{p!} \sum_{k_1+\cdots+k_p=m} \frac{w_\chi(k_1, \ldots, k_p)}{k_1 \cdots k_p} \right) = \frac{1}{m!} \sum_{g \in S_m} \chi(g) = a(\chi). \quad \blacksquare$$

In view of the above remarks, the left-hand sides of (7) and (1) equal $\sigma(m, w_{1_{S_m}})$ and $(-1)^m \sigma(m, w_{\mathrm{sgn}_{S_m}})$, respectively. Thus (7) and (1) follow from Theorem 2. We also have that $a(\mathrm{sgn}_{S_m}) = 0$. This is because 1_{S_m} cannot occur in the decomposition of the irreducible character sgn_{S_m}. Thus we have gained the following additional insight: *The original identity* (1) *is equivalent to the well-known fact that there are equally many even and odd permutations of* $\{1, 2, \ldots, m\}$.

This character-theoretic viewpoint also provides a new way to view (8). Namely, consider the action of S_m on the power set $\mathcal{P}_m$ of $\{1, 2, \ldots, m\}$. The corresponding character χ_{pow} has the property that $\chi_{\mathrm{pow}}(g) = 2^p$ where p is the number of cycles in g. Using [**5**, Example 7.18.8], we see that $a(\chi_{\mathrm{pow}}) = m + 1$. Also $a(\mathrm{sgn}_{S_m} \chi_{\mathrm{pow}}) = 0$. (Note that $\mathrm{sgn}_{S_m} \chi_{\mathrm{pow}}$ is a product of two characters and thus also a character.) Now (8) and (9) follow from Theorem 2.

We now discuss one final example of a "relative" of the original identity (1) which does not follow from Theorem 1. Consider the weight function w_1 defined by

$$w_1(k_1, \ldots, k_p) := |\{i \mid k_i = 1\}|.$$

That is, $w(k_1, \ldots, k_p)$ is just the number of 1's that appear among the arguments $k_1, \ldots, k_p$. Further, $w_1 = w_{\chi_{\mathrm{nat}}}$ where χ_{nat} is the character of S_m acting naturally on the set $\{1, 2, \ldots, m\}$. (Thus $\chi_{\mathrm{nat}}(g)$ equals the number of fixed points of g.) As noted in [**3**, Corollary 29.10], χ_{nat} is a sum of 1_{S_m} and another irreducible character. Thus $a(\chi_{\mathrm{nat}}) = 1$ and $a(\mathrm{sgn}_{S_m} \chi_{\mathrm{nat}}) = 0$. Therefore, we have the following "relatives" of (7) and (1) respectively:

THEOREM 3. *For all* $m \geq 2$,

$$\sum_{p=1}^{m} \left(\frac{1}{p!} \sum_{k_1+\cdots+k_p=m} \frac{|\{i \mid k_i = 1\}|}{k_1 \cdots k_p} \right) = 1,$$

and for all $m \geq 3$,

$$\sum_{p=1}^{m} \left(\frac{(-1)^p}{p!} \sum_{k_1+\cdots+k_p=m} \frac{|\{i \mid k_i = 1\}|}{k_1 \cdots k_p} \right) = 0.$$

We close by highlighting that the first equation in Theorem 3 is equivalent to the following combinatorial statement:

The total number of fixed points in all permutations of $\{1, 2, \ldots, m\}$ *equals* $m!$.

This is because the left-hand side of the equation is equal to $a(\chi_{\mathrm{nat}})$ and $\chi_{\mathrm{nat}}(g)$ counts the fixed points of g. We see also by (10) that in average the permutations of $\{1, 2, \ldots, m\}$ have one fixed point.

There is a simple direct proof of the combinatorial statement: List the $m!$ permutations in an $(m! \times m)$-matrix where the ith row contains the ith permutation in some arbitrary ordering of the permutations. For example, the corresponding matrix for the case $m = 3$ can be written as follows:

$$\begin{array}{ccc} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{1} & 3 & 2 \\ 2 & 1 & \mathbf{3} \\ 2 & 3 & 1 \\ 3 & 1 & 2 \\ 3 & \mathbf{2} & 1 \end{array}$$

The fixed points correspond to the occurrences of an integer $\mathbf{j}$ in the jth column of this matrix. Clearly each column contains each of the integers $1, 2, \ldots, m$ with the same multiplicity of $(m-1)!$. In particular the jth column contains j with this multiplicity. Thus there is a total of $m \cdot (m-1)! = m!$ fixed points in all the permutations of $\{1, 2, \ldots, m\}$.

Acknowledgment J. A. Sellers gratefully acknowledges the support of the Austrian American Educational Commission which supported him during the Summer Semester 2012 as a Fulbright Fellow at the Johannes Kepler University, Linz, Austria.

REFERENCES

1. M. Frumosu and A. Teodorescu-Frumosu, A remarkable combinatorial identity, *Math. Mag.* **85** (2012) 201–205. http://dx.doi.org/10.4169/math.mag.85.3.201
2. R. Graham, D. Knuth, and O. Patashnik, *Concrete Mathematics,* 2nd ed., Addison-Wesley, 1994.
3. G. James and M. Liebeck, *Representations and Characters of Groups,* 2nd ed., Cambridge University Press, 2001.
4. F. Roberts, *Applied Combinatorics,* Prentice Hall, 1984.
5. R. Stanley, *Enumerative Combinatorics,* Vol. 2, Cambridge University Press, 1999.

Summary In the June 2012 issue of this MAGAZINE, Frumosu and Teodorescu-Frumosu proved that, for all integers $m \geq 2$,

$$\sum_{p=1}^{m} \left(\frac{(-1)^p}{p!} \sum_{k_1+\cdots+k_p=m} \frac{1}{k_1 \cdots k_p} \right) = 0$$

where the inner sum is taken over all p-term ordered partitions of m. Their proof is calculus-based, relying on power series manipulations. In this note, we provide a combinatorial proof of this identity (which they requested at the end of their article) and we use the insights gained via this argument to prove other results of a similar type.

An Elementary Proof of the Two-sidedness of Matrix Inverses

FRANK SANDOMIERSKI
Kent State University
Kent, OH 44242
fsandomierski@neo.rr.com

Beauregard in [**1**] gave a short proof of the two-sidedness of an inverse of a square matrix over a field. Here we present another proof, which uses only basic notions from an elementary matrix algebra course. There, students learn how to find the reduced echelon form of the the matrix of a system of linear equations, from which it readily follows that a homogeneous system of linear equations with more unknowns than equations has a nontrivial solution.

Suppose A, B are $n \times n$ matrices over a field with $AB = I$, the identity $n \times n$ matrix.

Proof of $BA = I$. It is a well-known exercise that if B has both a left inverse and a right inverse, they are equal: If $AB = BC = I$, then $C = IC = ABC = AI = A$ and $BA = I$. So we need only find a right inverse of B.

We form the homogeneous system of n^2 linear equations, $BX - xI = 0$, where X is an $n \times n$ matrix. The system has $n^2 + 1$ unknowns, x and the entries of X, so it has a nontrivial solution $(X, x) = (D, d)$.

If $d = 0$ then $BD = 0$, and then $0 = ABD = ID = D$, a contradiction.

Thus, $d \neq 0$ and $BDd^{-1} = I$. Now $C = Dd^{-1}$ is a right inverse of B. ■

REFERENCES

1. R. A. Beauregard, A short proof of the two-sidedness of matrix inverses, *Math. Mag.* **80** (2007) 135–136.

Summary For square matrices, if $AB = I$ then $BA = I$. This is proved from the fact that a homogeneous system of linear equations with more unknowns than equations has a nontrivial solution.

Math. Mag. **85** (2012) 289. doi:10.4169/math.mag.85.4.289.

Heronian Triangles with Sides in Arithmetic Progression: An Inradius Perspective

HERB BAILEY
Rose-Hulman Institute of Technology
Terre Haute, IN 47803
bailey@rose-hulman.edu

WILLIAM GOSNELL
Amherst, MA 01002
campcounselor77.com

A triangle with rational sides and rational area is called an *Heronian triangle* after the Greek mathematician Hero (or Heron) of Alexandria (born ca. AD 10). In this note we consider Heronian triangles with sides in arithmetic progression, which we call *H.A.P. triangles.* H.A.P. triangles have been considered in a number of recent papers including those by Fleenor [**2**], MacDougall [**3**], and Beauregard and Suryanarayan [**1**]. Three examples are the triangles with sides $(3, 4, 5)$, $(25, 38, 51)$, and $(15/11, 26/11, 37/11)$. Their respective areas are 6, 456, and $156/121$, and their inradii are 1, 8, and $4/11$.

One special class of H.A.P. triangles consists of those with consecutive integer sides and integer area. These are called *Brahmagupta triangles* after the Indian mathematician Brahmagupta (born AD 598). After $(3, 4, 5)$ the next example is $(13, 14, 15)$, which has area 84 and inradius 4.

In this note we give a method to generate all H.A.P. triangles. We then show that all Brahmagupta triangles can be generated as solutions of a difference equation with certain initial conditions. The same difference equation with different initial conditions generates another special class of H.A.P. triangles in which, like Brahmagupta triangles, the side lengths differ from each other by 1 and from the inradius by integers, but whose side lengths are not themselves integers. The third example above is in this class of triangles.

We note below that if a triangle has consecutive integer sides, then it has integer area if and only if its inradius is an integer. Thus we might as well have defined a Brahmagupta triangle as one with consecutive integer sides and integer inradius. The computations for generating Brahmagupta triangles are made somewhat easier by focusing on inradius rather than area.

Preliminary results

If the sides a, b and c of a triangle are rational then the semiperimeter $s = (a + b + c)/2$ is also rational. Hence the area A of a triangle with rational side lengths is rational if and only if its inradius r is rational, since $A = sr$. The triangle's area is also given by Hero's (or Heron's) formula,

$$A = \sqrt{s(s-a)(s-b)(s-c)}.$$

For an H.A.P. triangle the sides can be expressed as $(a, b, c) = (b - d, b, b + d)$, where b and d are rational with $0 < d < b$. The semiperimeter is $s = 3b/2$ and the

Math. Mag. **85** (2012) 290–294. doi:10.4169/math.mag.85.4.290. © Mathematical Association of America

area is $A = rs = 3br/2$. Substituting the expressions $a = b - d, c = b + d, s = 3b/2$ into Heron's formula and simplifying gives

$$16A^2 = 3b^2(b - 2d)(b + 2d). \tag{1}$$

Now, substituting $A = 3br/2$ and simplifying further gives the key relationship

$$b^2 - 12r^2 = 4d^2. \tag{2}$$

Equations (1) and (2) hold for any triangle whose sides are in arithmetic progression, and for any such triangle, equation (2) gives the inradius r in terms of b and d. Any triple of rational numbers b, d, and r satisfying $0 < d < b$ and equation (2) defines an H.A.P. triangle.

From these equations we can prove a proposition about Brahmagupta triangles.

PROPOSITION 1. *A triangle with consecutive integer side lengths has integer area—that is, it is a Brahmagupta triangle—if and only if it has integer inradius.*

Proof. From the hypothesis, b is an integer and $d = 1$. From (2), if r is an integer then b is even, and thus s is an integer. Since $A = rs$ the area A must also be an integer. For the converse, if A is an integer then (1) shows that b is even. Also from (1) we have $A = (b/2)\sqrt{3(b/2)^2 - 3d^2}$, and since A is an integer then $3(b/2)^2 - 3d^2$ must be a square integer divisible by 9. Thus $r = A/s = (2A)/(3b) = (1/3)\sqrt{3(b/2)^2 - 3d^2}$ is an integer. ∎

Generating all H.A.P. triangles

If the sides of an H.A.P. triangle are divided by d, the resulting triangle has sides $(b/d - 1, b/d, b/d + 1)$ and the resulting inradius is rational. The scaled triangle is then also an H.A.P. triangle. Thus to find all H.A.P. triangles, we need only seek those whose sides differ by $d = 1$.

Hence by (2) we must find all rational pairs (b, r) satisfying $b^2 - 12r^2 = 4$.

These pairs are found by a method having several names, including "rational slope method" and "Diophantus chord method." The problem is to determine all rational points in the first quadrant on the hyperbola $b^2 - 12r^2 = 4$. The point $(r_0, b_0) = (0, 2)$ is on the hyperbola. For any value of m satisfying $0 < m < 2\sqrt{3}$, the line $b = mr + 2$ passes through $(0, 2)$ and one other point (r, b) on the hyperbola in the first quadrant. (The slope m must be less than $2\sqrt{3}$ because $b = 2\sqrt{3}\,r$ is an asymptote.)

Combining the equations of the line and the hyperbola gives $r = 4m/(12 - m^2)$ and $b = (24 + 2m^2)/(12 - m^2)$. Thus when m is rational, so are both r and b; when r and b are both rational, so is $m = r/(b - 2)$. Thus rational points on the hyperbola in the first quadrant are in one-to-one correspondence with rational numbers m satisfying $0 < m < 2\sqrt{3}$.

Replacing m by p/q, p and q positive relatively prime integers with $p/q < 2\sqrt{3}$, gives the first quadrant intersection

$$r = \frac{4pq}{12q^2 - p^2}, \qquad b = \frac{24q^2 + 2p^2}{12q^2 - p^2}. \tag{3}$$

Each pair p, q of positive, relatively prime integers with $p/q < 2\sqrt{3}$ defines an H.A.P. triangle with $d = 1$ by (3), and every H.A.P. triangle is a scaled version of one of these triangles.

The three examples in the introduction correspond to the (p, q) pairs $(2, 1)$, $(3, 2)$ and $(1, 1)$ respectively, where the second example requires a scaling factor of 13. TABLE 1 gives the values of r, b and A corresponding to all p and q from 1 to 3 with p and q relatively prime and $p/q < 2\sqrt{3}$.

TABLE 1: Properties of H.A.P. triangles with $d = 1$, for small values of p and q

p	q	b	r	A
1	1	26/11	4/11	156/121
1	2	98/47	8/47	1176/2209
1	3	218/107	12/107	3924/11449
2	1	4	1	6
2	3	28/13	3/13	126/169
3	1	14	4	84
3	2	38/13	8/13	456/169

H.A.P. triangles with $d = 1$ and $b - r$ an integer

We now classify H.A.P triangles with the two properties that

- $d = 1$, and
- the side lengths differ from the inradius by an integer; that is, $b - r = v$ for some positive integer v.

The first and third examples in the introduction satisfy these requirements. So do all Brahmagupta triangles, which have the additional property that the sides themselves are integers. In this case we cannot rely on scaling, because when $d = 1$ the difference $b - r$ scales to a rational but not necessarily to an integer.

Setting $d = 1$ and $b = r + v$ in (2) and solving the resulting quadratic for r gives

$$r = \frac{v + 2\sqrt{3v^2 - 11}}{11}. \tag{4}$$

Since r is rational, $3v^2 - 11$ must be a square integer, say u^2; and now $r = (v + 2u)/11$. We seek all pairs (u, v) with u and v positive integers such that

$$u^2 - 3v^2 = -11. \tag{5}$$

Each such pair (u, v) determines a triangle with the required properties.

A nice treatment of Diophantine equations of this type is given by T. Nagell [**4**]. A solution pair (u, v) is associated with the number $u + v\sqrt{3}$. All the solutions of (5) are determined by finding the fundamental solution $x_1 + y_1\sqrt{3}$ of $x^2 - 3y^2 = 1$ and the fundamental solutions, $u_1 + v_1\sqrt{3}$ and $u_2 + v_2\sqrt{3}$, for each of the two solution classes of $u^2 - 3v^2 = -11$. (A fundamental solution is one with the minimum positive integer value of the coefficient of $\sqrt{3}$ within its class.) By inspection, these solutions are $(x_1, y_1) = (2, 1)$, $(u_1, v_1) = (1, 2)$, and $(u_2, v_2) = (4, 3)$. All solutions $u + v\sqrt{3}$ of

(5), with u and v positive integers, are then given by one of the two following classes of solutions.

Class 1:

$$T_1(n) = (u_1 + v_1\sqrt{3})(x_1 + y_1\sqrt{3})^n = (1 + 2\sqrt{3})(2 + \sqrt{3})^n, \qquad n \geq 0;$$

Class 2:

$$T_2(n) = (u_2 + v_2\sqrt{3})(x_1 + y_1\sqrt{3})^n = (4 + 3\sqrt{3})(2 + \sqrt{3})^n, \qquad n \geq 0.$$

For example $(17)^2 - 3(10)^2 = -11$, hence $u = 17$ and $v = 10$ is a solution of (5). This solution, $17 + 10\sqrt{3}$, is in Class 2 with $n = 1$, since $T_2(1) = (4 + 3\sqrt{3})(2 + \sqrt{3}) = 17 + 10\sqrt{3}$. For this triangle $r = (v + 2u)/11 = 4$, $b = r + v = 14$, $a = b - 1 = 13$, $c = b + 1 = 15$, $s = 21$, and $A = rs = 84$. It is the second Brahmagupta triangle, which we mentioned above.

Focusing on Class 1, we have $T_1(0) = (1 + 2\sqrt{3})$, $T_1(1) = (1 + 2\sqrt{3})(2 + \sqrt{3})$, $T_1(n-1) = (2 + \sqrt{3})T_1(n-2)$, $n \geq 3$, and $T_1(n) = (2 + \sqrt{3})T_1(n-1)$, $n \geq 2$. Combining the expressions relating $T_1(n)$, $T_1(n-1)$, and $T_1(n-2)$ gives $T_1(n) = (2 + \sqrt{3})^2 T_1(n-2) = (8 + 4\sqrt{3})T_1(n-2) - T_1(n-2) = 4(2 + \sqrt{3})T_1(n-2) - T_1(n-2)$. Thus $T_1(n)$ satisfies the difference equation

$$T_1(n) = 4T_1(n-1) - T_1(n-2), \quad n \geq 2.$$

If T_1 is expressed as $T_1(n) = u_1(n) + v_1(n)\sqrt{3}$, then $u_1(n)$, $v_1(n)$, $r_1(n) = (v_1(n) + 2u_1(n))/11$ and $s_1(n) = 3(r_1(n) + v_1(n))/2 = (18v_1 + 3u_1)/11$ are also determined by the difference equation.

From the known values of $u_1(0)$, $v_1(0)$, $u_1(1)$, and $v_1(1)$, we can calculate the first two cases of r, s, and b. For example, since $u_1(0) = 1$ and $v_1(0) = 2$, we have $r_1(0) = (v_1(0) + 2u_1(0))/11 = 4/11$. Also $u_1(1) = 8$ and $v_1(1) = 5$ and thus $r_1(1) = 21/11$. The difference equation then generates the values of u, v, r, s, and b for all $n \geq 2$. Since $A = rs$, we also have the values of A for all n. Similar calculations give corresponding results for the triangles in Class 2. Class 1 consists of an infinite number of triangles and the values of $r_1(n)$, $s_1(n)$, $b_1(n)$, and $A_1(n)$ are all ratios of integers and therefore rational.

We now show that the $r_1(n)$ cannot be integers. If $p_1(n) = 11r_1(n) = v_1(n) + 2u_1(n)$, then $p_1(n)$ satisfies the difference equation. The first 12 least residues (mod 11) of $p_1(n)$ are (4, 10, 3, 2, 5, 7, 1, 8, 9, 6, 4, 10). The least residues at $n = 0$ and at $n = 10$ are equal, also the least residues at $n = 1$ and at $n = 11$ are equal. Since the least residues satisfy the difference equation, they repeat in blocks of length 10, and thus they are never 0. Hence $p_1(n)$ is never divisible by 11 and the values of $r_1(n)$, $n \geq 0$, are not integers. Similar calculations show that $s_1(n)$ and $A_1(n)$ are never integers.

For Class 2 triangles, the values of $r_2(0)$ and $r_2(1)$ are both integers and satisfy the difference equation; thus $r_2(n)$ are integers for $n \geq 0$. Since there are no triangles with integer inradius in Class 1, they are all in Class 2, and thus Class 2 triangles are all of the triangles that have consecutive integer side lengths and integer inradii. Hence triangles with consecutive integer side lengths are in Class 2 if and only if they have integer inradius. By Proposition 1, the areas must also be integers. These are exactly the Brahmagupta triangles.

Summarizing, we have the following propositions for H.A.P. triangles whose side lengths differ from their inradii by an integer.

PROPOSITION 2. (CLASS 1) *All H.A.P. triangles with side lengths $b(n) - 1$, $b(n)$, $b(n) + 1$, where the $b(n)$ are rational non-integers and n an integer, $n \geq 2$, are generated by applying the difference equation*

$$b(n) = 4b(n-1) - b(n-2)$$

with $b(0) = 26/11$ and $b(1) = 76/11$.

PROPOSITION 3. (CLASS 2) *All H.A.P. triangles with side lengths $b(n) - 1$, $b(n)$, $b(n) + 1$, where the b's and n are positive integers, are generated by applying the same difference equation with $b(0) = 4$ and $b(1) = 14$.*

The first example of the introduction is in Class 2 with $n = 0$. The third example is in Class 1 with $n = 0$. Not all H.A.P. triangles with side lengths differing from each other by 1 are included in Class 1. For example, the triangle with side lengths $(17/11, 28/11, 39/11)$ and $r = 5/11$ is not included in Class 1, since the side lengths and inradius do not differ by an integer.

TABLE 2 gives the values of b, A, and r for triangles in each class and for $0 \leq n \leq 3$. The initial conditions are the values at $n = 0$ and $n = 1$.

TABLE 2: Properties of H.A.P. triangles with $d = 1$, for small values of p and q

	b	r	A
Class 1:			
$n = 0$	26/11	4/11	156/121
$n = 1$	76/11	21/11	2394/121
$n = 2$	278/11	80/11	33360/121
$n = 3$	1036/11	299/11	464646/121
Class 2:			
$n = 0$	4	1	6
$n = 1$	14	4	84
$n = 2$	52	15	1170
$n = 3$	194	56	16296

Acknowledgment Our thanks to Underwood Dudley and John Rickert for very helpful suggestions.

REFERENCES

1. R. Beauregard and E. Suryanarayan, The Brahmagupta triangles, *College Math J.* **29** (1998) 13–17. http://dx.doi.org/10.2307/2687630
2. C. R. Fleenor, Heronian triangles with consecutive integer sides, *Jour. Rec. Math.* **28** (1996) 113–115.
3. J. A. MacDougall, Heron triangles with sides in arithmetic progression, *Jour. Rec. Math.* **31** (2003) 189–196.
4. Trygve Nagell, *Introduction to Number Theory,* 2nd ed., Chelsea, New York, 1964.

Summary In this note we give a method to generate all Heronian triangles with sides in arithmetic progression (H.A.P. triangles) and show that all Brahmagupta triangles can be generated as solutions of a difference equation with certain initial conditions. The same difference equation with different initial conditions generates all rational non-integer solutions, if H.A.P. triangles whose side lengths differ from each other by 1 and also differ from the inradius by integers.

PROBLEMS

BERNARDO M. ÁBREGO, *Editor*
California State University, Northridge

Assistant Editors: SILVIA FERNÁNDEZ-MERCHANT, California State University, Northridge; JOSÉ A. GÓMEZ, Facultad de Ciencias, UNAM, México; EUGEN J. IONASCU, Columbus State University; ROGELIO VALDEZ, Facultad de Ciencias, UAEM, México; WILLIAM WATKINS, California State University, Northridge

PROPOSALS

To be considered for publication, solutions should be received by March 1, 2013.

1901. *Proposed by Ángel Plaza, and César Rodríguez, Department of Mathematics, Universidad de las Palmas de Gran Canaria, Las Palmas, Spain.*

Let $f:[a,b]\to\mathbb{R}$ be an n times continuously differentiable function on (a,b), and let $a_1, a_2, \ldots, a_{n+1}$ be $n+1$ distinct numbers in (a,b). Prove that there exists c in (a,b) such that

$$\sum_{j=1}^{n+1} f(a_j) \prod_{\substack{1\le i\le n+1\\ i\neq j}} (a_j - a_i)^{-1} = \frac{f^{(n)}(c)}{n!}.$$

1902. *Proposed by H. A. ShahAli, Tehran, Iran.*

Let $a_0, a_1, \ldots, a_{2n}$ be positive real numbers. Prove that there are at least $(n!)^2$ different permutations σ of $\{0, 1, \ldots, 2n\}$ such that

$$a_{\sigma(0)} + a_{\sigma(1)}x + \cdots + a_{\sigma(2n)}x^{2n} > 0$$

for all $x \in \mathbb{R}$.

1903. *Proposed by Mihály Bencze, Brasov, Romania.*

Let x, y, $a_1 = a_4$, $a_2 = a_5$, and a_3 be positive real numbers. Prove that

$$8(a_1^3 + a_2^3 + a_3^3) \ge \sum_{k=1}^{3} \Big[(a_k + a_{k+1})^x (a_k + a_{k+2})^y\Big]^{\frac{3}{x+y}}$$
$$\ge 3(a_1 + a_2)(a_2 + a_3)(a_3 + a_1).$$

Math. Mag. **85** (2012) 295–301. doi:10.4169/math.mag.85.4.295.

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.

1904. *Proposed by Oskar Maria Baksalary, Adam Mickiewicz University, Poznań, Poland, and Götz Trenkler, Dortmund University of Technology, Dortmund, Germany.*

Let A be an $n \times n$ complex matrix of rank r, $0 < r \leq n$, and let s be the sum of the elements of the antidiagonal of A; that is, if $A = (a_{ij})$, then $s = a_{1,n} + a_{2,n-1} + \cdots + a_{n,1}$. Prove that if A is idempotent and Hermitian (i.e., an orthogonal projector), then $|s| < \sqrt{nr}$.

1905. *Proposed by Luis Gonzales, Maracaibo, Venezuela, and Cosmin Pohoata, Princeton University, Princeton, NJ.*

Let ℓ be an arbitrary line in the plane of a given triangle ABC. The three lines obtained as the reflections of ℓ with respect to the sidelines of $\triangle ABC$ bound a triangle, namely $\triangle XYZ$.

(a) Prove that the incenter of $\triangle XYZ$ lies on the circumcircle of $\triangle ABC$.

(b) If U and V are the incenter and the circumcenter of triangle XYZ, respectively, prove that the (second) intersection of the line UV with the circumcircle of $\triangle ABC$ is a fixed point E, independent of the position of ℓ. Moreover, prove that E is the Euler reflection point of $\triangle ABC$ (i.e., E is the intersection of the reflections of the Euler Line of $\triangle ABC$ with respect to its sidelines).

Quickies

Answers to the Quickies are on page 301.

Q1023. *Proposed by Dale Geer, Oshkosh, WI.*

Let x and y be positive real numbers such that $1 < x < y$ and $x^y = y^x$. For example, $(x, y) = (2, 4)$ or $(x, y) = (2.25, 3.375)$. Prove that $x < e < y$.

Q1024. *Proposed by Mowaffaq Hajja and Ahmad Hamdan, Mathematics Department, Yarmouk University, Irbid, Jordan*

A subring of a ring with multiplicative identity e can have a multiplicative identity different from e. For every positive integer n, find the number of subrings of $\mathbb{Z}/n\mathbb{Z}$ (counting $\mathbb{Z}/n\mathbb{Z}$ and the zero ring) that have a multiplicative identity.

Solutions

Factoring a homogeneous polynomial **October 2011**

1876. *Proposed by Roman Wituła, Edyta Hetmaniok, and Damian Słota, Institute of Mathematics, Silesian University of Technology, Gliwice, Poland.*

Prove that the following equality holds for $x, y \in \mathbb{C}$ and n a positive integer.

$$x^{2n} - x^n y^n + y^{2n} = \prod_{\substack{1 \leq k < 3n \\ \gcd(k,6)=1}} \left(x^2 - 2\cos\left(\frac{k\pi}{3n}\right) xy + y^2 \right).$$

Solution by Northwestern University Math Problem Solving Group, Department of Mathematics, Evanston, IL.

There is exactly one k verifying $\gcd(k, 6) = 1$ in $\{3m, 3m+1, 3m+2\}$, hence we have that the product on the right-hand side has exactly n factors, and both sides are homogeneous polynomials in x, y of degree $2n$.

If any of x or y is zero then the equality becomes trivial, so we may assume that $x, y \neq 0$. Dividing by y^{2n} and letting $z = x/y$ we can rewrite the equality like this:

$$z^{2n} - z^n + 1 = \prod_{\substack{1 \le k < 3n \\ \gcd(k,6)=1}} \left(z^2 - 2\cos\left(\frac{k\pi}{3n}\right) z + 1 \right).$$

Finally we prove that those two (monic) polynomials in z are equal by checking that they have the same roots, namely the $2n$ distinct complex numbers given by $z = e^{\pm k\pi i/3n}$, for $1 \le k < 3n$ and $\gcd(k, 6) = 1$. Substituting $z = e^{\pm k\pi i/3n}$ in the left-hand side gives

$$\begin{aligned} e^{\pm 2nk\pi i/3n} - e^{\pm nk\pi i/3n} + 1 &= e^{\pm k\pi i/3}(e^{\pm k\pi i/3} + e^{\mp k\pi i/3} - 1) \\ &= e^{\pm k\pi i/3} \left(2\cos\left(\pm\frac{k\pi}{3} \right) - 1 \right), \end{aligned}$$

and since $k \equiv \pm 1 \pmod 6$, then

$$e^{\pm 2nk\pi i/3n} - e^{\pm nk\pi i/3n} + 1 = e^{\pm k\pi i/3} \left(2\cos\left(\pm\frac{\pi}{3} \right) - 1 \right) = 0.$$

In the right-hand side, the same substitution on the following factor gives

$$\begin{aligned} z^2 - 2\cos\left(\frac{k\pi}{3n}\right) z + 1 &= e^{\pm 2k\pi i/3n} - 2\cos\left(\frac{k\pi}{3n}\right) e^{\pm k\pi i/3n} + 1 \\ &= e^{\pm k\pi i/3n} \left(e^{\pm k\pi i/3n} - 2\cos\left(\frac{k\pi}{3n}\right) + e^{\mp 2k\pi i/3n} \right) \\ &= e^{\pm k\pi i/3n} \left(2\cos\left(\frac{k\pi}{3n}\right) - 2\cos\left(\frac{k\pi}{3n}\right) \right) = 0, \end{aligned}$$

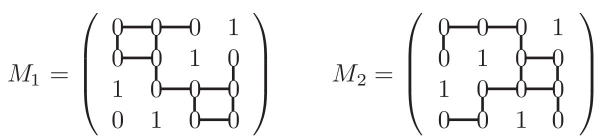
and this completes the proof.

Also solved by Michel Bataille (France), Claude Bégin (Canada), Francesco Bonesi (Italy) and Lorenzo Luzzi (Italy), Bruce S. Burdick, Robert Calcaterra, Stefan Chaładus (Poland), Hongwei Chen, Con Amore Problem Group (Denmark), Paul Deiermann, Robert L. Doucette, Dmitry Fleischman, Michael Goldenberg and Mark Kaplan, J. A. Grzesik, Eugene A. Herman, In Jae Hwang (Korea), John C. Kieffer, Tek Min Kim (Korea), Omran Kouba (Syria), Elias Lampakis (Greece), László Lipták, Peter McPolin (Northern Ireland), Jeremy Moore and John Zacharias, Ángel Plaza (Spain), Joel Schlosberg, Nicholas C. Singer, John H. Smith, H. T. Tang, Traian Viteam (Germany), Michael Vowe (Switzerland), Haohao Wang and Jerzy Wojdyło, and the proposers.

Disconnecting a permutation matrix — October 2011

1877. *Proposed by Daniel Edelman, Mason–Rice Elementary School, Newton Centre, MA, and Alan Edelman, MIT, Cambridge, MA.*

For each $n \times n$ permutation matrix M, consider the graph G_M where the vertices are the zero entries and two vertices are adjacent if their corresponding entries in the matrix are adjacent horizontally or vertically. We say that M *disconnects its zeros*, if G_M is disconnected. For example, M_1 has the bottom left zero disconnected, while M_2 does not disconnect its zeros.



Find a formula for the number of $n \times n$ permutation matrices that disconnect their zeros. Also find an asymptotic formula as $n \to \infty$ for the fraction of the $n!$ permutation matrices that disconnect their zeros.

Solution by Matteo Elia, Antonello Cirulli, and Cesare Gallozzi, Università di Roma "Tor Vergata", Roma, Italy.
Let $I_k = [a_{ij}]$ be the $k \times k$ matrix such that $a_{ij} = 1$ if $i = j$ and $a_{ij} = 0$ otherwise; and let $I'_k = [a_{ij}]$ be the $k \times k$ matrix such that $a_{ij} = 1$ if $i = k + 1 - j$ and $a_{ij} = 0$ otherwise. It is easy to verify that a permutation matrix M disconnects its zeros if and only if one and only one of the following conditions is satisfied:

1. M has a block I_k in the upper right corner or a block I_j in the lower left corner with $k \geq 2$, $j \geq 2$, and $k + j \leq n$;
2. M has a block I'_k in the upper left corner or a block I'_j in the lower right corner with $k \geq 2$, $j \geq 2$, and $k + j \leq n$.

Given one block I_k or I'_k in one of the four corners, the rest of the permutation matrix can be filled in $(n-k)!$ ways. Given two blocks I_k and I_j or I'_k and I'_j, the rest of the permutation matrix can be filled in $(n-k-j)!$ ways. Hence, by applying the inclusion–exclusion principle, the desired number is given by the following formula

$$d(n) = 4\sum_{k=2}^{n-1}(n-k)! - 2\sum_{k=2}^{n-2}\sum_{j=2}^{n-k}(n-k-j)! + 2 = 4\sum_{i=1}^{n-2} i! - 2\sum_{k=2}^{n-2}\sum_{i=0}^{n-k-2} i! + 2$$

$$= 4\sum_{i=1}^{n-2} i! - 2\sum_{i=0}^{n-4} i!(n-2-(i+1)) + 2$$

$$= 4\sum_{i=1}^{n-2} i! - 2(n-2)\sum_{i=0}^{n-4} i! + 2\sum_{i=1}^{n-3} i! + 2.$$

The values of $d(n)$ for $n = 2, \ldots, 10$ are 2, 6, 12, 32, 120, 580, 3392, 23244, and 182776. Since $\sum_{i=1}^{n} i!/n! \sim 1 + 1/n$, it follows that

$$\frac{d(n)}{n!} = \frac{4}{n^2} + \frac{8}{n^3} + \Theta\left(\frac{1}{n^4}\right).$$

Also solved by László Lipták, John Kieffer, Eugene A. Herman, Dmitry Fleischman, Robert Calcaterra, and the proposers. There were 4 incorrect or incomplete submissions.

An A.M.–G.M. inequality of ratios **October 2011**

1878. *Proposed by Pantelimon George Popescu, Politechnica University, Bucharest, Romania and José Luis Díaz-Barrero, Polytechnical University of Catalonia, Barcelona, Spain.*

Let $\alpha \geq -1$ and $\beta \geq 1$ be real numbers. Let $\{b_k : 1 \leq k \leq n\}$ be a set of real numbers in the interval $(0, 1]$. Prove that

$$\frac{A-\alpha}{A+\beta} \geq \frac{1}{n}\sum_{k=1}^{n}\frac{b_k-\alpha}{b_k+\beta} \geq \frac{G-\alpha}{G+\beta},$$

where A and G are the arithmetic and geometric mean, respectively, of the set $\{b_k : 1 \leq k \leq n\}$.

I. *Solution by Robert L. Doucette, Department of Mathematics, Computer Science, and Statistics, McNeese State University, Lake Charles, LA.*
The function $f(x) = (x-\alpha)/(x+\beta)$ has second derivative $f''(x) = -2(\alpha+\beta)(x+\beta)^{-3}$ and, since $\alpha+\beta \geq 0$, f is concave on the interval $(-\beta, \infty)$. The function $g(x) = f(e^x)$ has second derivative $g''(x) = (\alpha+\beta)e^x(\beta - e^x)(e^x+\beta)^{-3}$ and so g is convex on the interval $(-\infty, \ln\beta)$. Because $b_k \in (-\beta, \infty)$ and $\ln b_k \in (-\infty, \ln\beta)$, it follows by Jensen's Inequality that

$$f(A) \geq \frac{1}{n}\sum_{k=1}^{n} f(b_k) \quad \text{and} \quad \frac{1}{n}\sum_{k=1}^{n} g(\ln b_k) \geq g\left(\frac{1}{n}\sum_{k=1}^{n}\ln b_k\right).$$

Note that $g(\ln b_k) = f(b_k)$ and $g(\frac{1}{n}\sum_{k=1}^{n}\ln b_k) = g(\ln(\prod_{k=1}^{n} b_k)^{1/n}) = f(G)$. Thus

$$\frac{A-\alpha}{A+\beta} = f(A) \geq \frac{1}{n}\sum_{k=1}^{n}\frac{b_k-\alpha}{b_k+\beta} = \frac{1}{n}\sum_{k=1}^{n} f(b_k) \geq f(G) = \frac{G-\alpha}{G+\beta}.$$

II. *Solution by Michael Vowe, Therwil, Switzerland.*
The first inequality is proved as in the first solution. For the second inequality we use that $1/(b_k+\beta) \geq 1/(1+\beta)$ and that $A \geq G$ by the Arithmetic Mean–Geometric Mean Inequality. Therefore

$$\begin{aligned}\sum_{k=1}^{n}\left(\frac{b_k-\alpha}{b_k+\beta} - \frac{G-\alpha}{G+\beta}\right) &= \frac{\alpha+\beta}{G+\beta}\sum_{k=1}^{n}\frac{b_k-G}{b_k+\beta} \geq \frac{\alpha+\beta}{(G+\beta)(1+\beta)}\sum_{k=1}^{n}(b_k-G)\\ &= \frac{\alpha+\beta}{(G+\beta)(1+\beta)}\cdot n(A-G) \geq 0,\end{aligned}$$

which is equivalent to

$$\frac{1}{n}\sum_{k=1}^{n}\frac{b_k-\alpha}{b_k+\beta} \geq \frac{G-\alpha}{G+\beta}.$$

Also solved by Michel Bataille (France), Matteo Elia (Italy), Dmitry Fleischman, Michael Goldenberg and Mark Kaplan, Ahmad Habil (Syria, only left inequality), John C. Kieffer, Tek Min Kim (Korea), Omran Kouba (Syria), Elias Lampakis (Greece), Paolo Perfetti (Italy), Greg Ronsse, Nicholas C. Singer, Traian Viteam (Germany), and the proposers.

Using residues to integrate **October 2011**

1879. *Proposed by Wong Fook Sung, Temasek Polytechnic, Singapore.*

Let m and n be positive integers such that $m < n$ and let a and b be positive real numbers. Evaluate

$$\int_0^\infty \frac{x^{2(n-m)}(x^2-1)^{2m}}{ax^{2n}+b(x^2-1)^{2n}}\,dx.$$

Solution by John Zacharias, Melbourne, FL.
The requested integral can be written as

$$\begin{aligned}\int_0^\infty \frac{x^{2(n-m)}(x^2-1)^{2m}}{ax^{2n}+b(x^2-1)^{2n}}\,dx &= \int_0^\infty \frac{(x-1/x)^{2m}}{a+b(x-1/x)^{2n}}\,dx\\ &= \int_0^1 \frac{(x-1/x)^{2m}}{a+b(x-1/x)^{2n}}\,dx + \int_1^\infty \frac{(x-1/x)^{2m}}{a+b(x-1/x)^{2n}}\,dx. \qquad (1)\end{aligned}$$

In the second integral we use the change of variable $u = 1/x$ to get

$$\int_1^\infty \frac{(x-1/x)^{2m}}{a+b(x-1/x)^{2n}}\,dx = \int_0^1 \frac{(1/u-u)^{2m}(u^{-2})}{a+b(1/u-u)^{2n}}\,du.$$

Using this in Equation (1), the original integral becomes

$$\int_0^1 \frac{(x-1/x)^{2m}}{a+b(x-1/x)^{2n}}\,dx + \int_0^1 \frac{(x-1/x)^{2m}(x^{-2})}{a+b(x-1/x)^{2n}}\,dx$$
$$= \int_0^1 \frac{(1+x^{-2})(x-1/x)^{2m}}{a+b(x-1/x)^{2n}}\,dx.$$

Now the change of variable $u = x - 1/x$ transforms this into

$$\int_0^\infty \frac{u^{2m}}{a+bu^{2n}}\,du = \frac{1}{2b}\int_{-\infty}^\infty \frac{x^{2m}}{(a/b)+x^{2n}}\,dx.$$

For $r > 0$, define

$$F(r) = \int_{-\infty}^\infty \frac{x^{2m}}{r+x^{2n}}\,dx.$$

We evaluate $F(r)$ by the method of residues and contour integration. That is,

$$F(r) = \int_{-\infty}^\infty \frac{x^{2m}}{r+x^{2n}}\,dx = \lim_{R\to\infty}\int_{C_R} \frac{z^{2m}}{r+z^{2n}}\,dz = 2\pi i \sum_{k=1}^{n} \operatorname{Res}\left(\frac{z^{2m}}{r+z^{2n}}, z_k\right),$$

where C_R is the upper half of the circle of radius R centered at the origin together with the part of the real axis that is a diameter of C_R and $\{z_k\}$ are the roots of the denominator of $z^{2m}/(r+z^{2n})$ that lie in the upper half-plane. These roots are $z_k = r^{1/2n}\exp((2k-1)\pi i/2n)$, $k = 1, 2, \ldots, n$, and they are all simple poles. Therefore

$$F(r) = 2\pi i \sum_{k=1}^{n} \operatorname{Res}\left(\frac{z^{2m}}{r+z^{2n}}, z_k\right) = 2\pi i \sum_{k=1}^{n} \frac{z_k^{2m}}{2nz_k^{2n-1}}$$
$$= \frac{\pi i}{n}\sum_{k=1}^{n} z_k^{2m+1-2n} = \frac{\pi i}{n}\sum_{k=1}^{n}\left(r^{1/2n}\exp\left(\frac{2k-1}{2n}\pi i\right)\right)^{2m-2n+1}$$
$$= \frac{\pi i}{n} r^{\frac{2m+1-2n}{2n}} \exp\left(\frac{-(2m-2n+1)}{2n}\pi i\right)\sum_{k=1}^{n}\left[\exp\left(\frac{2m-2n+1}{n}\pi i\right)\right]^k$$
$$= \frac{\pi i r^{\frac{2m+1}{2n}}}{nr}\cdot i\csc\left(\frac{2m-2n+1}{2n}\pi\right) = \frac{\pi r^{\frac{2m+1}{2n}}}{nr}\csc\left(\frac{2m+1}{2n}\pi\right).$$

Therefore the original integral equals

$$\frac{F(a/b)}{2b} = \int_0^\infty \frac{x^{2(n-m)}(x^2-1)^{2m}}{ax^{2n}+b(x^2-1)^{2n}}\,dx = \frac{\pi}{2na}\cdot\left(\frac{a}{b}\right)^{\frac{2m+1}{2n}}\cdot\csc\left(\tfrac{2m+1}{2n}\pi\right).$$

Also solved by Michel Bataille (France), Hongwei Chen, Joonyong Choi (Korea),Paul Deiermann, Robert L. Doucette, Michael Goldenberg and Mark Kaplan, Eugene A. Herman, John C. Kieffer, Omran Kouba (Syria), Elias Lampakis (Greece), Kee-Wai Lau (China), Bob Mallison, Matthew McMullen, Samih Obaid, Paolo Perfetti (Italy), Nicholas C. Singer, Traian Viteam (Germany), Michael Vowe (Switzerland), and the proposer.

From exponential to geometric distributions **October 2011**

1880. *Proposed by Richard Stephens, Department of Mathematics, Columbus State University, Columbus, GA.*

Let X be a positive continuous random variable and for any $\alpha > 0$, let Y_α be the random variable defined by $Y_\alpha = n$ for $n = 1, 2, 3, \ldots$ if and only if $(n-1)\alpha < X < n\alpha$. Prove that X has an exponential distribution if and only if Y_α has a geometric distribution for every $\alpha > 0$.

Solution by Robert Calcaterra, University of Wisconsin–Platteville, Platteville, WI.
Suppose X has an exponential distribution with mean μ, $\alpha > 0$, and $p = 1 - e^{-\alpha/\mu}$. Then

$$P\left(Y_\alpha = n\right) P((n-1)\alpha < X < n\alpha) = -e^{-n\alpha/\mu} + e^{-(n-1)\alpha/\mu} = p(1-p)^{n-1}$$

and so Y_α has a geometric distribution.

Next, suppose the probability function of Y_α is $p_\alpha(1-p_a)^{n-1}$ for some $p_\alpha \in (0, 1)$. Let F be the cumulative distribution function of X, $x > 0$, m and n positive integers, and $q = n/m$. Then

$$\begin{aligned} F(x) = P(X < n(x/n)) &= P(Y_{x/n} \le n) \\ &= \sum_{k=1}^{n} p_{x/n}(1 - p_{x/n})^{k-1} = 1 - (1 - p_{x/n})^n. \end{aligned}$$

Hence, $1 - F(q) = (1 - p_{1/m})^n = ((1 - p_{1/m})^m)^q = (1 - F(1))^q$. But the rational numbers are dense in the real numbers and thus $1 - F(x) = (1 - F(1))^x$ for all positive real numbers x. Therefore, F is the cumulative distribution function of an exponential random variable and the proof is complete.

Also solved by Francesco Bonesi, Matteo Elia, and Alessio Podda (Italy); Bruce S. Burdick; Dmitry Fleishman; John C. Kieffer; Omran Kouba (Syria); Greg Ronsse; Carl M. Russell; Joel Schlosberg; and the proposer. There was one incomplete solution.

Answers

A1023. Since x and y are positive, $x^y = y^x$ if and only if $x \ln y = y \ln x$. It follows that $\ln x/x = \ln y/y$, and that equation further implies that

$$\frac{\ln y - \ln x}{y - x} = \frac{\ln x}{x} = \frac{\ln y}{y}.$$

Let $f(t) = \ln t$. By the Mean Value Theorem, the expression on the left is equal to $f'(z) = 1/z$, for some z such that $x < z < y$. Thus $z = x/\ln x = y/\ln y$ and because $0 < x < z < y$, it follows that $x/\ln x > x$ and $y/\ln y < y$. Thus $\ln x < 1 < \ln y$ and so $x < e < y$.

A1024. The subrings of $\mathbb{Z}/n\mathbb{Z}$ are precisely $d(\mathbb{Z}/n\mathbb{Z})$, where d divides n. Suppose that $n = dm$.

If $d(\mathbb{Z}/n\mathbb{Z})$ has a (multiplicative) identity ud, then $(ud)d \equiv d \pmod{n}$ and hence $ud \equiv 1 \pmod{m}$. Therefore $\gcd(d, m) = 1$. Conversely, if $\gcd(d, m) = 1$, then there exists u such that $ud \equiv 1 \pmod{m}$. Then for every $a \in \mathbb{Z}/n\mathbb{Z}$, $(ud)(ad) \equiv ad \pmod{n}$, and thus ud is an identity of $d(\mathbb{Z}/n\mathbb{Z})$.

Therefore $d(\mathbb{Z}/n\mathbb{Z})$ has an identity if and only if $\gcd(d, n/d) = 1$. The number of such divisors d (including 1 and n) is clearly the number of subsets of the set of prime divisors of n, i.e., 2^t, where t is the number of prime divisors of n.

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.

Kaplan, Daniel T., *Statistical Modeling: A Fresh Approach*, 2nd ed., Project Mosaic, 2011; 388 pp, $54.95 (P). ISBN 978-0-9839658-7-9.

I have taught introductory statistics and mathematical statistics many times. The standard syllabi are heavy on calculation and technique, just like those for calculus and for differential equations. Students can "succeed" in any of those courses without becoming able to apply their learning to a real-world situation. What tends to be missing in such courses is the learning of *modeling*, the step from a real-world problem to a mathematical formulation. It is a hard step, because it requires—besides a mathematical toolbox—some understanding of the context of the particular problem, plus the interest and curiosity to try to solve it. This book teaches the modeling step in statistics—in marvelous fashion!—without assuming any previous knowledge of statistics. The last section of each chapter is devoted to "computational technique," but there are no traditional statistical formulas. Computations are accomplished in a carefully chosen subset of R, through the front end RStudio (both are free for all platforms, hence have the advantage of being accessible to students throughout their lives). Each chapter has a few reading questions; exercises, software, and datasets are available at `http://www.mosaci-web.org/StatisticalModeling/` .

Adam, John A., *X and the City: Modeling Aspects of Urban Life*, Princeton University Press, 2012; xviii + 319 pp, $29.95. ISBN 978-0-691-15464-0.

The author of *A Mathematical Nature Walk* (2009) takes to the streets to model aspects of city life: its size, chance encounters, estimation problems (how many squirrels in Central Park?, etc.), average trip length, car-following, and more. Lots of math gets used (calculus and more)! (The author has just received an Allendoerfer Award for a recent article in this MAGAZINE; see page 319.)

Posamentier, Alfred S., and Ingmar Lehmann, *The Secrets of Triangles: A Mathematical Journey*, Prometheus Books, 2012; 321 pp, $26. ISBN 978-1-61614-587-3.

This is a book for those who love plane geometry, have a good appetite for proof, and can interpret intricate diagrams. At a few points, trigonometry is used, as well as logarithms (in connection with fractal dimension). There is more here than you ever knew about triangles!

Stewart, Ian, *In Pursuit of the Unknown: 17 Equations That Changed the World*, Basic Books, 2012; x + 342 pp, $26.99. ISBN 978-0-465-029723-0.

The first equation in the book is the Pythagorean theorem, and the last is the Black-Scholes equation ("the Midas formula"). In between, you will find 15 more of your favorites, with their significance described and expounded in enlightening fashion by a author well known for his "genius for explanation." The book is for the general reader; virtually no other equations appear except the ones celebrated.

Math. Mag. **85** (2012) 302–303. `doi:10.4169/math.mag.85.4.302`.

Mackenzie, Dana, *The Universe in Zero Words: The Story of Mathematics as Told through Equations*, Princeton University Press, 2012; 224 pp, $27.95. ISBN 978-0-691-15282-0.

Compared to Stewart's book reviewed above, this book includes more equations (24, starting with $1 + 1 = 2$), has a lavish layout with color illustrations (are some colorized?), is roughly the same length, and costs about as much. But there is little overlap, because this book is devoted to equations that had an effect on mathematics. Hence here Fermat's Last Theorem occurs, as does the Galois group of a polynomial, the prime number theorem, and the continuum hypothesis.

Haldane, Andrew G., and Robert M. May, Systemic risk in banking ecosystems, *Nature* 469 (30 January 2011) 351–355.

Stewart's book reviewed above mentions this article, which compares the banking system to biological ecosystems. Haldane and May assert that in both kinds of systems,"too much complexity implies instability." While biological systems have been subjected to survival of the fittest, evolution of financial systems has often proceeded according to "survival of the fattest." For example, U.S. policy has been to save individual banks, without considering equilibrium and risk of the system as a whole. "[A]ll banks doing the same thing—can minimize risk for each individual bank, but maximize the probability of the entire system collapsing." The authors conclude with public policy recommendations.

Weinstein, Lawrence, *Guesstimation 2.0: Solving Today's Problems on the Back of a Napkin*, Princeton University Press, 2012; xv + 353 pp, $19.95 (P). ISBN 978-0-691-15080-2.

Paper or plastic? How much energy does it take to transport 1 ton of cargo across the U.S. by car, truck bicycle, train, or plane? or to ship one year's worth of a person's groceries? How easy is it to shear the head off a $5/16''$ lag screw, as the author did? This book, a successor to *Guesstimation* (2009) (by the same author with John A. Adam), is another book of problems about estimation and approximation ("Fermi" problems). My favorite (I've always enjoyed physics): "Given that radiation from proton decay does not kill us, what is the minimum possible proton lifetime?" Fortunately, the author gives hints and answers.

Langkamp, Greg, and Joseph Hull, *Quantitative Reasoning and the Environment: Mathematical Modeling in Context*, Pearson Prentice-Hall, 2007; $93.33; xiv + 356 pp (P). ISBN 0-13-148257-X.

In the recent college tradition of entrance requirements morphing into graduation requirements, mathematics remediation has shape-shifted to a graduation requirement in QL/QR (quantitative literacy/quantitative reasoning). Indeed, the MAA has a Special Interest Group on Quantitative Literacy and has published several accounts about what it is, why it matters, and how to achieve it. This book lives up to its title, in terms of both involvement with environmental issues and fostering mathematical modeling. The first chapter, on measurements and units, has examples featuring gasoline additives, oil spills, pH, and earthquakes; it begins with an account of Chinese coal-fired power plants and the resulting mercury in the diet of the Inuit of Greenland. Part 1 ("Essential Numeracy") of the book treats units, scientific notation, ratios, percentages, logarithms; Part 2 ("Function Modeling") explores linear functions (including line of best fit) and exponential and power functions (and fractals); Part 3 ("Difference Equation Modeling") introduces sequences, writing difference equations, solving recurrences, logistic difference equations, and systems models; and Part 4 ("Elementary Statistics") does the usual descriptive statistics, normal distributions, and confidence intervals. All of this is done in the lively context of an astonishing variety of environmental issues (ice cap melting, garbage disposal, prairie dog survival, hurricane frequency, water use, manatee mortality, energy demand, microloans, airborne lead, measles) and concomitant modeling problems (equilibria, resource harvesting, chaotic behavior, population dynamics, epidemics). I marvel at the magnificent cornucopia of issues brought up, and I think students will be impressed that the mathematics that they are learning can play a role. The exercises are rooted in real situations and use real data; there are no pretend applications with made-up numbers. A supplementary Website at `www.enviromath.com` contains data set files plus source material for chapter projects, including maps, data, and TI calculator programs (and even field exercises). Of course, this would make an outstanding textbook for a high school mathematics course, too.

NEWS AND LETTERS

41st USA Mathematical Olympiad
3rd USA Junior Mathematical Olympiad

JACEK FABRYKOWSKI
Youngstown State University
Youngstown, OH 44555
jfabrykowski@ysu.edu

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

The MAA gave the 41st United States of America Mathematical Olympiad (USAMO) and 3rd United States of America Junior Mathematical Olympiad (USAJMO) competitions to 267 and 243 students respectively on April 24 and 25, 2012. The event was conducted by the Committee on the American Mathematics Competitions. For the third time, the USAJMO was offered for students in 10th grade and below, providing a nicely balanced link between the computational character of the AIME problems and proof-oriented problems of USAMO. Both Mathematical Olympiads contained three problems for each of two days, with an allowed time of 4.5 hours each day. Problems JMO1, JMO3 on Day 1, and problems JMO4, JMO5 on Day 2 were different from the USAMO problems, but JMO2 and JMO6 were the same as USAMO1 and USAMO5 respectively.

USAMO Problems

1. Find all integers $n \geq 3$ such that among any n positive real numbers $a_1, a_2, \ldots, a_n$ with

$$\max(a_1, a_2, \ldots, a_n) \leq n \cdot \min(a_1, a_2, \ldots, a_n),$$

there exist three that are the side lengths of an acute triangle.

2. A circle is divided into 432 congruent arcs by 432 points. The points are colored in four colors such that some 108 points are colored Red, some 108 points are colored Green, some 108 points are colored Blue, and the remaining 108 points are colored Yellow. Prove that one can choose three points of each color in such a way that the four triangles formed by the chosen points of the same color are congruent.
3. Determine which integers $n > 1$ have the property that there exists an infinite sequence $a_1, a_2, a_3, \ldots$ of nonzero integers such that the equality

$$a_k + 2a_{2k} + \cdots + na_{nk} = 0$$

holds for every positive integer k.

Math. Mag. **85** (2012) 304–312. doi:10.4169/math.mag.85.4.304.

4. Find all functions $f : \mathbb{Z}^+ \to \mathbb{Z}^+$ (where $\mathbb{Z}^+$ is the set of positive integers) such that $f(n!) = f(n)!$ for all positive integers n and such that $m - n$ divides $f(m) - f(n)$ for all distinct positive integers m, n.

5. Let P be a point in the plane of $\triangle ABC$, and γ a line passing through P. Let A', B', C' be the points where the reflections of lines PA, PB, PC with respect to γ intersect lines BC, AC, AB, respectively. Prove that A', B', C' are collinear.

6. For integer $n \geq 2$, let $x_1, x_2, \ldots, x_n$ be real numbers satisfying

$$x_1 + x_2 + \cdots + x_n = 0, \qquad \text{and} \qquad x_1^2 + x_2^2 + \cdots + x_n^2 = 1.$$

For each subset $A \subseteq \{1, 2, \ldots, n\}$, define

$$S_A = \sum_{i \in A} x_i.$$

(If A is the empty set, then $S_A = 0$.)

Prove that for any positive number λ, the number of sets A satisfying $S_A \geq \lambda$ is at most $2^{n-3}/\lambda^2$. For what choices of $x_1, x_2, \ldots, x_n, \lambda$ does equality hold?

Solutions

1. An integer n has this property if and only if $n \geq 13$.

For any n, suppose that $a_1, a_2, \ldots, a_n$ satisfy $\max(a_1, a_2, \ldots, a_n) \leq n \cdot \min(a_1, a_2, \ldots, a_n)$, and that we cannot find three that are the side lengths of an acute triangle. We may assume that $a_1 \leq a_2 \leq \ldots \leq a_n$. Then $a_{i+2}^2 \geq a_i^2 + a_{i+1}^2$ for all $i \leq n-2$. Let $\{F_n\}$ be the Fibonacci sequence, with $F_1 = F_2 = 1$ and $F_{n+1} = F_n + F_{n-1}$. These numbers satisfy $F_n \leq n^2$ for $n \leq 12$ and $F_n > n^2$ for $n \geq 13$ (the last inequality being proved by induction). The inequality $a_{i+2}^2 \geq a_i^2 + a_{i+1}^2$ and the fact that $a_1 \leq a_2 \leq \cdots \leq a_n$ imply that $a_i^2 \geq F_i \cdot a_1^2$ for all $i \leq n$. Hence, if $n \geq 13$, we obtain $a_n^2 > n^2 \cdot a_1^2$, contradicting the hypothesis. This shows that any $n \geq 13$ has the required property.

If $n \leq 12$, take $a_i = \sqrt{F_i}$ for $1 \leq i \leq n$. Then $1 = a_1 \leq \cdots \leq a_n = \sqrt{F_n} \leq n$, so $\max(a_1, a_2, \ldots, a_n) \leq n \cdot \min(a_1, a_2, \ldots, a_n)$, but no three a_i's can be the side lengths of an acute triangle. Therefore no integer $n \leq 12$ has the property.

This problem and solution were suggested by Titu Andreescu.

2. We will use R, G, B, Y to denote the sets of Red, Green, Blue, Yellow points, respectively, and r, g, b, y to denote generic Red, Green, Blue, Yellow points, respectively. For any integer n with $0 \leq n \leq 431$, we let $\mathcal{T}_n$ denote the counterclockwise rotation of n steps—that is, $(360n/432)$ degrees around the center of the circle. Finally, for any set S, we let $|S|$ denote the number of elements in S.

First, we claim that there is some index i such that $|\mathcal{T}_i(R) \cap G| \geq 28$. Indeed, for every pair (r, g) consisting of one red point and one green point, there is a nonzero n for which $g = \mathcal{T}_n(r)$. There are 108^2 such pairs, and only 431 rotations $\mathcal{T}_n$ with $n \neq 0$, so by the pigeonhole principle at least one of these rotations must satisfy $g = \mathcal{T}_n(r)$ for at least $\lceil 108^2/431 \rceil = 28$ of the pairs (r, g). Call that rotation $\mathcal{T}_i$. Let RG denote the set of green points that appear in these pairs—that is, $RG = \mathcal{T}_i(R) \cap G$, and let rg denote a generic (green) point in RG.

Second, we claim that there is some index j such that $|\mathcal{T}_j(RG) \cap B| \geq 8$. Again, for every pair (rg, b) consisting of one green point in RG and one blue point, there is a nonzero n for which $b = \mathcal{T}_n(rg)$. There are at least $28 \cdot 108$ such pairs, and only 431 rotations $\mathcal{T}_k$ with $n \neq 0$, so at least one of these rotations must play that role for at least $\lceil 28 \cdot 108/431 \rceil = 8$ of the pairs. Call that rotation $\mathcal{T}_j$. Let RGB denote the set of green points that appear in these pairs—that is, $RGB = \mathcal{T}_j(RG) \cap B$, and let rgb denote a generic (blue) point in RGB.

Third, we claim that there is some index k such that $|\mathcal{T}_k(RGB) \cap Y| \geq 3$. Again, for every pair (rgb, y) consisting of one blue point in RGB and one yellow point, there is a nonzero n for which $y = \mathcal{T}_n(rgb)$. There are at least $8 \cdot 108$ such pairs, and only 431 rotations $\mathcal{T}_n$ with $n \neq 0$, so at least one of these rotations must play that role for at least $\lceil 8 \cdot 108/431 \rceil = 3$ of the pairs. Call that rotation $\mathcal{T}_k$.

Finally, let y_1, y_2, y_3 be three points in $\mathcal{T}_k(RGB) \cap Y$. Then

$$\begin{aligned}
&(y_1, y_2, y_3),\\
&(b_1, b_2, b_3) = \mathcal{T}_k^{-1}(y_1, y_2, y_3),\\
&(g_1, g_2, g_3) = \mathcal{T}_j^{-1}(b_1, b_2, b_3), \text{ and}\\
&(r_1, r_2, r_3) = \mathcal{T}_i^{-1}(g_1, g_2, g_3)
\end{aligned}$$

are twelve points that we are looking for.

This problem and solution were suggested by Gregory Galperin.

3. We will prove that the requested sequences exist for all $n \geq 3$. Bertrand's Postulate states that for every integer $n \geq 1$, there is a prime p such that $n < p \leq 2n$. It implies the following lemma.

LEMMA 1. *For every positive integer $n \geq 5$ there exist primes p and q such that $p < q \leq n$, $p \leq n < p^2$ and $q \leq n < 2q$.*

Proof. We treat the cases with $n \leq 13$ directly. When $5 \leq n \leq 8$ we may choose $p = 3$ and $q = 5$, and when $9 \leq n \leq 13$ we may choose $p = 5$ and $q = 7$.

For larger n we apply Bertrand's postulate to $\lfloor n/2 \rfloor$ to find a prime q satisfying $\lfloor n/2 \rfloor < q \leq 2\lfloor n/2 \rfloor$, and apply the postulate again to $\lfloor n/4 \rfloor$ to find a prime p satisfying $\lfloor n/4 \rfloor < p \leq 2\lfloor n/4 \rfloor$. These ranges never overlap, so $p < q$, and the prime q always satisfies $q \leq n < 2q$. When $n \geq 14$ (or even when $n \geq 12$) the prime p satisfies $p \leq n < p^2$, and this completes the proof of Lemma 1. ■

A function $f : \mathbb{Z}^+ \to \mathbb{Z}^*$ is a *totally multiplicative function* if $f(mn) = f(m)f(n)$ for all $m, n \in \mathbb{Z}^+$. (Here, $\mathbb{Z}^+$ and $\mathbb{Z}^*$ denote the sets of positive integers and non-zero integers respectively.) In this case $f(1) = 1$ and f is uniquely determined by its values at the primes.

LEMMA 2. *For every integer $n \geq 3$ there exists a totally multiplicative function $f : \mathbb{Z}^+ \to \mathbb{Z}^*$, such that:*

$$f(1) + 2f(2) + \cdots + nf(n) = 0. \tag{1}$$

Proof. If $n = 3$, define $f(n) = (-1)^\alpha$, where 3^α is the largest power of 3 that divides n. For example, $f(3) = f(6) = -1$ and $f(9) = +1$. Now

$$f(1) + 2f(2) + 3f(3) = 1 + 2(1) + 3(-1) = 0$$

as required.

If $n = 4$, define $f(n) = 2^\alpha(-7)^\beta$, where 2^α is the largest power of 2 that divides n and 3^β is the largest power of 3 that divides n. Now

$$f(1) + 2f(2) + 3f(3) + 4f(4) = 1 + 2(2) + 3(-7) + 4(4) = 0$$

as required.

For larger n, we apply Lemma 1 to find primes p and q such that $p < q \leq n$, $p \leq n < p^2$ and $q \leq n < 2q$. Write (1) as

$$W = \sum_{m \leq n} mf(m)$$
$$= \left(\sum_{\substack{m \leq n, \\ p \nmid m, q \nmid m}} mf(m) \right) + [pf(p) + 2pf(2p) \cdots tpf(tp)] + qf(q) \qquad (2)$$

with $t < p$. Define $f(m) = 1$ for all m not divisible by p or q. Then (2) reduces to

$$W = \left(\sum_{\substack{m \leq n, \\ p \nmid m, q \nmid m}} mf(m) \right) + \frac{t(t+1)}{2} pf(p) + qf(q).$$

Since $\gcd(\frac{t(t+1)}{2} p, q) = 1$, Bezout's theorem guarantees that we can choose $f(p)$ and $f(q)$ in such a way that $W = 0$, thus proving Lemma 2. ■

Now for all $n \geq 3$, let $a_n = f(n)$. Then $a_n \neq 0$ for all n, and for any positive integer k

$$a_k + 2a_{2k} + \cdots + na_{nk} = f(k) + 2f(2k) + \cdot + nf(nk)$$
$$= f(k)[f(1) + 2f(2) + \cdots + nf(n)] = 0$$

as desired.

If $n = 2$, the condition $a_k + 2a_{2k} = 0$ implies that $a_1 = -2a_2 = 4a_4 = \cdots$ indicating that a_1 is divisible by arbitrary large powers of 2. Therefore, no such sequence is possible when $n = 2$.

This problem was suggested by Gabriel Carroll. The solution was suggested by student Xiaoyu He.

4. There are three solutions: the identity function and the constant functions 1 and 2. Let us prove that these are the only ones. Consider such a function f and suppose, first, that there exists $a > 2$ such that $f(a) = a$. Then $a!$, $(a!)!$, ... are all fixed points of f. So there is an increasing sequence $(a_k)_{k \geq 0}$ of fixed points. If n is any positive integer, $a_k - n$ divides $a_k - f(n) = f(a_k) - f(n)$ for all k, and so it also divides $f(n) - n$ for all k. Thus $f(n) = n$ and since this holds for any n, f must be the identity function.

   Now suppose that f has no fixed points greater than 2. From the hypothesis that $f(1)! = f(1!)$ we know that $f(1)$ must be 1 or 2. Let $p > 3$ be a prime and observe that $(p-2)! \equiv 1 \pmod{p}$ (by Wilson's theorem), thus $f(p-2)! - f(1) = f((p-2)!) - f(1)$ is a multiple of p. As $p > 3$, the fact that p divides $f(p-2)! - f(1)$ implies that $f(p-2) < p$. Since $(p-1)! - f(1)$ is not a multiple of p (again by Wilson), we deduce that actually $f(p-2) \leq p-2$. On the other hand, $p-3$ divides $f(p-2) - f(1) \leq f(p-2) - 1$. Thus either $f(p-2) = f(1)$ or $f(p-2) = p-2$. As $p-2 > 2$, the last case is excluded and so $f(p-2) = f(1)$. This holds for all primes $p > 3$. Taking n any positive integer, we deduce that $p-2-n$ divides $f(1) - f(n)$ for all large primes p. Thus $f(n) = f(1)$ and f is constant. We have seen that the constant must be $f = 1$ or $f = 2$.

   This problem and solution were suggested by Titu Andreescu and Gabriel Dospinescu.

5. It is easy to see (say by the Law of Sines) that

$$\frac{AC'}{BC'} = \frac{AP \sin \angle APC'}{BP \sin \angle BPC'}, \quad \frac{BA'}{CA'} = \frac{BP \sin \angle BPA'}{CP \sin \angle CPA'}, \quad \frac{CB'}{AB'} = \frac{CP \sin \angle CPB'}{AP \sin \angle APB'}.$$

The construction of A', B', C' by reflections implies that

$$\sin \angle APC' = \sin \angle CPA', \ \sin \angle BPC' = \sin \angle CPB', \ \sin \angle BPC' = \sin \angle CPB'.$$

Hence,

$$\frac{AC'}{BC'} \cdot \frac{BA'}{CA'} \cdot \frac{CB'}{AB'} = 1,$$

and the proof is complete by Menelaus' theorem.

This problem was suggested by Titu Andreescu and Cosmin Pohoata. The solution was suggested by Li Zhou, Polk State College, Winter Haven FL.

6. This problem is a form of Chebyshev's inequality for random variables. For each set $A \subseteq \{1, 2, \ldots, n\}$, define

$$\Delta_A = 2S_A = \sum_{i \in A} x_i - \sum_{i \in \{1,2,\ldots,n\} \setminus A} x_i = \sum_{i=1}^{n} \epsilon_A(i) x_i,$$

where $\epsilon_A(i) = 1$ if $i \in A$ and -1 otherwise. Squaring, we have

$$\Delta_A^2 = \sum_{i=1}^{n} x_i^2 + \sum_{\substack{i,j \in \{1,\ldots,n\} \\ i \neq j}} \epsilon_A(i) \epsilon_A(j) x_i x_j. \tag{3}$$

Now sum the Δ_A^2's over all 2^n possible choices of A. For each pair $i \neq j$, there are 2^{n-2} sets A with $i, j \in A$, and another 2^{n-2} sets with $i, j \notin A$; each of these sets contributes a term of $+x_i x_j$ to the sum in (3). There are also 2^{n-2} sets A with $i \in A$, $j \notin A$, and 2^{n-2} sets with $i \notin A$, $j \in A$. Each of these sets contributes a term of $-x_i x_j$ to (3). Hence, $x_i x_j$ appears 2^{n-1} times with a $+$ sign and 2^{n-1} times with a $-$ sign. Therefore all of these terms cancel, and we find

$$\sum_{A \subseteq \{1,2,\ldots,n\}} \Delta_A^2 = 2^n (x_1^2 + \cdots + x_n^2) = 2^n. \tag{4}$$

Now let $\lambda > 0$. From (4), there cannot be more than $2^{n-2}/\lambda^2$ terms Δ_A^2 with value greater than or equal to $4\lambda^2$. Hence, there can be at most $2^{n-2}/\lambda^2$ sets A such that $|S_A| \geq \lambda$. (Recall that $\Delta_A = 2S_A$). Moreover, these sets can be arranged into complementary pairs because $S_A = -S_{\{1,\ldots,n\} \setminus A}$. In each of these pairs, exactly one of the two members is positive. Therefore there are at most $2^{n-3}/\lambda^2$ sets A with $S_A \geq \lambda$.

For equality to hold, it must be the case that all positive values of Δ_A^2 are equal to $4\lambda^2$; otherwise we would again have a contradiction because the sum of all Δ_A^2 would exceed 2^n. In particular, all positive values of Δ_A^2 must be the same. Thus all positive values of x_A must be the same. This will be the case only if at most one of the x_i is positive and at most one of the x_i is negative. Because we must have at least one of each, there must be exactly one positive term and one negative term. Thus it must be the case that one $x_k = \sqrt{2}/2$ for some k, one is $x_j = -\sqrt{2}/2$ for some $j \neq k$, and all other $x_i = 0$. Then the assumption that every positive $\Delta_A^2 = 4\lambda^2$ yields $\lambda = \sqrt{2}/2$.

Conversely, with the x_i and λ as described, we have exactly $2^{n-2} = 2^{n-3}/\lambda^2$ sets A such that $x_A \geq \lambda$ (namely, those sets A that contain the $\sqrt{2}/2$ term and do not contain the $-\sqrt{2}/2$ term.) So these are exactly the cases in which equality holds.

This problem and solution were suggested by Gabriel Carroll.

USAJMO Problems

1. Given a triangle ABC, let P and Q be points on segments $\overline{AB}$ and $\overline{AC}$, respectively, such that $AP = AQ$. Let S and R be distinct points on segment $\overline{BC}$ such that S lies between B and R, $\angle BPS = \angle PRS$, and $\angle CQR = \angle QSR$. Prove that P, Q, R, S are concyclic (in other words, these four points lie on a circle).
2. Same as USAMO 1.
3. Let a, b, c be positive real numbers. Prove that
$$\frac{a^3 + 3b^3}{5a + b} + \frac{b^3 + 3c^3}{5b + c} + \frac{c^3 + 3a^3}{5c + a} \geq \frac{2}{3}(a^2 + b^2 + c^2).$$
4. Let α be an irrational number with $0 < \alpha < 1$, and draw a circle in the plane whose circumference has length 1. Given any integer $n \geq 3$, define a sequence of points $P_1, P_2, \ldots, P_n$ as follows. First select any point P_1 on the circle, and for $2 \leq k \leq n$ define P_k as the point on the circle for which the length of arc $P_{k-1}P_k$ is α, when traveling counter-clockwise around the circle from P_{k-1} to P_k. Suppose that P_a and P_b are the nearest adjacent points on either side of P_n. Prove that $a + b \leq n$.
5. For distinct positive integers $a, b < 2012$, define $f(a, b)$ to be the number of integers k with $1 \leq k < 2012$ such that the remainder when ak divided by 2012 is greater than that of bk divided by 2012. Let S be the minimum value of $f(a, b)$, where a and b range over all pairs of distinct positive integers less than 2012. Determine S.
6. Same as USAMO 5.

Solutions

1. We use the following lemma.

LEMMA. *Given a triangle ABC, X, Y, Z are points on BC, CA, AB respectively. Then three perpendicular lines of BC, CA, AB which go through X, Y, Z respectively are concurrent if and only if $AY^2 + BZ^2 + CX^2 = AZ^2 + BX^2 + CY^2$.*

Proof. If the lines are concurrent, let P be the point on the three lines. From $BX^2 - CX^2 = (PB^2 - PX^2) - (PC^2 - PX^2) = PB^2 - PC^2$ and so on, we obtain the desired result. Conversely, if $AY^2 + BZ^2 + CX^2 = AZ^2 + BX^2 + CY^2$ holds, let Q be the intersection the lines perpendicular to BC and CA which go through X and Y respectively. Then as we have seen $BX^2 - CX^2 = QB^2 - QC^2$ and $CY^2 - AY^2 = QC^2 - QA^2$ holds. Summing these equations, we have $AZ^2 - BZ^2 = QA^2 - QB^2$. This implies that QZ and AB are perpendicular, as desired. ■

Let M be the midpoint of SR. We show that $AP^2 + BM^2 + CQ^2 = AQ^2 + BP^2 + CM^2$. Since $AP = AQ$, $CQ^2 = CR \cdot CS$, $BP^2 = BS \cdot BR$, and $BM^2 - CM^2 = (BM + CM)(BM - CM) = BC(BS - RC)$, we have $(AP^2 + BM^2 + CQ^2) - (AQ^2 + BP^2 + CM^2) = BC(BS - RC) - BS \cdot BR + CR \cdot CS = BS \cdot CR - CR \cdot BC = 0$. Thus there exists a point O such that $OP \perp BC$, $OQ \perp AC$, $OM \perp BC$. Then O is the center of a circumcircle of PRS, since the circle is tangent to AB at P.

Similarly, O is the center of a circumcircle of QRS, which implies that P, Q, R, S are on a circle.

This problem and solution were suggested by Sungyoon Kim and Inseok Seo.

3. We start by noting that $(a^2+b^2+c^2) \geq (ab+bc+ca)$, with equality only when $a=b=c$. This follows from the identity

$$(a^2+b^2+c^2)-(ab+bc+ca) = \frac{1}{2}\left[(a-b)^2+(b-c)^2+(c-a)^2\right].$$

Now in the Cauchy-Schwarz inequality, $(u_1^2+u_2^2+u_3^2)(v_1^2+v_2^2+v_3^2) \geq (u_1v_1+u_2v_2+u_3v_3)^2$, we can substitute $u_i = x_i/\sqrt{y_i}$ and $v_i = \sqrt{y_i}$ to obtain the equivalent form

$$\frac{x_1^2}{y_1}+\frac{x_2^2}{y_2}+\frac{x_3^2}{y_3} \geq \frac{(x_1+x_2+x_3)^2}{y_1+y_2+y_3}.$$

In the following calculation we apply the last result with the substitutions $x_1 = a^2$ and $y_1 = (5a^2+ab)$, etc.

$$\begin{aligned}\frac{a^3}{5a+b}+\frac{b^3}{5b+c}+\frac{c^3}{5c+a} &= \frac{a^4}{5a^2+ab}+\frac{b^4}{5b^2+bc}+\frac{c^4}{5c^2+ca}\\ &\geq \frac{(a^2+b^2+c^2)^2}{5(a^2+b^2+c^2)+(ab+bc+ca)}\\ &\geq \frac{(a^2+b^2+c^2)^2}{5(a^2+b^2+c^2)+(a^2+b^2+c^2)}\\ &= \frac{1}{6}(a^2+b^2+c^2).\end{aligned}$$

An almost identical calculation shows that

$$\frac{3b^3}{5a+b}+\frac{3c^3}{5b+c}+\frac{3a^3}{5c+a} \geq \frac{3}{6}(a^2+b^2+c^2),$$

and combining these results gives the inequality in the problem.

Equality holds only if $(a^2+b^2+c^2) = (ab+bc+ca)$, which requires that $a=b=c$.

This problem and solution were suggested by Titu Andreescu.

4. Observe that since α is irrational no two of the points will coincide. It will be useful to define the auxiliary point P_0 such that the length of arc P_0P_1 is α, when traveling counter-clockwise around the circle from P_0 to P_1. We begin by noting that for any $n \geq 3$, if $a+b=n$ then P_0 lies on the arc from P_a to P_b containing P_n. For if we travel back (clockwise) around the circle through a distance of $b\alpha$ from P_n then we reach P_a. The same translation must map P_b to P_0, and since P_n is situated between P_a and P_b, we deduce that P_0 must be also.

The claim is clearly true for $n=3$. Now suppose to the contrary that for some value of n we have $a+b>n$ and consider the minimal such counterexample. If in fact $a+b>n+1$, then we may translate the three points P_a, P_b, and P_n clockwise around the circle through a distance α to find points P_{a-1} and P_{b-1} adjacent to P_{n-1} on either side. But then we would have $(a-1)+(b-1)>(n-1)$ for this trio of points, which contradicts our assumption that n was the minimal counterexample.

Therefore we must have $a + b = n + 1$. Again we translate points P_a, P_b, and P_n clockwise around the circle through a distance α to obtain points P_{a-1} and P_{b-1} adjacent to P_{n-1} on either side with $(a-1)+(b-1)=(n-1)$. By our earlier observation this implies that P_0 lies on the arc from P_{a-1} to P_{b-1} containing P_{n-1}. But now translating forward again, we conclude that P_1 lies on the arc from P_a to P_b containing P_n, contradicting the fact that P_a and P_b were the nearest adjacent points to P_n on either side. This completes the proof.

This problem and solution were suggested by Sam Vandervelde.

5. For simplicity, we will define $g(n)$ to be $n \pmod{2012}$. Note that $g(ak) + g(a(2012-k))$ is either 0 or 2012; it is 0 exactly when 2012 divides ak. This means that for $1 \le k \le 1005$, the number of elements i in $\{k, 2012-k\}$ such that $ai \pmod{2012} > bi \pmod{2012}$ is

$$\begin{cases} 0 & \text{if } g(ak)=0 \text{ or } g(ak)=g(bk); \\ 2 & \text{if } g(bk)=0 \text{ and } g(ak)\neq 0; \\ 1 & \text{otherwise.} \end{cases}$$

Let $T = \{1, 2, \ldots, 1005\}$. Note that the condition $g(ak) = g(bk)$ is equivalent to $g((a-b)k) = 0$. We will try to choose a, b so as to maximize the number of numbers k in T such that the first of the three cases occurs. From the prime factorization $2012 = 2 \cdot 2 \cdot 503$, the proper divisors of 2012 are 1, 2, 4, 503, and 1006. We shall choose a and $a - b$ to be multiples of some of these numbers. It is not hard to verify that we can choose a to be a multiple of 1006 and $a - b$ to be a multiple of 4. We will take $a = 1006$ and $b = 1002$.

With this choice of a and b, the second of the three cases (i.e., $g(bk) = 0$ and $g(ak) \neq 0$) never occurs, hence minimizing the number of elements i in $T - \{1006\}$ such that $ai \pmod{2012} > bi \pmod{2012}$. Moreover, $g(1006a) = 0$, meaning that $g(1006a) > g(1006b)$ does not hold. This means that our choice of a and b minimizes $f(a, b)$.

Note that $g(1006k) = 0$ occurs for 502 values in T, and $g(1006k) = g(1002k)$ occurs for 1 value in T. No value in T satisfies both condition. Hence $S = 1005 - 502 - 1 = 502$.

This problem and solution were suggested by Warut Suksompong.

The top twelve students on the 2012 USAMO were (in alphabetical order):

Andre Arlsan	11	Hunter College High School	New York, NY
Joshua Brakensiek	10	homeschool	Chandler, AZ
Calvin Deng	11	North Carolina School of Science and Mathematics	Durham, NC
Xiaoyu He	12	Acton-Boxborough Regional High School	Acton, MA
Ravi Jagadeesan	10	Phillips Exeter Academy	Exeter, NH
Mitchell Lee	12	homeschool	Oakton, VA
Zhuo Qun (Alex) Song	9	Phillips Exeter Academy	Exeter, NH
Thomas Swayze	11	Canyon Crest Academy	San Diego, CA
Victor Wang	11	Ladue Horton Watkins High School	St. Louis, MO
David Yang	11	Phillips Exeter Academy	Exeter, NH
Samuel Zbarsky	11	Montgomery Blair High School	Silver Spring, MD
Alex Zhu	12	Academy for the Advancement of Science and Technology	Hackensack, NJ

The top fourteen students on the 2012 USAJMO were (in alphabetical order):

Ernest Chiu	10	West Windsor Plainsboro High School	Plainsboro, NJ
Paolo Gentili	10	Canyon Crest Academy	San Diego, CA
Courtney Guo	10	International School of Beijing	Beijing, China
Steven Hao	10	Lynbrook High School	San Jose, CA
Andrew He	9	Monta Vista High School	Cupertino, CA
Calvin Huang	10	Henry M Gunn High School	Palo Alto, CA
Shashwat Kishore	9	Unionville High School	Kennett Square, PA
Laura Pierson	6	Berkeley Math Circle, University of California	Berkeley, CA
Tahsin Saffat	9	Westview High School	Portland, OR
David Stoner	9	South Aiken High School	Aiken, SC
Ashwath Thirumalai	9	Harker High School	San Jose, CA
Jerry Wu	10	Mission San Jose High School	Fremont, CA
Isaac Xia	10	Concord-Carlisle Regional High School	Concord, MA
Jesse Zhang	9	University of Colorado	Boulder, CO

53rd International Mathematical Olympiad

ZUMING FENG
Phillips Exeter Academy
Exeter, NH 03833
zfeng@exeter.edu

YI SUN
Massachusetts Institute of Technology
Cambridge, MA 02139
yisun@mit.edu

Problems (Day 1)

1. Given triangle ABC the point J is the centre of the excircle opposite the vertex A. This excircle is tangent to the side BC at M, and to the lines AB and AC at K and L, respectively. The lines LM and BJ meet at F, and the lines KM and CJ meet at G. Let S be the point of intersection of the lines AF and BC, and let T be the point of intersection of the lines AG and BC.

   Prove that M is the midpoint of ST.

   (The *excircle* of ABC opposite the vertex A is the circle tangent to the line segment BC, to the ray AB beyond B, and to the ray AC beyond C.)

2. Let $n \geq 3$ be an integer, and let $a_2, a_3, \ldots, a_n$ be positive real numbers such that $a_2 a_3 \cdots a_n = 1$. Prove that

$$(1 + a_2)^2 (1 + a_3)^3 \cdots (1 + a_n)^n > n^n.$$

Math. Mag. **85** (2012) 312–317. doi:10.4169/math.mag.85.4.312.

3. The *liar's guessing game* is a game played between two players A and B. The rules of the game depend on two positive integers k and n which are known to both players.

   At the start of the game A chooses integers x and N with $1 \le x \le N$. Player A keeps x secret, and truthfully tells N to player B. Player B now tries to obtain information about x by asking player A questions as follows: each question consists of B specifying an arbitrary set S of positive integers (possibly one specified in some previous question), and asking A whether x belongs to S. Player B may ask as many such questions as he wishes. After each question, player A must immediately answer it with *yes* or *no*, but is allowed to lie as many times as she wants; the only restriction is that, among any $k+1$ consecutive answers, at least one answer must be truthful.

   After B has asked as many questions as he wants, he must specify a set X of at most n positive integers. If x belongs to X, then B wins; otherwise, he loses. Prove that:

   1. If $n \ge 2^k$, then B can guarantee a win.
   2. For all sufficiently large k, there exists an integer $n \ge 1.99^k$ such that B cannot guarantee a win.

Problems (Day 2)

4. Find all functions $f : \mathbb{Z} \to \mathbb{Z}$ such that, for all integers a, b, c that satisfy $a+b+c=0$, the following equality holds:

$$f(a)^2 + f(b)^2 + f(c)^2 = 2f(a)f(b) + 2f(b)f(c) + 2f(c)f(a).$$

   (Here $\mathbb{Z}$ denotes the set of integers.)

5. Let ABC be a triangle with $\angle BCA = 90°$, and let D be the foot of the altitude from C. Let X be a point in the interior of segment CD. Let K be the point on the segment AX such that $BK = BC$. Similarly, let L be the point on the segment BX such that $AL = AC$. Let M be the point of intersection of AL and BK.

   Show that $MK = ML$.

6. Find all positive integers n for which there exist non-negative integers $a_1, a_2, \ldots, a_n$ such that

$$\frac{1}{2^{a_1}} + \frac{1}{2^{a_2}} + \cdots + \frac{1}{2^{a_n}} = \frac{1}{3^{a_1}} + \frac{2}{3^{a_2}} + \cdots + \frac{n}{3^{a_n}} = 1.$$

Solutions

1. Notice that $\angle KAJ = \frac{A}{2}$ and that

$$\angle KGJ = \angle MCJ - \angle GMC = \angle MCJ - \angle KMB = \frac{A+B}{2} - \frac{B}{2} = \frac{A}{2},$$

   so $\angle KAJ = \angle KGJ$, hence $AKJG$ is cyclic. In particular, $\angle AGC = \angle AKJ = 90°$, meaning that $AG \perp GJ$, so $AG \parallel ML$. Now, CML is an isosceles triangle with altitude CJ, so because $AT \parallel ML$, ACT is isosceles with altitude LG. In the same way, we can show that ABS is isosceles with altitude BF.

   Notice now that $\angle SAT = A + \angle SAB + \angle CAT = A + \frac{B}{2} + \frac{C}{2}$, where we have used the fact that SAB and CAT are isosceles. On the other hand, we see that $\angle MGT = 90 + \angle KGJ = 90 + \angle KAJ = 90 + \frac{A}{2}$, which implies that $\angle MGT = \angle SAT$, hence $SA \parallel MG$. Because G is the midpoint of AT, this implies that MG is the midline of triangle AST, so M is the midpoint of ST.

   This problem was proposed by Evangelos Psychas of Greece.

2. By the AM-GM inequality, for every k with $2 \le k \le n$, we have

$$(1+a_k)^k = \left(\frac{1}{k-1} + \frac{1}{k-1} + \cdots + \frac{1}{k-1} + a_k\right)^k \ge \frac{k^k a_k}{(k-1)^{k-1}},$$

where equality holds if and only if $a_k = \frac{1}{k-1}$. Multiplying these inequalities for each k from 2 to n yields

$$\begin{aligned}(1+a_2)^2(1+a_3)^3 \cdots (1+a_n)^n &\ge 2^2 a_2 \cdot \frac{3^3 a_3}{2^2} \cdots \frac{n^n a_n}{(n-1)^{n-1}} \\ &= n^n a_2 a_3 \cdots a_n = n^n.\end{aligned}$$

Equality holds only if $a_k = \frac{1}{k-1}$ for each k, implying that $a_2 a_3 \cdots a_n = \frac{1}{(n-1)!}$, an impossibility for $n \ge 3$. Thus, the inequality is strict, as needed.

This problem was proposed by Angelo di Pasquale of Australia.

3. (1) Let T be the set of possible values of x given the answers to B's questions. We give a strategy for B to reduce $|T|$ to at most 2^k, upon which he can specify the set T to guarantee a win.

Suppose $|T| > 2^k$, and let $t_0, t_1, \ldots, t_{2^k}$ be $2^k + 1$ distinct elements of T. Let B start by asking repeatedly about the set $\{t_{2^k}\}$. If A says *no* to the first $k+1$ of these questions, then t_{2^k} is excluded, reducing the size of T by 1.

If A says *yes* to any question, B stops asking about $\{t_{2^k}\}$ and asks about the sets $U_0, U_1, \ldots, U_{k-1}$, where

$$U_i = \{t_j \mid j \text{ has a 0 in the } i\text{th digit in binary}\}.$$

Construct the binary number $d = \overline{d_{k-1}d_{k-2}\cdots d_0}$ by $d_i = 0$ if A said *no* to U_i and $d_i = 1$ otherwise. If x were equal to t_d, then A would have lied in her answers to the previous $k+1$ questions. So $x \ne t_d$. In this case also, B has reduced $|T|$ by one. He can repeat until $|T| = 2^k$ and then specify the set T to win.

(2) Let λ be a real number with $1.99 < \lambda < 2$. Because $\lambda > 1.99$, for sufficiently large k, we have $1.99^k + 2 < (2-\lambda)\lambda^{k+1}$. Choose n to be an integer so that $1.99^k \le n < 1.99^k + 1$, meaning that $n + 1 < (2-\lambda)\lambda^{k+1}$.

Player A will choose $N = n+1$ and an arbitrary x. Let $m_i(t)$ be the number of consecutive answers, ending at the tth answer, which are inconsistent with $x = i$ for $i = 1, 2, \ldots, n+1$, with $m_i(0) = 0$. Define

$$L(t) = \sum_{i=1}^{n+1} \lambda^{m_i(t)}.$$

Player A will use the strategy of giving the answer which minimizes $L(t)$, irrespective of her choice of x. We will show that this is a valid strategy and that B can guarantee that any number was not chosen by A.

We first show by induction that $L(t) < \lambda^{k+1}$. For the base case, we have $L(0) = n + 1 < (2-\lambda)\lambda^{k+1}$. Suppose now that $L(t) < \lambda^{k+1}$; if B asks about a set S, A chooses between two possible values for $L(t+1)$:

$$L_1 = |S| + \sum_{i \notin S} \lambda^{m_i(t)+1}, \qquad L_2 = (n+1-|S|) + \sum_{i \in S} \lambda^{m_i(t)+1}.$$

Recalling that $n+1 < (2-\lambda)\lambda^{k+1}$ and $\lambda L(t) < \lambda^{k+2}$, we have

$$\frac{L_1+L_2}{2} = \frac{n+1}{2} + \frac{1}{2}\sum_{i=1}^{n+1} \lambda^{m_i(t)+1} = \frac{n+1+\lambda L(t)}{2} < \lambda^{k+1},$$

so we see that $L(t+1) = \min\{L_1, L_2\} < \lambda^{k+1}$, completing the induction.

It is therefore impossible for any $m_i(t)$ to reach a value $k+1$ or higher, so $m_i(t) \le k$ for all i and t. This means that A's strategy never violates the rules. Because A's answers are independent of x, there is no number that B can guarantee that A chose, so B cannot guarantee a win.

This problem was proposed by David Arthur and Jacob Tsimerman of Canada.

4. There are three classes of solutions, namely:

- for a fixed integer m, the function $f(n) = mn^2$,
- for a fixed non-zero integer m, the function

$$f(n) = \begin{cases} 0 & \text{if } n \text{ is even,} \\ m & \text{if } n \text{ is odd,} \end{cases} \quad \text{and}$$

- for a fixed non-zero integer m, the function

$$f(n) = \begin{cases} 0 & \text{if } n \equiv 0 \pmod 4, \\ 4m & \text{if } n \equiv 2 \pmod 4, \\ m & \text{if } n \text{ is odd.} \end{cases}$$

All of these functions satisfy the given identity. It remains to show they are the only solutions. Setting $a = b = c = 0$ in the identity yields $3f(0)^2 = 6f(0)^2$, hence $f(0) = 0$. Setting $a = 0$ and $c = -b$ in the identity yields $f(b)^2 + f(-b)^2 = 2f(b)f(-b)$, hence $(f(b) - f(-b))^2 = 0$, so $f(b) = f(-b)$ for every $b \in \mathbb{Z}$.

Setting $c = -a-b$ in the given and using the fact that $f(-a-b) = f(a+b)$, we obtain

$$f(a)^2 + f(b)^2 + f(a+b)^2 = 2f(a)f(b) + 2f(b)f(a+b) + 2f(a+b)f(a).$$

Rearranging and factoring yields

$$(f(a+b) - f(a) - f(b))^2 = 4f(a)f(b). \tag{1}$$

Now, if $f(a) = 0$ for any a, then from (1) (or from the original identity) we get $f(a+b) = f(b)$ for every $b \in \mathbb{Z}$, so f has period a. If $f(1) = 0$ this means that f is identically zero, and we have a solution in the first class. If $f(2) = 1$ the periodicity means that f is in the second class. Otherwise we may assume that $f(1) = m \neq 0$ and $f(2) \neq 0$.

In this case applying (1) with $a = b = 1$ implies that $f(2) = 4m$, and then applying (1) with $a = 2$ and $b = 1$ implies that $f(3) = m$ or $f(3) = 9m$. If $f(3) = m$, then applying (1) with $a = 3$ and $b = 1$ implies that $f(4) \in \{0, 4m\}$. On the other hand, applying (1) with $a = b = 2$ shows that $f(4) \in \{0, 16m\}$. Because m is non-zero, this shows that $f(4) = 0$. Now the periodicity of f places it the third class.

In the only remaining case, we have $f(0) = 0$, $f(1) = m$, $f(2) = 4m$, and $f(3) = 9m$. We claim by induction on n that $f(n) = mn^2$ for all positive integers n. The base cases $n = 0, 1, 2, 3$ hold by assumption. Suppose that $f(\ell) = m\ell^2$ for all $\ell \le n$ and that $n \ge 3$. Now equation (1) with $a = n$ and $b = 1$ implies that $f(n+1) \in \{(n+1)^2m, (n-1)^2m\}$, while equation (1) with $a = n-1$ and $b = 2$ implies that $f(n+1) \in \{(n+1)^2m, (n-3)^2m\}$. Because $n \geq 3$, $(n-3)^2m \neq (n-1)^2m$, so we must have $f(n+1) = (n+1)^2m$, completing the induction and placing f in the first class of solutions.

This problem was proposed by Liam Baker of South Africa.

5. Extend segment AX through X to a point E satisfying $BE \perp AX$. Extend segment BX through X to a point F satisfying $AF \perp BX$. Let lines BE and AF meet at Y so that Y is the orthocenter of triangle ABX.

It is easy to see that $BDFY$ is cyclic, and so $AD \cdot AB = AF \cdot AY$ by power of a point. Consequently, using the fact that triangles ACD and ABC are similar, we have $AL^2 = AC^2 = AD \cdot AB = AF \cdot AY$. It follows that triangles AFL and ALY are similar, implying that $\angle ALY = \angle AFL = 90°$ and $YL^2 = YF \cdot YA$. In exactly the same way, we can show that $\angle BKY = 90°$ and $YK^2 = YE \cdot YB$.

Because X is the orthocenter of triangle ABY, $ABEF$ is cyclic, from which it follows that $YF \cdot YA = YE \cdot YB$ by power of a point. Combining this with our previous observations, we find that $YL^2 = YF \cdot YA = YE \cdot YB = YK^2$. Now, notice that YKM and YLM are right triangles which share side YM and have $YK = YL$, so they are congruent, implying that $MK = ML$.

This problem was proposed by Josef Tkadlec of the Czech Republic.

6. There exist such numbers $a_1, a_2, \ldots, a_n$ if and only if $n \equiv 1, 2 \pmod 4$.

We first show this is necessary. For $a_1, a_2, \ldots, a_n$ satisfying the conditions, let $M = \max\{a_1, a_2, \ldots, a_n\}$ and let $N = \sum_{i=1}^{n} i \cdot 3^{M-a_i}$ so that $\frac{1}{3^{a_1}} + \frac{2}{3^{a_2}} + \cdots + \frac{n}{3^{a_n}} = \frac{N}{3^M}$, which means that $N = 3^M$ is odd. Thus, we have

$$1 + 2 + \cdots + n \equiv \sum_{i=1}^{n} i \cdot 3^{M-a_i} = N \equiv 1 \pmod 2,$$

which can happen only if $n \equiv 1, 2 \pmod 4$.

We now show that $n \equiv 1, 2 \pmod 4$ is sufficient. Call a sequence $b_1, b_2, \ldots, b_n$ *feasible* if there exist nonnegative integers $a_1, a_2, \ldots, a_n$ such that

$$\frac{1}{2^{a_1}} + \frac{1}{2^{a_2}} + \cdots + \frac{1}{2^{a_n}} = \frac{b_1}{3^{a_1}} + \frac{b_2}{3^{a_2}} + \cdots + \frac{b_n}{3^{a_n}} = 1.$$

We wish to show the sequence $1, 2, \ldots, n$, which we denote by α_n, is feasible for $n \equiv 1, 2 \pmod 4$. We first give a method of generating feasible sequences.

LEMMA. *Let $b_1, b_2, \ldots, b_n$ be feasible. For non-negative integers u and v with $u + v = 3b_k$, the sequence $b_1, b_2, \ldots, b_{k-1}, u, v, b_{k+1}, \ldots, b_n$ is feasible.*

Proof. Let $a_1, \ldots, a_n$ be exponents making $b_1, \ldots, b_n$ feasible, meaning that

$$\left(\frac{1}{2^{a_1}} + \cdots + \frac{1}{2^{a_{k-1}}}\right) + \frac{1}{2^{a_k}} + \left(\frac{1}{2^{a_{k+1}}} + \cdots + \frac{1}{2^{a_n}}\right) = 1$$

$$\left(\frac{b_1}{3^{a_1}} + \cdots + \frac{b_{k-1}}{3^{a_{k-1}}}\right) + \frac{b_k}{3^{a_k}} + \left(\frac{b_{k+1}}{3^{a_{k+1}}} + \cdots + \frac{b_n}{3^{a_n}}\right) = 1.$$

Observe that

$$\frac{1}{2^{a_k+1}} + \frac{1}{2^{a_k+1}} = \frac{1}{2^{a_k}} \qquad \text{and} \qquad \frac{u}{3^{a_k+1}} + \frac{v}{3^{a_k+1}} = \frac{b_k}{3^{a_k}},$$

so the previous relations show that

$$\left(\frac{1}{2^{a_1}}+\cdots+\frac{1}{2^{a_{k-1}}}\right)+\frac{1}{2^{a_k+1}}+\frac{1}{2^{a_k+1}}+\left(\frac{1}{2^{a_{k+1}}}+\cdots+\frac{1}{2^{a_n}}\right)=1$$

$$\left(\frac{b_1}{3^{a_1}}+\cdots+\frac{b_{k-1}}{3^{a_{k-1}}}\right)+\frac{u}{3^{a_k+1}}+\frac{v}{3^{a_k+1}}+\left(\frac{b_{k+1}}{3^{a_{k+1}}}+\cdots+\frac{b_n}{3^{a_n}}\right)=1,$$

so the exponents $a_1, \ldots, a_{k-1}, a_k+1, a_k+1, a_{k+1}, \ldots, a_n$ show feasibility. ■

We induct on n to show that α_n is feasible when $n \equiv 1, 2 \pmod 4$. First, for the base cases $n = 1, 2, 5, 6$, we may take the sequences of exponents to be (0), $(1, 1)$, $(2, 2, 2, 3, 3)$, and $(2, 2, 3, 3, 3, 3)$, respectively. Now, suppose that for some $n \geq 9$, α_k is feasible for all $k < n$. If $n \equiv 2 \pmod 4$, then $n-1 \equiv 1 \pmod 4$, so α_{n-1} is feasible by the inductive hypothesis. Applying the lemma to α_{n-1} with $u = \frac{n}{2}$ and $v = n$, we see that α_n is feasible. If $n \equiv 1 \pmod 4$, then $n-7 \equiv 2 \pmod 4$ with $n-7 > 0$ and $\frac{n-5}{2} \geq n-7$. By the inductive hypothesis, α_{n-7} is feasible. Now, beginning with α_{n-7}, we apply the lemma seven times with $(b_k, u, v) = \left(\frac{n-5}{2}, \frac{n-3}{2}, n-6\right)$, $\left(\frac{n-3}{2}, \frac{n-3}{2}, n-3\right)$, $\left(\frac{n-3}{2}, \frac{n-1}{2}, n-4\right)$, $\left(\frac{n-1}{2}, \frac{n-1}{2}, n-1\right)$, $\left(\frac{n-1}{2}, \frac{n-3}{2}, n\right)$, $\left(\frac{n-3}{2}, \frac{n-5}{2}, n-2\right)$, and $\left(\frac{n-5}{2}, \frac{n-5}{2}, n-5\right)$ to obtain that α_n is feasible, completing the induction.

This problem was proposed by Dušan Djukić of Serbia.

Results

The IMO was held in Mar del Plata, Argentina, on July 10–11, 2011. There were 548 competitors from 100 countries and regions. On each day contestants were given four and a half hours for three problems.

On this challenging exam, a perfect score was achieved by only one student, Jeck Lim (Singapore). The USA team won 5 gold medals and 1 silver medal, placing third behind Korea and China. The students' individual results were as follows.

- Xiaoyu He, who finished 12th grade at Acton-Boxborough Regional High School in Acton, MA, won a silver medal.
- Ravi Jagadeesan, who finished 10th grade at Phillips Exeter Academy in Exeter, NH, won a gold medal.
- Mitchell Lee from Oakton, VA, who finished 12th grade (homeschooled), won a gold medal.
- Bobby Shen, who finished 11th grade at Dulles High School in Sugar Land, TX, won a gold medal and placed third overall with a score of 39/42.
- Thomas Swayze, who finished 11th grade at Canyon Crest Academy in San Diego, CA, won a gold medal.
- David Yang, who finished 11th grade at Phillips Exeter Academy in Exeter, NH, won a gold medal.

2012 Carl B. Allendoerfer Awards

The Carl B. Allendoerfer Awards, established in 1976, are made to authors of expository articles published in *Mathematics Magazine*. The Awards are named for Carl B. Allendoerfer, a distinguished mathematician at the University of Washington and President of the Mathematical Association of America, 1959–60.

P. Mark Kayll, "Integrals Don't Have Anything to Do with Discrete Math, Do They?," *Mathematics Magazine*, **84:2** (2011), p. 108–119.

Mathematical work is often so highly specialized that mathematicians in one field can find it difficult to understand research in other areas. These divisions are frequently reflected in mathematics education where courses such as "Discrete Mathematics" suggest the compartmentalization of discrete and continuous topics. However, there are many examples of contemporary problems that combine disparate mathematical fields—such as algebra, geometry, topology, and combinatorics—or that connect mathematics to seemingly unrelated disciplines, such as biology. It is in this context that Mark Kayll enthusiastically reminds us that integrals do have something to do with discrete mathematics.

The beauty of Kayll's article lies in its exposition of some not-so-well known integral formulas for the number of perfect matchings in a graph. Deftly alternating between discrete and continuous topics, the author expresses the number of perfect matchings in a complete bipartite graph in terms of the gamma function. After this initial combination of the discrete and continuous, he expands his collection of improper integrals with the introduction of rook polynomials and derangements. This development of topics continues to a final full refutation of the author's title, a proof that the number of perfect matchings in a complete graph on n vertices is the nth moment of a standard normal random variable.

Kayll's well-written article provides us with engaging examples in which discrete and continuous mathematics come together. It reminds us that elements such as the gamma function are interesting in their own right, and it elegantly illustrates some of the ways in which continuous mathematics can be used to study discrete concepts. Enough details are included for the reader to follow the story, and comprehensive references are provided for those who want to learn more. Readers will finish the article with an increased appreciation of how surprising connections can exist between the discrete and the continuous, and of how the teaching of these subjects as distinct entities can be misleading to students.

Response from Mark Kayll. Thanks first to the citation author(s), whose kind words suggest that my goals in writing the article were in some measure achieved. We've all heard the phrase 'it takes a village,' and my experiences with this paper bear this out. The considered, constructive, and generous advice from both referees contributed substantively to improving the manuscript. A few colleagues, particularly Karel Stroethoff, also volunteered helpful and insightful input. *Mathematics Magazine* editor Walter Stromquist masterfully guided me in further polishing the piece. It was a privilege to work with these fine scholars—anonymous and otherwise—who selflessly shared their wisdom.

Math. Mag. **85** (2012) 318–320. doi:10.4169/math.mag.85.4.318.

Finally, thank you to the MAA for valuing expository writing and to the selection committee for finding the article worthy of merit. I feel both honored and delighted to receive the Allendoerfer Award, especially since the paper began as an experiment. (In a burst of eccentricity, I had even chosen a pseudonym, 'Kal M. Karply'.) To see it through to its final incarnation has been rewarding and humbling.

Biographical Note

Mark Kayll grew up in North Vancouver, British Columbia. After earning mathematics degrees from Simon Fraser University (B.Sc. 1987) and Rutgers University (Ph.D. 1994), he joined the faculty at the University of Montana in Missoula. He's enjoyed sabbaticals in Slovenia (University of Ljubljana, 2001–02) and Canada (Université de Montréal, 2008–09).

His publications fall in the discrete realm and have touched on combinatorics, graph theory, number theory, and probability. Mark's musical interests, such as playing the banjo, have motivated him in recent years to develop a general education course on mathematics and music for non-math majors.

He lives in Missoula with his wife, Jennifer (an excellent editor), and two beautiful children, Samuel and Leah.

John A. Adam, "Blood Vessel Branching: Beyond the Standard Calculus Problem," *Mathematics Magazine*, **84:3** (2011), p. 196–207.

What optimality principles determine the structure of the arteries, veins, and capillaries that comprise the human circulatory system? How reasonable are estimates that the total length of all the blood vessels within the body is on the order of 50,000 miles? This insightful and intellectually rich article offers an approach to these questions by studying the flow, branching, and maintenance of this important biological tree. To carry out his analysis, the author considers a sequence of increasingly sophisticated models. The first of these models employs only standard calculus tools to determine the optimal branching from a straight blood vessel to a nearby point. Later models use techniques from the calculus of variations to optimize a configuration for a combination of flow and volume.

The author first lays the groundwork by discussing the underlying biological setting, specifying his simplifying assumptions, and introducing some necessary equations from fluid dynamics. He then develops a sequence of models for blood vessel branching based upon a series of ever more comprehensive "cost functionals." The most sophisticated of these models implies certain empirical laws for vascular branching proposed by Wilhelm Roux in 1878, and also yields estimates for the total length of the vascular system.

This well-written article provides an excellent example of mathematical modeling in a context that is accessible and of obvious importance. It clearly shows the interaction of appropriate mathematical techniques with relevant scientific principles and illustrates the complexity of the modeling process. The reader is left with a deeper understanding of the power of mathematics to shed light on natural phenomena.

Response from John A. Adam. I am delighted and honored to receive the Allendoerfer Award. After undergoing open-heart surgery in 1996, it is perhaps not surprising that I started to develop an interest in the biophysics of the blood circulatory system! However, it was not until more than a decade later that I returned to review my old notes on the subject. In almost every calculus textbook these days there is an optimization problem about vascular branching, and as a result of devoting some class time to this topic, my appetite was further whetted to see if more sophisticated (and accurate) models of branching and bifurcation existed. Apart from a passing reference to empirical 'laws' of branching (proposed by Roux in 1878), cited by D'Arcy Thompson in his 1942 book *On Growth and Form* (Dover edition 1992), I could find only a brief but valuable summary by Rosen (1967). Consequently, I determined to try and reproduce the stated results, and re-develop them in the more pedagogic context of 'mathematical modeling'.

I must point out that the paper in its final form owes much to the valiant 'wordsmithing' efforts of past *Mathematics Magazine* editor, Frank Farris, and current editor, Walter Stromquist. The paper was in pre-press status during the transition from Frank's editorship to that of Walter, so it was very thoroughly vetted by each one! I am very grateful for their guidance and assistance.

Biographical Note

John Adam is Professor of Mathematics at Old Dominion University. He received his Ph.D. in theoretical astrophysics from the University of London in 1975. He is author of approximately 100 papers in several areas of applied mathematics and mathematical modeling. His first book, *Mathematics in Nature: Modeling Patterns in the Natural World*, was published in 2003 by Princeton University Press (paperback in 2006). He enjoys spending time with his family, especially his (thus far) five grandchildren, walking, nature photography, and is a frequent contributor to the *Earth Science Picture of the Day* (EPOD: http://epod.usra.edu/).

In 2007 he was a recipient of the State Council of Higher Education of Virginia's Outstanding Faculty Award. He coauthored *Guesstimation: Solving the World's Problems on the Back of a Cocktail Napkin*, published by Princeton University Press in 2008. More recently he has authored *A Mathematical Nature Walk* (2009, paperback version in 2011) and *X and the City: Modeling Aspects of Urban Life* (2012), both published by Princeton.

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

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS